

**Skript zur Vorlesung**

**Lineare Algebra 2 (BA/GYM)**

**(Sommersemester 2017)**

Dieses Geheft enthält in kompakter Form die wesentlichen Inhalte, wie sie in der Vorlesung „Lineare Algebra 2 (BA/GYM)“ vorgestellt werden.

Es ist zum Gebrauch neben der Vorlesung gedacht und erhebt nicht den Anspruch, „in sich selbst verständlich“ oder vollständig zu sein.

S. Hilger

# Inhaltsverzeichnis

<b>10 Matrixgruppen</b>	<b>5</b>
10.1 Reguläre Matrizen . . . . .	5
10.2 Berechnung von inversen Matrizen . . . . .	7
10.3 Untergruppen . . . . .	11
10.4 Gruppenhomomorphismen . . . . .	12
10.5 Matrixgruppen . . . . .	13
<b>11 Euklidische und unitäre Vektorräume</b>	<b>16</b>
11.1 Bilinearformen . . . . .	16
11.2 Sesquilinearformen . . . . .	18
11.3 Das Skalarprodukt . . . . .	20
11.4 Die euklidische Norm . . . . .	22
11.5 Orthogonalität . . . . .	25
11.6 Der unorientierte Winkel . . . . .	30
11.7 Das Vektorprodukt im euklidischen $\mathbb{R}^3$ . . . . .	34
11.8 Allgemeinere Normen . . . . .	36
<b>12 Basiswechsel</b>	<b>38</b>
12.1 Vektoren beim Basiswechsel . . . . .	38
12.2 Homomorphismen beim Basiswechsel . . . . .	41
12.3 Äquivalenz von Matrizen . . . . .	43
12.4 Beispielaufgaben: Orthogonale Polynome . . . . .	45
<b>13 Polynomringe</b>	<b>50</b>
13.1 Polynome . . . . .	50
13.2 Polynomfunktionen . . . . .	52
13.3 Teilbarkeit in Polynomringen . . . . .	53
13.4 Nullstellen und Linearfaktoren . . . . .	55
13.5 Polynome über $\mathbb{R}$ und $\mathbb{C}$ . . . . .	56
13.6 Das Minimalpolynom eines Ideals . . . . .	58
<b>14 Die Polynomialgebra einer quadratischen Matrix</b>	<b>60</b>
14.1 Einführung . . . . .	60
14.2 Das Minimalpolynom einer quadratischen Matrix . . . . .	61
14.3 Das charakteristische Polynom einer quadratischen Matrix . . . . .	64
14.4 Eigenwerte . . . . .	66
14.5 Der Satz von Cayley-Hamilton . . . . .	67
14.6 Ähnlichkeit von quadratischen Matrizen . . . . .	70
14.7 Direkte Summe von quadratischen Matrizen . . . . .	72
<b>15 Endomorphismen</b>	<b>74</b>
15.1 Endomorphismen beim Basiswechsel . . . . .	74
15.2 Invarianten von Endomorphismen . . . . .	78
15.3 Direkte Summe von Endomorphismen . . . . .	79

<b>16 Nilpotente Endomorphismen</b>	<b>80</b>
16.1 Young-Tableaus . . . . .	80
16.2 Nilpotente Endomorphismen . . . . .	82
16.3 Die Potenzkern-Kette . . . . .	84
16.4 Das Lemma von Fitting . . . . .	89
<b>17 Die Jordan-Normalform eines Endomorphismus</b>	<b>92</b>
17.1 Die $\lambda$ -Verschiebung eines Endomorphismus . . . . .	92
17.2 Die Hauptraumzerlegung . . . . .	94
17.3 Die Jordan-Normalform . . . . .	95
17.4 Algorithmus zur Transformation in Jordan-Normalform . . . . .	96
17.5 Beispiel . . . . .	98
17.6 Charakterisierung der Ähnlichkeitsklassen . . . . .	100
<b>18 Diagonalisierung</b>	<b>101</b>
18.1 Eigenvektoren . . . . .	101
18.2 Diagonalisierbarkeit . . . . .	102
18.3 Algorithmus zur Diagonalisierung . . . . .	105
18.4 Beispiele . . . . .	107
<b>19 Normale Endomorphismen</b>	<b>112</b>
19.1 Vorbemerkung / Vorbereitung . . . . .	112
19.2 Der adjungierte Homomorphismus . . . . .	114
19.3 Der adjungierte Endomorphismus . . . . .	116
19.4 Normale Endomorphismen . . . . .	117
19.5 Unitäre Endomorphismen . . . . .	119
19.6 Hermitesche und schief-hermitesche Endomorphismen . . . . .	123
<b>20 Endomorphismen im euklidischen Vektorraum</b>	<b>126</b>
20.1 Orthogonale Endomorphismen . . . . .	126
20.2 Orthogonale Endomorphismen im euklidischen $\mathbb{R}^2$ . . . . .	128
20.3 Normalform eines orthogonalen Endomorphismus . . . . .	132
20.4 Symmetrische und schiefsymmetrische Endomorphismen . . . . .	135
<b>21 Dualität</b>	<b>137</b>
21.1 Der duale Vektorraum . . . . .	137
21.2 Der duale Homomorphismus . . . . .	139
21.3 Der biduale Vektorraum . . . . .	142
<b>22 Affine Räume</b>	<b>144</b>
22.1 Einführung . . . . .	144
22.2 Affine Unterräume . . . . .	147
<b>23 Euklidische affine Räume</b>	<b>149</b>
23.1 Einführung . . . . .	149
23.2 Abstand zu einem Unterraum . . . . .	150
23.3 Hyperebenen und Normalenvektoren . . . . .	152
23.4 Isometrien . . . . .	153

23.5 Anwendungen in der Elementargeometrie . . . . .	154
<b>24 Nachtrag: Orientierung</b>	<b>157</b>
24.1 Orientierung eines $\mathbb{R}$ -Vektorraums . . . . .	158
24.2 Orientierung des $\mathbb{R}^n$ . . . . .	160
24.3 Orientierung eines euklidischen Raums . . . . .	161
24.4 Orientierung der zweidimensionalen „Zeichenebene“ . . . . .	165
24.5 Orientierung des dreidimensionalen „Anschauungsraums“ . . . . .	166

## 10 Matrixgruppen

### 10.1 Reguläre Matrizen

#### 10.1.1 Satz und Definition: Reguläre Matrizen

Für eine quadratische Matrix  $A \in \mathbb{K}^{n \times n}$  bzw. die zugehörige lineare Abbildung  $\ell_A : \mathbb{K}^n \rightarrow \mathbb{K}^n$  sind die folgenden Aussagen äquivalent:

- (A) Die Matrix heißt *regulär*.
- (B)  $A$  ist (multiplikativ) invertierbar, d.h. es existiert eine Matrix  $A^{-1}$  mit
 
$$A \cdot A^{-1} = I \quad \text{und} \quad A^{-1} \cdot A = I.$$
- (C) Die zugehörige lineare Abbildung  $\ell_A$  ist bijektiv.
- (D) Die zugehörige lineare Abbildung  $\ell_A$  ist injektiv.
- (E) Die zugehörige lineare Abbildung  $\ell_A$  ist surjektiv.
- (F) Der Rang von  $A$  ist gleich  $n$ .
- (G) Für beliebiges  $b \in \mathbb{K}^n$  ist das LGS  $A \cdot x = b$  eindeutig lösbar.
- (H) Die Matrix  $A$  lässt sich durch elementare Zeilenumformungen in die Einheitsmatrix  $I$  überführen.
- (J) Die Determinante von  $A$  ist ungleich Null.

Eine Matrix, die nicht regulär ist, heißt auch *singulär*.

#### 10.1.2 Beweis

- (1) Die Äquivalenz  $(B) \Leftrightarrow (C)$  ist offensichtlich. Man erinnere sich an Satz 1.7.14.
- (2) Für den Rang einer Matrix  $A$  gilt nach Satz und Definition 7.6.1.

$$\text{rang } A = \dim(\text{im } A) = n - \dim(\text{ker } A)$$

Daraus folgt direkt die Äquivalenz  $(D) \Leftrightarrow (E) \Leftrightarrow (F) \Leftrightarrow$  und damit auch die zu  $(C)$ .

- (3) Die Äquivalenz  $(B) \Leftrightarrow (G)$  ist ebenfalls direkt einsehbar.

$(F) \Rightarrow (H)$ : Gemäß Satz 2.3.2 (i) kann die Matrix  $A$  durch elementare Zeilenumformungen in eine Matrix  $\tilde{A}$  überführt werden, so dass

- die Matrix  $\tilde{A}$  eine obere Dreiecksmatrix ist,
- die Diagonaleinträge  $\tilde{a}_{jj}$  gleich Null oder Eins sind,
- falls ein Diagonaleintrag  $\tilde{a}_{jj} = 1$  ist, auch alle Einträge oberhalb dieses Eintrags gleich Null sind.

Wir wissen aus Satz 7.6.3 (ii), dass bei diesen elementaren Zeilenumformungen der Rang gleich bleibt, also

$$\text{rang } \tilde{A} = n.$$

Eine quadratische obere Dreiecksmatrix kann aber nur dann Vollrang  $n$  haben, wenn alle Diagonalelemente ungleich Null sind. Man kann also die Einheitsmatrix erreichen.

$(H) \Rightarrow (F)$ : Gemäß  $(D)$  lässt sich die erweiterte Matrix  $(A|b)$  durch elementare Zeilen-Äquivalenzumformungen in die erweiterte Matrix  $(I|\tilde{b})$  überführen.  $\tilde{b}$  ist dann die einzige Lösung des zugehörigen LGS  $I \cdot x = \tilde{b}$  und damit die einzige Lösung des ursprünglichen LGS.

$(G) \Rightarrow (C)$  ist direkt einsehbar.

$(B) \Rightarrow (J)$ : Wenn  $A$  invertierbar ist, so gilt

$$\det A \cdot \det(A^{-1}) = \det(A \cdot A^{-1}) = \det I = 1,$$

also muss  $\det A \neq 0$  sein.

$(J) \Rightarrow (B)$ : Gemäß Satz 9.5.2 gilt für die Kofaktor-Matrix

$$A^{\diamond T} \cdot A = \det A \cdot I.$$

Setzt man also

$$A^{-1} = \frac{1}{\det A} \cdot A^{\diamond T},$$

so haben wir die inverse Matrix gefunden.

Im Beweis sind — mehr oder weniger versteckt — einige Berechnungsmethoden aufgetreten, die wir noch einmal eigens beschreiben wollen.

## 10.2 Berechnung von inversen Matrizen

### 10.2.1 Satz: Inverse Matrix mittels Determinante

Die Inverse einer regulären Matrix  $A$  ist gegeben durch

$$\begin{aligned} A^{-1} &= \frac{1}{\det A} \cdot A^{\diamond T} \\ &= \frac{1}{\det A} \cdot \begin{pmatrix} + \det A^{(11)} & - \det A^{(21)} & \cdots & \cdots & (-1)^{n+1} \det A^{(n1)} \\ - \det A^{(21)} & + \det A^{(22)} & \cdots & \cdots & (-1)^{n+2} \det A^{(n2)} \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ (-1)^{1+n} \det A^{(1n)} & (-1)^{2+n} \det A^{(2n)} & \cdots & \cdots & + \det A^{(nn)} \end{pmatrix} \end{aligned}$$

Der Eintrag an der Position  $(j, k)$  der inversen Matrix ist

$$(A^{-1})_{jk} = (-1)^{j+k} \cdot \frac{\det A^{(kj)}}{\det A}.$$

Beachte die Vertauschung  $(jk) \rightarrow (kj)$  in dieser Formel.

### 10.2.2 Die Cramer'sche Regel

Ist die Matrix  $A \in \mathbb{K}^{n \times n}$  in dem LGS

$$Ax = b$$

regulär, so ist die eindeutige Lösung  $x$  gegeben durch

$$x_j = \frac{1}{\det A} \cdot \det \begin{pmatrix} a_{11} & \cdots & a_{1,j-1} & b_1 & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & b_n & a_{n,j+1} & \cdots & a_{nn} \end{pmatrix}, \quad j = 1, \dots, n.$$

**10.2.3 Beweis** Der Beweis besteht im Nachrechnen:

$$\frac{1}{\det A} \cdot \det \begin{pmatrix} a_{11} & \cdots & a_{1,j-1} & b_1 & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & b_n & a_{n,j+1} & \cdots & a_{nn} \end{pmatrix} =$$

(Entwicklung nach der  $j$ -ten Spalte)

$$\begin{aligned} &\frac{1}{\det A} \cdot \sum_{i=1}^n b_i \cdot (-1)^{i+j} \det A^{(ij)} = \\ &\frac{1}{\det A} \cdot \sum_{i=1}^n (A^{\diamond T})_{ji} b_i = \frac{1}{\det A} \cdot (A^{\diamond T} b)_j = (A^{-1} b)_j = x_j. \end{aligned}$$

### 10.2.4 Satz: Inverse Matrix mittels Gauß-Jordan-Algorithmus

- (i) Ist die Matrix  $A \in \mathbb{K}^{n \times n}$  regulär, so lässt sich in  $\mathbb{K}^{n \times 2n}$  durch elementare Zeilenumformungen die Transformation

$$(A|I) \longrightarrow (I|A^{-1})$$

durchführen.

- (ii) Die Zeilenumformungen entsprechen der Multiplikation mit  $A^{-1}$  von links.  
 (iii) Jede reguläre Matrix lässt sich als (endliches) Produkt von Matrizen

$$T_{jk}, Q_{j,\alpha}, R_{jk} \quad \text{oder sogar nur von} \quad Q_{j,\alpha}, R_{jk}$$

(vgl. Abschnitt 2.2.2) darstellen.

### 10.2.5 Beweis

- (1) Betrachte dazu das LGS mit  $n^2$  in der Matrix  $X$  zusammengefassten unbekanntem Zahlen:

$$A \cdot X = I$$

- (2) Gemäß Satz 10.1.1 (H) kann die Matrix  $A$  durch eine Abfolge geeigneter Zeilenumformungen in die Einheitsmatrix überführt werden.

- (3) Diese Abfolge von Zeilenumformungen entspricht der Multiplikation von  $A$  mit einer Matrix  $T$  von links. Multiplizieren wir das LGS mit  $T$  von links, so stellt sich heraus:

$$T \cdot A \cdot X = T$$

Wegen  $T \cdot A = I$  bedeutet dies aber  $X = T = A^{-1}$ .

- (4) Das bedeutet aber, dass sich die rechte Seite  $I$  in der erweiterten Matrix  $(A|I)$  durch Multiplikation mit  $T$ , das heißt durch die durchgeführten Zeilenumformungen in die Matrix  $A^{-1}$  verwandelt hat.

- (5) Zunächst ist, wie gerade bemerkt,  $A = T^{-1}$  eine Abfolge von Zeilenumformungen und deshalb ein Produkt von Matrizen  $T_{jk}, Q_{j,\alpha}, R_{jk}$ .

Bezeichnen wir die Vektoren der  $j$ -ten bzw.  $k$ -ten Zeile in  $A$  mit  $b$  bzw.  $c$ , so wirken sich die Linksmultiplikationen mit den angegebenen Matrizen wie folgt aus:

$$\begin{pmatrix} * \\ b \\ c \end{pmatrix} \xrightarrow{R_{kj}} \begin{pmatrix} * \\ b \\ c+b \end{pmatrix} \xrightarrow{Q_{k,-1}} \begin{pmatrix} * \\ b \\ -c-b \end{pmatrix} \xrightarrow{R_{jk}} \begin{pmatrix} * \\ -c \\ -c-b \end{pmatrix} \xrightarrow{Q_{j,-1}} \begin{pmatrix} * \\ c \\ -c-b \end{pmatrix} \xrightarrow{R_{kj}} \begin{pmatrix} * \\ c \\ -b \end{pmatrix} \xrightarrow{Q_{k,-1}} \begin{pmatrix} * \\ c \\ b \end{pmatrix}$$

Insgesamt wurden also die  $j$ -te und  $k$ -te Zeile getauscht. Es ist

$$T_{jk} = Q_{k,-1} \cdot R_{kj} \cdot Q_{j,-1} \cdot R_{jk} \cdot Q_{k,-1} \cdot R_{kj}.$$



### 10.2.6 Beispiel: Berechnung der Inversen mittels Determinante

Als Beispiel betrachten wir die Matrix

$$A = \begin{pmatrix} 5 & 2 & -1 \\ 0 & 1 & -4 \\ -3 & 4 & 2 \end{pmatrix}$$

Ihre Determinante ist gemäß Regel von Sarrus

$$\det A = 10 + 24 + 0 - 3 + 80 - 0 = 111.$$

Die Matrizen  $A^\diamond$ ,  $A^{\diamond T}$  und  $A^{-1}$  sind dann

$$A^\diamond = \begin{pmatrix} 18 & 12 & 3 \\ -8 & 7 & -26 \\ -7 & 20 & 5 \end{pmatrix}, \quad A^{\diamond T} = \begin{pmatrix} 18 & -8 & -7 \\ 12 & 7 & 20 \\ 3 & -26 & 5 \end{pmatrix},$$

$$A^{-1} = \begin{pmatrix} \frac{18}{111} & -\frac{8}{111} & -\frac{7}{111} \\ \frac{12}{111} & \frac{7}{111} & \frac{20}{111} \\ \frac{3}{111} & -\frac{26}{111} & \frac{5}{111} \end{pmatrix}.$$

Es empfiehlt sich, diese drei Schritte trotz erhöhten Schreibaufwands tatsächlich durchzuführen. Eine Abkürzung ist mit deutlich erhöhter Fehlergefahr verbunden.

### 10.2.7 Beispiel: Berechnung der Inversen mittels Gauß-Jordan-Algorithmus

5	2	-1	1	0	0	
0	1	-4	0	1	0	
-3	4	2	0	0	1	$+\frac{3}{5} \text{ I}$
5	2	-1	1	0	0	
0	1	-4	0	1	0	
0	$\frac{26}{5}$	$\frac{7}{5}$	$\frac{3}{5}$	0	1	$5 \cdot$
5	2	-1	1	0	0	
0	1	-4	0	1	0	
0	26	7	3	0	5	$-26 \text{ II}$
5	2	-1	1	0	0	
0	1	-4	0	1	0	
0	0	111	3	-26	5	$\frac{1}{111} \cdot$
5	2	-1	1	0	0	$+ \text{ III}$
0	1	-4	0	1	0	$+4 \text{ II}$
0	0	1	$\frac{3}{111}$	$-\frac{26}{111}$	$\frac{5}{111}$	
5	2	0	$\frac{114}{111}$	$-\frac{26}{111}$	$\frac{5}{111}$	$-2 \text{ II}$
0	1	0	$\frac{12}{111}$	$\frac{7}{111}$	$\frac{20}{111}$	
0	0	1	$\frac{3}{111}$	$-\frac{26}{111}$	$\frac{5}{111}$	
5	0	0	$\frac{90}{111}$	$-\frac{40}{111}$	$-\frac{35}{111}$	$\frac{1}{5} \cdot$
0	1	0	$\frac{12}{111}$	$\frac{7}{111}$	$\frac{20}{111}$	
0	0	1	$\frac{3}{111}$	$-\frac{26}{111}$	$\frac{5}{111}$	
1	0	0	$\frac{18}{111}$	$-\frac{8}{111}$	$-\frac{7}{111}$	
0	1	0	$\frac{12}{111}$	$\frac{7}{111}$	$\frac{20}{111}$	
0	0	1	$\frac{3}{111}$	$-\frac{26}{111}$	$\frac{5}{111}$	

## 10.3 Untergruppen

### 10.3.1 Definition: Untergruppe

Es sei  $(G, *)$  eine Gruppe. Die folgenden Aussagen über eine nicht-leere Teilmenge  $U \subseteq G$  sind äquivalent:

- (A) (Def)  $U$  heißt *Untergruppe* von  $G$ .
- (B)  $U$  bildet mit der Verknüpfung  $*$  eine Gruppe.
- (C) Für beliebige  $g, h \in U$  gilt  $g \cdot h \in U$  und für beliebige  $g \in U$  gilt  $g^{-1} \in U$ .
- (D) Für beliebige  $g, h \in U$  ist  $g \cdot h^{-1} \in U$ .

### 10.3.2 Beispiele

(1) Die folgenden Teilmengen von  $\mathbb{R}$  bilden (additive) Untergruppen der abelschen Gruppe  $(\mathbb{R}, +)$ .

$$\{0\}, \quad \mathbb{Z}, \quad h\mathbb{Z}, \quad \mathbb{Q}, \quad \left\{ \frac{m}{n^\ell} \in \mathbb{Q} \mid m \in \mathbb{Z}, \ell \in \mathbb{N}_0 \right\}, \quad \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}.$$

$h > 0$  und  $n \in \mathbb{N}$  sind dabei fixierte reelle Zahlen.

(2) Die folgenden Teilmengen von  $\mathbb{R} \setminus \{0\}$  bilden (multiplikative) Untergruppen der abelschen Gruppe  $(\mathbb{R} \setminus \{0\}, \cdot)$

$$\{1\}, \quad \{-1, +1\}, \quad \mathbb{R}^+, \quad \mathbb{Q} \setminus \{0\}, \quad \{q^\ell \in \mathbb{R}^+ \mid \ell \in \mathbb{Z}\}$$

$$\{a + b\sqrt{2} \in \mathbb{R} \setminus \{0\} \mid a, b \in \mathbb{Q}\}.$$

$q \in \mathbb{R}^+$  ist dabei eine fixierte positive Zahl.

(3) Die folgenden Teilmengen von  $\mathbb{C}^\times := \mathbb{C} \setminus \{0\}$  bilden (multiplikative) Untergruppen der abelschen Gruppe  $(\mathbb{C} \setminus \{0\}, \cdot)$

$$\mathbb{R}^\times := \mathbb{R} \setminus \{0\} \quad \text{und alle Untergruppen von } (\mathbb{R} \setminus \{0\}, \cdot)$$

$$E = \{z \in \mathbb{C} \mid |z| = 1\}$$

$$Z_n := \{z \in \mathbb{C} \mid z^n = 1\}$$

$n \in \mathbb{N}$  ist dabei eine fixierte natürliche Zahl.

(4) Die Teilmenge  $\mathcal{A}_n$  der geraden Permutationen bildet eine Untergruppe der Gruppe  $\mathcal{S}_n$  der Permutationen auf  $n$  Symbolen.

## 10.4 Gruppenhomomorphismen

### 10.4.1 Definition: Gruppenhomomorphismus

Es seien  $(G, *)$  und  $(H, \star)$  zwei Gruppen. Eine Abbildung  $\varphi : G \rightarrow H$  heißt *Gruppenhomomorphismus*, wenn für je zwei Elemente  $g, \tilde{g} \in G$  gilt:

$$\varphi(g * \tilde{g}) = \varphi(g) \star \varphi(\tilde{g}).$$

**10.4.2 Eigenschaften** Es sei  $\varphi : G \rightarrow H$  ein Gruppenhomomorphismus. Dann gilt

- (i)  $\varphi(e_G) = e_H$  (Neutrale Elemente)
- (ii)  $\varphi(g^{-1}) = (\varphi(g))^{-1}$  (Inverse Elemente)
- (iii) Ist  $\varphi$  ein bijektiver Gruppenhomomorphismus, so ist die Umkehrabbildung wieder ein Gruppenhomomorphismus.
- (iv) Der Kern des Gruppenhomomorphismus, definiert durch

$$\ker \varphi := \{g \in G \mid \varphi(g) = e_H\}$$

ist eine Untergruppe von  $G$ .

### 10.4.3 Beispiele

(1) Die Betragsbildung  $|\cdot| : \mathbb{C} \setminus \{0\} \rightarrow \mathbb{R}^+$  ist ein Gruppenhomomorphismus:

$$|z \cdot w| = |z| \cdot |w|$$

Sein Kern ist die Menge der komplexen Zahlen vom Betrag 1 (Einheitskreis in der Gaußschen Zahlenebene)

(2) Die Signums-Abbildung

$$\text{sgn} : \begin{cases} \mathbb{R}^\times & \rightarrow \{-1, +1\} \\ r & \mapsto \frac{r}{|r|} \end{cases}$$

ist ein Gruppenhomomorphismus. Sein Kern ist  $\mathbb{R}^+$ .

(3) Für festes  $a > 0$  ist die Exponentialabbildung zur Basis  $a$  (siehe Analysis)

$$\begin{cases} (\mathbb{R}, +) & \rightarrow (\mathbb{R}^+, \cdot) \\ x & \mapsto a^x \end{cases}$$

ein bijektiver Gruppenhomomorphismus. Der zugehörige Logarithmus, definiert als die zugehörige Umkehrabbildung,

$$\begin{cases} (\mathbb{R}^+, \cdot) & \rightarrow (\mathbb{R}, +) \\ x & \mapsto \log_a x \end{cases}$$

ist ebenfalls ein Gruppenhomomorphismus.

(4) Die Signumsabbildung

$$\sigma : \begin{cases} \mathcal{S}_n & \rightarrow \{-1, +1\} \\ \pi & \mapsto \sigma(\pi) \end{cases}$$

ist ein Gruppenhomomorphismus.

## 10.5 Matrixgruppen

### 10.5.1 Definition: Die allgemeine lineare Gruppe

Es sei

$$\mathbf{GL}(n, \mathbb{K}) = \{A \in \mathbb{K}^{n \times n} \mid A \text{ regulär}\}$$

die Teilmenge der regulären  $n \times n$ -Matrizen über dem Körper  $\mathbb{K}$ .

Sie bildet unter Matrixmultiplikation eine Gruppe. Das neutrale Element ist die Einheitsmatrix, das zu einer Matrix  $A$  inverse Element ist die inverse Matrix  $A^{-1}$ . Für  $n \geq 2$  ist die Gruppe nicht kommutativ.

Diese Gruppe heißt die *Allgemeine Lineare Gruppe* (**General Linear Group**) vom Grad  $n$  über  $\mathbb{K}$ .

### 10.5.2 Definition Matrixgruppe

Eine Gruppe  $G$  heißt „Matrixgruppe“ (= „Matrix-Liegruppe“), wenn

mit  $n \in \mathbb{N}$  und  $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$  die Gruppe  $G$  eine Untergruppe von  $\mathbf{GL}(n, \mathbb{K})$  ist

und dabei die folgende Bedingung erfüllt ist:

Ist  $(A_n)_{n \in \mathbb{N}}$  eine konvergente Folge von Matrizen in  $G$ , deren Grenzwert (definiert durch die Grenzwerte der Einträge) in  $\mathbf{GL}(n, \mathbb{K})$  enthalten ist, so ist der Grenzwert bereits in  $G$  enthalten.

Die zweite Bedingung kommt aus der Analysis, sie spielt in der Theorie der Liegruppen und in der Differentialgeometrie eine große Rolle. Sie ist nicht Thema der Linearen Algebra und ist hier nur der Vollständigkeit halber genannt.

### 10.5.3 Beispiele von Matrixgruppen

(1) Die Teilmenge der Diagonalmatrizen mit nicht-verschwindenden Diagonaleinträgen

$$\{A \in \mathbf{GL}(n, \mathbb{K}) \mid a_{jk} = 0 \Leftrightarrow j \neq k\}$$

bildet eine kommutative Matrixgruppe.

(2) Die Teilmenge der oberen Dreiecksmatrizen mit nicht-verschwindenden Diagonaleinträgen

$$\{A \in \mathbf{GL}(n, \mathbb{K}) \mid a_{jk} = 0 \text{ für } j > k, a_{jk} \neq 0 \text{ für } j = k\}$$

bildet eine Matrixgruppe.

### 10.5.4 Beispiele von reellen Matrixgruppen

(1) Die Determinante als Funktion  $\mathbf{GL}(n, \mathbb{R}) \rightarrow \mathbb{R}^\times$  ist ein Gruppenhomomorphismus:

$$\det(A \cdot B) = \det A \cdot \det B.$$

(2) Die Teilmenge der Matrizen mit positiver Determinante

$$\mathbf{GL}(n, \mathbb{R})^+ := \{A \in \mathbf{GL}(n, \mathbb{R}) \mid \det A > 0\}$$

bildet eine Matrixgruppe.

(3) Die Teilmenge der Matrizen mit Determinante vom Betrag 1

$$\{A \in \mathbf{GL}(n, \mathbb{R}) \mid \det A \in \{-1, +1\}\}$$

bildet eine Matrixgruppe.

(4) Die Teilmenge der Matrizen mit Determinante gleich 1

$$\mathbf{SL}(n, \mathbb{R}) := \{A \in \mathbf{GL}(n, \mathbb{R}) \mid \det A = 1\}$$

bildet eine Matrixgruppe. Sie heißt die *Spezielle Lineare Gruppe*.

(5) Die Teilmenge der orthogonalen Matrizen

$$\mathbf{O}(n) := \{A \in \mathbf{GL}(n, \mathbb{R}) \mid A^T = A^{-1}\}$$

bildet eine Matrixgruppe. Sie heißt die *Orthogonale Gruppe*.

(6) Die Teilmenge der orthogonalen Matrizen mit Determinante gleich 1

$$\mathbf{SO}(n) := \{A \in \mathbf{GL}(n, \mathbb{R}) \mid A^T = A^{-1}, \det A = 1\}$$

bildet eine Matrixgruppe. Sie heißt die *Spezielle Orthogonale Gruppe*, manchmal auch *Drehgruppe*.

### 10.5.5 Beispiele von komplexen Matrixgruppen

Bevor wir zu der Liste mit Beispielen schreiten, noch erst eine Definition:

Wir definieren zu einer komplexen Matrix die *hermitesch adjungierte* Matrix durch

$$A^\dagger := (\overline{A})^T.$$

Dabei ist  $\overline{A}$  die Matrix, die aus  $A$  durch Komplex-Konjugation aller Einträge hervorgeht.

(1) Die Determinante als Funktion  $\mathbf{GL}(n, \mathbb{C}) \rightarrow \mathbb{C}^\times$  ist ein Gruppenhomomorphismus:

$$\det(A \cdot B) = \det A \cdot \det B.$$

(2) Zunächst ist

$$\mathbf{GL}(n, \mathbb{R}) \subseteq \mathbf{GL}(n, \mathbb{C}).$$

Damit sind alle reellen Matrixgruppen auch komplexe Matrixgruppen.

(3) Die Teilmenge der Matrizen mit Determinante vom Betrag 1

$$\{A \in \mathbf{GL}(n, \mathbb{C}) \mid |\det A| = 1\}$$

bildet eine Matrixgruppe.

(4) Die Teilmenge der Matrizen mit Determinante gleich 1

$$\mathbf{SL}(n, \mathbb{C}) := \{A \in \mathbf{GL}(n, \mathbb{C}) \mid \det A = 1\}$$

bildet eine Matrixgruppe. Es ist die komplexe *Spezielle Lineare Gruppe* über  $\mathbb{C}$ .

(5) Die Teilmenge der Matrizen

$$\mathbf{U}(n) := \{A \in \mathbf{GL}(n, \mathbb{C}) \mid A^\dagger = A^{-1}\}$$

bildet eine Matrixgruppe. Sie heißt die *Unitäre Gruppe*.

(6) Die Teilmenge der unitären Matrizen mit Determinante gleich 1

$$\mathbf{SU}(n) := \{A \in \mathbf{GL}(n, \mathbb{C}) \mid A^\dagger = A^{-1}, \det A = 1\}$$

bildet eine Matrixgruppe. Sie heißt die *Spezielle Unitäre Gruppe*.

Die Gruppe  $\mathbf{SU}(2)$  spielt eine prominente Rolle in der Atomphysik, insbesondere in der Theorie des Wasserstoff-Atoms. Die sogenannten „irreduziblen Darstellungen“ der Gruppe  $\mathbf{SU}(3)$  bilden die mathematische Grundlage für die Theorie der Quarks.

# 11 Euklidische und unitäre Vektorräume

## 11.1 Bilinearformen

### 11.1.1 Definition

Es sei  $V$  ein  $\mathbb{K}$ -Vektorraum. Wir betrachten eine Abbildung

$$\langle \cdot, \cdot \rangle : \begin{cases} V \times V & \rightarrow \mathbb{K} \\ (v, w) & \mapsto \langle v, w \rangle \end{cases}$$

mit zwei Argumenten in  $V$  und Werten in  $\mathbb{K}$ . Es gibt auch andere Bezeichnungen anstelle von  $\langle \cdot, \cdot \rangle$ :

$$\langle v, w \rangle = \langle v|w \rangle = (v|w) = v \cdot w.$$

Die Version ganz rechts ist beim schulischen Skalarprodukt üblich, aber sehr ungünstig, weil sie die Gültigkeit eines Assoziativgesetzes suggeriert:  $(v \cdot w) \cdot u \stackrel{\text{falsch}}{=} v \cdot (w \cdot u)$ . Das ist aber Quatsch!

- Die Abbildung  $\langle \cdot, \cdot \rangle$  heißt *bilinear* oder eine *Bilinearform*, wenn sie in jedem einzelnen Argument linear ist, d.h. wenn für alle  $v, \tilde{v}, w, \tilde{w} \in V$  und  $\alpha, \tilde{\alpha} \in \mathbb{K}$  gilt

$$\begin{aligned} \langle \alpha v + \tilde{\alpha} \tilde{v}, w \rangle &= \alpha \langle v, w \rangle + \tilde{\alpha} \langle \tilde{v}, w \rangle \\ \langle v, \alpha w + \tilde{\alpha} \tilde{w} \rangle &= \alpha \langle v, w \rangle + \tilde{\alpha} \langle v, \tilde{w} \rangle. \end{aligned}$$

Anders ausgedrückt: Für festes  $v \in V$  sind die beiden Abbildungen

$$\langle \cdot, w \rangle : V \rightarrow \mathbb{K} \quad \text{und} \quad \langle v, \cdot \rangle : V \rightarrow \mathbb{K}$$

linear.

- Die Abbildung  $\langle \cdot, \cdot \rangle$  heißt *symmetrisch*, wenn für alle  $v, w \in V$  gilt:

$$\langle v, w \rangle = \langle w, v \rangle.$$

- Die Abbildung  $\langle \cdot, \cdot \rangle$  heißt *schief-symmetrisch* (= *antisymmetrisch* oder *alternierend*), wenn für alle  $v, w \in V$  gilt:

$$\langle v, w \rangle = -\langle w, v \rangle.$$

### 11.1.2 Standard-Beispiel

Es seien  $V = \mathbb{K}^n$  und  $B = \{e_1, \dots, e_n\}$  die Standard-Basis in  $\mathbb{K}^n$ .

Dann ist durch

$$\langle v, w \rangle = \left\langle \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \right\rangle := v_1 w_1 + v_2 w_2 + \dots + v_n w_n = v^T \cdot w$$

eine symmetrische Bilinearform definiert.



### 11.1.3 Matrizen induzieren Bilinearformen

Wir betrachten wieder den Vektorraum  $\mathbb{K}^n$  mit kanonischer Basis  $B = \{e_1, \dots, e_n\}$ .

Eine Matrix  $\mathbf{B} = (b_{jk}) \in \mathbb{K}^{n \times n}$  induziert dann vermöge

$$\langle v, w \rangle_{\mathbf{B}} := \sum_{j,k=1}^n v_j b_{jk} w_k = v^T \cdot \mathbf{B} \cdot w$$

eine Bilinearform auf  $\mathbb{K}^n$ .

Beachte, dass die Matrix  $\mathbf{B}$  hier in ihrer Bedeutung als quadratisches Zahlenschema auftritt, ihr aber keine lineare Abbildung  $\ell_{\mathbf{B}}$  zugeordnet ist, sondern eben diese Bilinearform.

Man kann sich leicht überlegen, dass die Bilinearform  $B$  genau dann symmetrisch bzw. schiefsymmetrisch ist, wenn das für die Matrix  $\mathbf{B}$  zutrifft.

### 11.1.4 Beispiel 1

Wählt man im  $\mathbb{K}^4$  die Einheitsmatrix,  $\mathbf{B} = I$ , so erhält man die Bilinearform aus Abschnitt 11.1.2

$$\begin{aligned} \left\langle \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix}, \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix} \right\rangle_{\mathbf{B}} &= (v_1 \ v_2 \ v_3 \ v_4) \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix} \\ &= v_1 w_1 + v_2 w_2 + v_3 w_3 + v_4 w_4. \end{aligned}$$

### 11.1.5 Beispiel 2

Um ein konkretes Beispiel für den  $\mathbb{R}^3$  anzugeben, sei

$$\mathbf{B} = \begin{pmatrix} 4 & -2 & 3 \\ \frac{5}{2} & 9 & 0 \\ -5 & 3,5 & \frac{4}{3} \end{pmatrix}$$

eine x-beliebige reelle  $3 \times 3$ - Matrix. Die zugehörige Bilinearform ist dann

$$\begin{aligned} &\left\langle \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}, \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \right\rangle_{\mathbf{B}} \\ &= (v_1 \ v_2 \ v_3) \cdot \begin{pmatrix} 4 & -2 & 3 \\ \frac{5}{2} & 9 & 0 \\ -5 & 3,5 & \frac{4}{3} \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \\ &= 4v_1 w_1 - 2v_1 w_2 + 3v_1 w_3 + \frac{5}{2}v_2 w_1 + 9v_2 w_2 - 5v_3 w_1 + 3,5v_3 w_2 + \frac{4}{3}v_3 w_3. \end{aligned}$$

## 11.2 Sesquilinearformen

### 11.2.1 Definition

Es sei  $V$  ein  $\mathbb{C}$ -Vektorraum. Wir betrachten wieder eine Abbildung

$$\langle \cdot, \cdot \rangle : \begin{cases} V \times V & \rightarrow \mathbb{C} \\ (v, w) & \mapsto \langle v, w \rangle \end{cases}$$

mit zwei Argumenten in  $V$  und Werten in  $\mathbb{C}$ .

- Die Abbildung  $\langle \cdot, \cdot \rangle$  heißt *sesquilinear* ( $= \frac{3}{2}$ -linear) oder eine *Sesquilinearform*, wenn sie im ersten Argument linear und im zweiten Argument *semilinear* ( $= \frac{1}{2}$ -linear) ist, d.h. wenn für alle  $v, \tilde{v}, w, \tilde{w} \in V$  und  $\alpha, \tilde{\alpha} \in \mathbb{C}$  gilt

$$\begin{aligned} \langle \alpha v + \tilde{\alpha} \tilde{v}, w \rangle &= \alpha \langle v, w \rangle + \tilde{\alpha} \langle \tilde{v}, w \rangle \\ \langle v, \alpha w + \tilde{\alpha} \tilde{w} \rangle &= \bar{\alpha} \langle v, w \rangle + \bar{\tilde{\alpha}} \langle v, \tilde{w} \rangle. \end{aligned}$$

- Die Abbildung  $\langle \cdot, \cdot \rangle$  heißt *hermitesch*, wenn sie sesquilinear ist und für alle  $v, w \in V$  gilt:

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

- Die Abbildung  $\langle \cdot, \cdot \rangle$  heißt *schiefhermitesch*, wenn für alle  $v, w \in V$  gilt:

$$\langle v, w \rangle = -\overline{\langle w, v \rangle}.$$

### 11.2.2 Standard-Beispiel

Es seien  $V = \mathbb{C}^n$  und  $B = \{e_1, \dots, e_n\}$  die Standard-Basis.

Dann ist durch

$$\langle v, w \rangle = \left\langle \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \right\rangle := v_1 \bar{w}_1 + v_2 \bar{w}_2 + \dots + v_n \bar{w}_n = v^T \cdot \bar{w}$$

eine hermitesche Sesquilinearform definiert.

### 11.2.3 Matrizen induzieren Sesquilinearformen

Wir betrachten den Vektorraum  $\mathbb{C}^n$  mit kanonischer Basis  $B = \{e_1, \dots, e_n\}$ .

Eine Matrix  $\mathbf{B} = (b_{jk}) \in \mathbb{C}^{n \times n}$  induziert dann vermöge

$$\langle v, w \rangle_{\mathbf{B}} := \sum_{j,k=1}^n v_j b_{jk} \overline{w_k} = v^T \cdot \mathbf{B} \cdot \overline{w}$$

eine Sesquilinearform auf  $\mathbb{C}^n$ .

Man kann sich wieder leicht überlegen, dass die Sesquilinearform  $B$  genau dann hermitesch bzw. schiefhermitesch ist, wenn das für die Matrix  $\mathbf{B}$  zutrifft. Vgl. auch Abschnitt 10.5.5.

$$\begin{aligned} \mathbf{B} \text{ hermitesch} &\iff \mathbf{B}^\dagger := \overline{\mathbf{B}}^T = \mathbf{B} \\ \mathbf{B} \text{ schiefhermitesch} &\iff \mathbf{B}^\dagger := \overline{\mathbf{B}}^T = -\mathbf{B}. \end{aligned}$$

### 11.2.4 Beispiel 1

Wählt man im  $\mathbb{C}^2$  die Einheitsmatrix,  $\mathbf{B} = I$ , so erhält man die Bilinearform aus Abschnitt 11.1.2

$$\begin{aligned} \left\langle \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \right\rangle_{\mathbf{B}} &= \begin{pmatrix} v_1 & v_2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \overline{w_1} \\ \overline{w_2} \end{pmatrix} \\ &= v_1 \overline{w_1} + v_2 \overline{w_2}. \end{aligned}$$

### 11.2.5 Beispiel 2

Um ein konkretes Beispiel für den  $\mathbb{C}^3$  anzugeben, sei

$$\mathbf{B} = \begin{pmatrix} 2 & -2+i & 3-i \\ 1-\frac{1}{2}i & 7i & 0 \\ -5 & 3,5+2i & \frac{4}{3}-\frac{1}{4}i \end{pmatrix}$$

eine x-beliebige komplexe  $3 \times 3$ - Matrix. Die zugehörige Bilinearform ist dann

$$\begin{aligned} &\left\langle \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}, \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \right\rangle_{\mathbf{B}} \\ &= \begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix} \cdot \begin{pmatrix} 2 & -2+i & 3-i \\ 1-\frac{1}{2}i & 7i & 0 \\ -5 & 3,5+2i & \frac{4}{3}-\frac{1}{4}i \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \\ &= 2v_1w_1 + (-2+i)v_1w_2 + (3-i)v_1w_3 + (1-\frac{1}{2}i)v_2w_1 \\ &\quad + 7iv_2w_2 - 5v_3w_1 + (3,5+2i)v_3w_2 + (\frac{4}{3}-\frac{1}{4}i)v_3w_3. \end{aligned}$$

## 11.3 Das Skalarprodukt

Es sei ab jetzt

- $V$  ein  $\mathbb{R}$ -Vektorraum und  $\langle \cdot, \cdot \rangle$  eine Bilinearform  $V \times V \rightarrow \mathbb{R}$  oder
- $V$  ein  $\mathbb{C}$ -Vektorraum und  $\langle \cdot, \cdot \rangle$  eine Sesquilinearform  $V \times V \rightarrow \mathbb{C}$ .

### 11.3.1 Definitionen

- Die Abbildung  $\langle \cdot, \cdot \rangle$  heißt *positiv definit*, wenn für alle  $v \in V \setminus \{\vec{0}\}$  gilt:

$$\langle v, v \rangle > 0.$$

- Die Abbildung  $\langle \cdot, \cdot \rangle$  heißt ein *Skalarprodukt* oder ein *Inneres Produkt*, wenn sie die folgenden drei Eigenschaften hat:
  - im Fall  $\mathbb{K} = \mathbb{R}$ : bilinear, symmetrisch und positiv definit
  - im Fall  $\mathbb{K} = \mathbb{C}$ : sesquilinear, hermitesch und positiv definit
- Ein endlich-dimensionaler Vektorraum  $V$ , auf dem ein Skalarprodukt  $\langle \cdot, \cdot \rangle$  definiert ist, heißt ...
  - im Fall  $\mathbb{K} = \mathbb{R}$ : ein *euklidischer* oder *Euklid'scher Vektorraum*.
  - im Fall  $\mathbb{K} = \mathbb{C}$ : ein *unitärer Vektorraum*.

Man schreibt dann auch  $(V, \langle \cdot, \cdot \rangle)$ .

### 11.3.2 Standard-Skalarprodukt auf $\mathbb{R}^n$

Wir betrachten auf dem Vektorraum  $V = \mathbb{R}^n$  die Bilinearform aus Beispiel 11.1.2. Für beliebiges  $v \in \mathbb{R}^n \setminus \{\vec{0}\}$  gilt

$$\langle v, v \rangle = \left\langle \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right\rangle := v_1^2 + v_2^2 + \dots + v_n^2 > 0$$

Also ist diese Bilinearform positiv definit. Da sie auch symmetrisch ist, ist sie ein Skalarprodukt.

### 11.3.3 Standard-Skalarprodukt auf $\mathbb{C}^n$

Wir betrachten auf dem Vektorraum  $V = \mathbb{C}^n$  die Bilinearform aus Beispiel 11.2.2. Für beliebiges  $v \in \mathbb{C}^n \setminus \{\vec{0}\}$  gilt

$$\begin{aligned} \langle v, v \rangle &= \left\langle \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right\rangle = v_1 \cdot \overline{v_1} + v_2 \cdot \overline{v_2} + \dots + v_n \cdot \overline{v_n} \\ &= |v_1|^2 + |v_2|^2 + \dots + |v_n|^2 > 0 \end{aligned}$$

Also ist diese Sesquilinearform positiv definit. Da sie auch hermitesch ist, ist sie ein Skalarprodukt.

### 11.3.4 Beispiel (Staatsexamen GS/HS/RS: F00 T2, A3a)

Aufgabenstellung: Auf dem  $\mathbb{R}^3$  sei die Bilinearform

$$(x, y) := x_1y_1 + 3x_2y_2 + 4x_3y_3 + x_1y_2 + x_2y_1 + x_1y_3 + x_3y_1 + x_2y_3 + x_3y_2$$

gegeben. Zeigen Sie, dass diese Bilinearform ein inneres Produkt auf dem  $\mathbb{R}^3$  definiert.

Lösung: Bei einer Vertauschung von  $x$  und  $y$  gilt:

$$(y, x) := y_1x_1 + 3y_2x_2 + 4y_3x_3 + y_1x_2 + y_2x_1 + y_1x_3 + y_3x_1 + y_2x_3 + y_3x_2$$

Dieser Ausdruck ist offenbar mit dem obigen identisch, also ist die Abbildung symmetrisch.

Die Linearität im ersten Argument ist einfach nachzurechnen:

$$\begin{aligned} (\alpha x, y) &= \\ \alpha x_1y_1 + 3\alpha x_2y_2 + 4\alpha x_3y_3 + \alpha x_1y_2 + \alpha x_2y_1 + \alpha x_1y_3 + \alpha x_3y_1 + \alpha x_2y_3 + \alpha x_3y_2 &= \\ \alpha(x, y) \end{aligned}$$

$$\begin{aligned} (x + z, y) &= \\ (x_1 + z_1)y_1 + 3(x_2 + z_2)y_2 + 4(x_3 + z_3)y_3 + (x_1 + z_1)y_2 + (x_2 + z_2)y_1 + \\ &\quad (x_1 + z_1)y_3 + (x_3 + z_3)y_1 + (x_2 + z_2)y_3 + (x_3 + z_3)y_2 = \\ x_1y_1 + z_1y_1 + 3x_2y_2 + 3z_2y_2 + 4x_3y_3 + 4z_3y_3 + x_1y_2 + z_1y_2 + x_2y_1 + z_2y_1 + \\ &\quad x_1y_3 + z_1y_3 + x_3y_1 + z_3y_1 + x_2y_3 + z_2y_3 + x_3y_2 + z_3y_2 = \\ (x, y) + (z, y). \end{aligned}$$

Bezüglich der Linearität im zweiten Argument genügt es, auf die bereits bewiesene Symmetrie zu verweisen.

Zum Schluss muss die Positiv-Definitheit gezeigt werden:

$$\begin{aligned} (x, x) &= x_1x_1 + 3x_2x_2 + 4x_3x_3 + x_1x_2 + x_2x_1 + x_1x_3 + x_3x_1 + x_2x_3 + x_3x_2 \\ &= (x_1 + x_2 + x_3)^2 + 2x_2^2 + 3x_3^2. \end{aligned}$$

Dieser Ausdruck ist niemals negativ. Ist er gleich Null, so kann man in der folgenden Reihenfolge schließen:

$$x_2 = 0 \quad \text{und} \quad x_3 = 0 \quad \implies \quad x_1 = 0.$$

## 11.4 Die euklidische Norm

### 11.4.1 Definitionen

Es sei  $(V, \langle \cdot, \cdot \rangle)$  ein euklidischer ( $\mathbb{K} = \mathbb{R}$ ) oder unitärer ( $\mathbb{K} = \mathbb{C}$ ) Vektorraum.

- Die Abbildung

$$\|\cdot\| : \begin{cases} V & \rightarrow \mathbb{R}_0^+ \\ v & \mapsto \|v\| := \sqrt{\langle v, v \rangle} \end{cases}$$

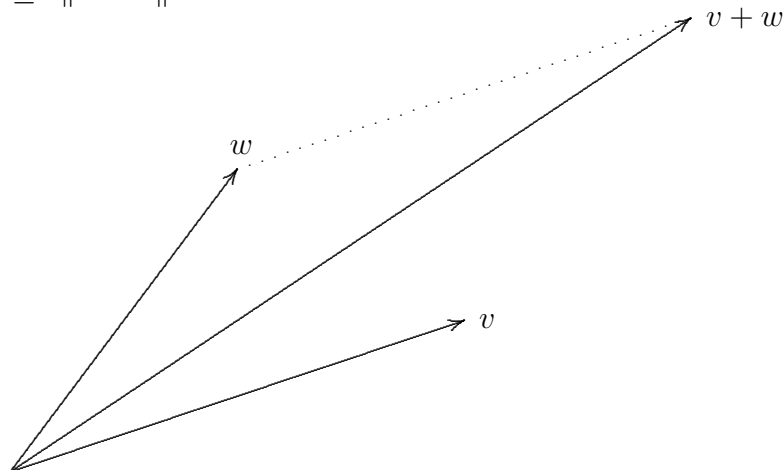
heißt (*euklidische*) Norm oder (*euklidische*) Länge auf  $(V, \langle \cdot, \cdot \rangle)$ .

- Ein Vektor  $v \in V$  heißt *normiert* oder *Einheitsvektor*, wenn er „Einheitslänge“ hat:  $\|v\| = 1$ .

### 11.4.2 Satz: Eigenschaften der euklidischen Norm

Für  $v, w \in V$  und  $\alpha \in \mathbb{K}$  sind die folgenden Eigenschaften erfüllt.

- (i)  $\|v\| = 0 \implies v = \vec{0}$ .
- (ii)  $\|\alpha \cdot v\| = |\alpha| \cdot \|v\|$
- (iii)  $|\langle v, w \rangle| \leq \|v\| \cdot \|w\|$  (Ungleichung von Cauchy–Schwarz)
- (iv) Gleichheit tritt in (iii) genau dann ein, wenn  $v$  und  $w$  linear abhängig sind.
- (v)  $\|v + w\| \leq \|v\| + \|w\|$  (Dreiecksungleichung)
- (vi)  $\left| \|v\| - \|w\| \right| \leq \|v - w\|$ .



### 11.4.3 Beweis

(1) Wir führen den Beweis nur für den unitären Fall ( $\mathbb{K} = \mathbb{C}$ ) durch.

Der euklidische Fall  $\mathbb{K} = \mathbb{R}$  ist einfacher insofern, als die auftretenden komplexen Operationen auf  $\mathbb{R}$  trivial sind. Für  $\alpha \in \mathbb{R} \subseteq \mathbb{C}$  gilt nämlich

$$\operatorname{Re} \alpha = \alpha \quad \bar{\alpha} = \alpha \quad \alpha \cdot \bar{\alpha} = |\alpha|^2 = \alpha^2$$

(2) Die folgende Gleichheit für  $v, w \in V$  und  $\alpha, \beta \in \mathbb{C}$  werden wir zweimal benötigen:

$$\begin{aligned} \|\alpha v + \beta w\|^2 &= \langle \alpha v + \beta w, \alpha v + \beta w \rangle \\ &= \langle \alpha v, \alpha v \rangle + \langle \alpha v, \beta w \rangle + \langle \beta w, \alpha v \rangle + \langle \beta w, \beta w \rangle \\ &= \alpha \bar{\alpha} \langle v, v \rangle + \alpha \bar{\beta} \langle v, w \rangle + \beta \bar{\alpha} \langle w, v \rangle + \beta \bar{\beta} \langle w, w \rangle \\ &= |\alpha|^2 \|v\|^2 + 2 \operatorname{Re}(\alpha \bar{\beta} \langle v, w \rangle) + |\beta|^2 \|w\|^2. \end{aligned}$$

(3) Die Aussage (i) folgt direkt aus der Positiv-Definitheit des Skalarprodukts.

(4) Es ist

$$\begin{aligned} \|\alpha \cdot v\| &= \sqrt{\langle \alpha \cdot v, \alpha \cdot v \rangle} = \sqrt{\alpha \cdot \langle v, \alpha \cdot v \rangle} = \sqrt{\alpha \cdot \bar{\alpha} \cdot \langle v, v \rangle} \\ &= \sqrt{\alpha \cdot \bar{\alpha}} \cdot \sqrt{\langle v, v \rangle} = |\alpha| \cdot \|v\|. \end{aligned}$$

und damit ist (ii) gezeigt.

(5) Für  $w = \vec{0}$  ist die Aussage (iii) trivial, so dass wir  $\vec{w} \neq \vec{0}$  annehmen können.

(6) Mit  $\beta := -\frac{\langle v, w \rangle}{\|w\|^2}$  gilt wegen (1):

$$\begin{aligned} 0 &\leq \|v + \beta w\|^2 = \|v\|^2 + 2 \operatorname{Re}(\bar{\beta} \cdot \langle v, w \rangle) + |\beta|^2 \|w\|^2 \\ &= \|v\|^2 + 2 \operatorname{Re}\left(-\frac{\overline{\langle v, w \rangle}}{\|w\|^2} \cdot \langle v, w \rangle\right) + \left|\frac{\langle v, w \rangle}{\|w\|^2}\right|^2 \|w\|^2 \\ &= \|v\|^2 - 2 \frac{|\langle v, w \rangle|^2}{\|w\|^2} + \frac{|\langle v, w \rangle|^2}{\|w\|^2} \\ &= \|v\|^2 - \frac{|\langle v, w \rangle|^2}{\|w\|^2} \end{aligned}$$

Die Multiplikation dieser Ungleichung mit  $\|w\|^2 > 0$  liefert

$$0 \leq \|v\|^2 \cdot \|w\|^2 - |\langle v, w \rangle|^2$$

und dann

$$|\langle v, w \rangle|^2 \leq \|v\|^2 \cdot \|w\|^2.$$

Zieht man daraus die Wurzel, so hat man die Ungleichung (iii).

(7) Eine Durchsicht des Schrittes (6) zeigt, dass Gleichheit in (iii) genau dann eintritt, wenn in der ersten Zeile Gleichheit gilt, wenn also

$$\vec{0} = v + \beta w = v - \frac{\langle v, w \rangle}{\|w\|^2} w.$$

Aus der Gleichheit folgt die lineare Abhängigkeit der Vektoren  $v$  und  $w$ .

Sind umgekehrt  $v$  und  $w$  linear abhängig,  $v = \gamma w$ , so ist das obige

$$\beta = -\frac{\langle v, w \rangle}{\langle w, w \rangle} = -\frac{\gamma \langle w, w \rangle}{\langle w, w \rangle} = -\gamma,$$

also  $v + \beta w = \vec{0}$  in Schritt (6).

(8) Mit (1) gilt

$$\begin{aligned} \|v + w\|^2 &= \|v\|^2 + 2 \operatorname{Re}(\langle v, w \rangle) + \|w\|^2 \\ &\quad (\operatorname{Re} z \leq |z| \text{ für alle } z \in \mathbb{C}) \\ &\leq \|v\|^2 + 2|\langle v, w \rangle| + \|w\|^2 \\ &\quad (\text{Cauchy-Schwarz-Ungleichung}) \\ &\leq \|v\|^2 + 2\|v\| \cdot \|w\| + \|w\|^2 \\ &= (\|v\| + \|w\|)^2. \end{aligned}$$

Wurzelziehen ergibt die Dreiecksungleichung (v).

(9) Die Ungleichung (vi) wird in der Übung bewiesen.



## 11.5 Orthogonalität

### 11.5.1 Definitionen: Orthogonalität und Orthonormalität

Es sei  $(V, \langle \cdot, \cdot \rangle)$  wieder ein euklidischer oder ein unitärer Vektorraum.

- Zwei einzelne Vektoren  $u, w \in V$  heißen *orthogonal (zueinander)*, wenn  $\langle u, w \rangle = 0$ .

In diesem Fall schreibt man auch  $u \perp w$ .

- Zwei Teilmengen  $U, W \subseteq V$  heißen *orthogonal (zueinander)*, wenn gilt:

$$\langle u, w \rangle = 0 \quad \text{für alle } u \in U, w \in W.$$

- Ist  $W \subseteq V$  eine Teilmenge von  $V$ , so heißt die Menge

$$W^\perp := \{v \in V \mid \langle v, w \rangle = 0 \text{ für alle } w \in W\}$$

das *orthogonale Komplement* von  $W$ .

- Eine Teilmenge  $W \subseteq V$  heißt *orthonormiert* oder *orthonormal*, wenn die Vektoren in  $W$  normiert und paarweise orthogonal sind:

$$\begin{aligned} \|w\| &= 1 \quad \text{für alle } w \in W, \\ \langle w, \tilde{w} \rangle &= 0 \quad \text{für alle } w, \tilde{w} \in W \text{ mit } w \neq \tilde{w}. \end{aligned}$$

- Eine Basis  $B$  von  $V$  heißt *Orthonormalbasis (ONB)*, wenn sie orthonormiert ist.

### 11.5.2 Satz: Erste Eigenschaften

(i) Sind

$$\begin{aligned} v &= \alpha_1 w_1 + \dots + \alpha_n w_n \\ \tilde{v} &= \beta_1 w_1 + \dots + \beta_n w_n \end{aligned}$$

die Darstellungen zweier Vektoren bzgl. einer ONB, so ist

$$\begin{aligned} \langle v, \tilde{v} \rangle &= \alpha_1 \overline{\beta_1} + \dots + \alpha_n \overline{\beta_n} \\ \|v\| &= \sqrt{|\alpha_1|^2 + \dots + |\alpha_n|^2}. \end{aligned}$$

(ii) Eine orthonormale Menge  $W = \{w_1, w_2, \dots, w_m\}$  ist linear unabhängig.

(iii) Sind  $U, W$  orthogonal zueinander, so gilt

$$U \cap W \subseteq \{\vec{0}\}.$$

### 11.5.3 Beweis

(i) Wir rechnen einfach nach:

$$\begin{aligned} \langle v, \tilde{v} \rangle &= \langle \alpha_1 w_1 + \dots + \alpha_n w_n, \beta_1 w_1 + \dots + \beta_n w_n \rangle \\ &= \alpha_1 \overline{\beta_1} \langle w_1, w_1 \rangle + \alpha_1 \overline{\beta_2} \langle w_1, w_2 \rangle + \dots + \alpha_1 \overline{\beta_n} \langle w_1, w_n \rangle + \\ &\quad \vdots \\ &\quad \alpha_n \overline{\beta_1} \langle w_n, w_1 \rangle + \alpha_n \overline{\beta_2} \langle w_n, w_2 \rangle + \dots + \alpha_n \overline{\beta_n} \langle w_n, w_n \rangle \\ &= \alpha_1 \overline{\beta_1} + \dots + \alpha_n \overline{\beta_n} \end{aligned}$$

Die zweite Gleichung folgt direkt aus der ersten.

(ii) Wir stellen den Nullvektor als Linearkombination von  $\{w_1, w_2, \dots, w_m\}$  dar

$$\alpha_1 w_1 + \alpha_2 w_2 + \dots + \alpha_m w_m = \vec{0} \quad (*)$$

Es sei jetzt  $j \in \{1, \dots, m\}$  beliebig. Wir bilden auf beiden Seiten der Gleichung (\*) das Skalarprodukt mit  $w_j$ .

$$\alpha_1 \langle w_1, w_j \rangle + \alpha_2 \langle w_2, w_j \rangle + \dots + \alpha_m \langle w_m, w_j \rangle = \langle \vec{0}, w_j \rangle = 0.$$

Da die Vektoren  $w_i$  paarweise orthogonal sind, gilt  $\langle w_k, w_j \rangle = 0$  für alle  $k \neq j$ . Daraus folgt

$$\alpha_j \langle w_j, w_j \rangle = 0 \implies \alpha_j = 0.$$

$j$  war beliebig, es ist also  $\alpha_j = 0$  für alle  $j$ . Also ist  $W$  linear unabhängig.

(iii) Ist  $v \in U \cap W$ , so gilt aufgrund der Definition von zueinander orthogonalen Teilmengen  $\langle v, v \rangle = 0$  und damit  $v = \vec{0}$ .

### 11.5.4 Satz: Orthonormalisierung

Es sei  $(V, \langle \cdot, \cdot \rangle)$  ein euklidischer oder unitärer Vektorraum.

(i) (Orthonormalisierungsverfahren von Erhard Schmidt)

Ist die Teilmenge  $W \subseteq V$  endlich und linear unabhängig, so kann man eine gleichmächtige orthonormale Teilmenge  $U \subseteq V$  finden, so dass

$$\text{span } U = \text{span } W.$$

(ii) Es sei  $\dim V = n$ . Jede orthonormale Teilmenge  $W = \{w_1, w_2, \dots, w_m\} \subseteq V$  kann durch eine Menge  $W' = \{w_{m+1}, \dots, w_n\}$  von  $n - m$  Vektoren zu einer ONB von  $V$  ergänzt werden. Es gilt dann:

$$(\text{span } W)^\perp = \text{span } W' \quad \text{und} \quad \text{span } W \oplus \text{span } W' = V.$$

(iii) Jeder euklidische oder unitäre Vektorraum  $V$  endlicher Dimension hat eine ONB.

### 11.5.5 Beweis

Wir setzen  $m = |W|$  und  $W = \{w_1, \dots, w_m\}$ .

(i) Der Fall  $m = 1$  ist einfach: Man setzt

$$u_1 = \frac{w_1}{\|w_1\|}.$$

Wir nehmen jetzt als Induktionsvoraussetzung an, dass bereits eine orthonormale Menge  $\{u_1, \dots, u_k\}$ ,  $k < m$ , mit

$$\text{span}\{u_1, \dots, u_k\} = \text{span}\{w_1, \dots, w_k\}$$

konstruiert ist. Wir setzen jetzt

$$\tilde{u}_{k+1} := w_{k+1} - \langle w_{k+1}, u_1 \rangle u_1 - \dots - \langle w_{k+1}, u_k \rangle u_k.$$

Damit ist für  $\ell = 1, \dots, k$

$$\begin{aligned} \langle \tilde{u}_{k+1}, u_\ell \rangle &= \langle w_{k+1} - \langle w_{k+1}, u_1 \rangle u_1 - \dots - \langle w_{k+1}, u_k \rangle u_k, u_\ell \rangle \\ &= \langle w_{k+1}, u_\ell \rangle - \langle w_{k+1}, u_1 \rangle \cdot \langle u_1, u_\ell \rangle - \dots - \langle w_{k+1}, u_k \rangle \cdot \langle u_k, u_\ell \rangle \\ &= \langle w_{k+1}, u_\ell \rangle - \langle w_{k+1}, u_\ell \rangle = 0. \end{aligned}$$

Es ist also  $\tilde{u}_{k+1}$  orthogonal zu  $\{u_1, \dots, u_k\}$ . Setzen wir jetzt noch

$$u_{k+1} := \frac{\tilde{u}_{k+1}}{\|\tilde{u}_{k+1}\|},$$

so ist  $\{u_1, \dots, u_{k+1}\}$  orthonormal.

Aufgrund der Induktionsvoraussetzung ist  $\text{span}\{u_1, \dots, u_k\} = \text{span}\{w_1, \dots, w_k\}$  und deshalb

$$u_{k+1} \in \text{span}\{w_1, \dots, w_{k+1}\}.$$

Dann ist aber

$$\text{span}\{u_1, \dots, u_{k+1}\} \subseteq \text{span}\{w_1, \dots, w_{k+1}\}$$

Nach Satz 11.5.2 (i) ist die linke Menge linear unabhängig, die rechte ist es aufgrund der Voraussetzung. Das impliziert aber

$$\text{span}\{u_1, \dots, u_{k+1}\} = \text{span}\{w_1, \dots, w_{k+1}\}.$$

(ii) Man ergänze zunächst gemäß Basisergänzungssatz 5.7.1 die Menge  $W$  zu einer Basis von  $V$  und wende dann das Verfahren aus (i) auf diese Basis von  $V$  an. Die beiden Aussagen über  $W$  und  $W'$  sind leicht einzusehen.

(iii) ist in (ii) enthalten.

### 11.5.6 Algorithmus

Bei der konkreten Durchführung des Verfahrens zur Ersetzung einer Basis  $\{w_1, \dots, w_m\}$  durch eine ONB  $\{u_1, \dots, u_m\}$  „genügt“ es, rekursiv für  $(k = 1, \dots, m)$  die Vektoren

$$\begin{aligned} u_1 &= \frac{w_1}{\|w_1\|} \\ u_2 &= \frac{w_2 - \langle w_2, u_1 \rangle u_1}{\|w_2 - \langle w_2, u_1 \rangle u_1\|} \\ u_3 &= \frac{w_3 - \langle w_3, u_1 \rangle u_1 - \langle w_3, u_2 \rangle u_2}{\|w_3 - \langle w_3, u_1 \rangle u_1 - \langle w_3, u_2 \rangle u_2\|} \\ &\vdots \\ u_m &= \frac{w_m - \sum_{k=1}^{m-1} \langle w_m, u_k \rangle u_k}{\|w_m - \sum_{k=1}^{m-1} \langle w_m, u_k \rangle u_k\|} \end{aligned}$$

anzugeben. Die Orthonormalität von  $\{u_1, \dots, u_m\}$  ist durch den abstrakten Beweis sichergestellt.

Bei der konkreten Durchführung ohne elektronische Unterstützung empfiehlt es sich, die Berechnungen in viele kleine Teilschritte zu zerlegen. Das folgende Beispiel zeigt dies auf.

### 11.5.7 Aufgabe: Staatsexamen F99 T1 A2

Es seien

$$w_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \quad w_2 = \begin{pmatrix} 0 \\ -1 \\ 2 \\ 0 \end{pmatrix}$$

Vektoren des euklidischen  $\mathbb{R}^4$  und  $U$  der von ihnen aufgespannte Unterraum. Bestimmen Sie eine Orthonormalbasis für das orthogonale Komplement  $U^\perp \subset \mathbb{R}^4$  von  $U$ .

Im Original heißen die Vektoren  $a_1 = w_1$  und  $a_2 = w_2$

Im Unterschied zum obigen Algorithmus sind Basis-Vektoren  $w_3, w_4$ , die das Komplement  $U^\perp$  erzeugen, noch nicht vorgegeben.

Wir werden in der folgenden (zu ausführlichen) Lösung eine ONB  $\{u_1, u_2, u_3, u_4\}$  für den  $\mathbb{R}^4$  erstellen, so dass  $\text{span}\{u_1, u_2\} = U$  und demzufolge  $\text{span}\{u_3, u_4\} = U^\perp$  ist.

### 11.5.8 Lösung: Staatsexamen F99 T1 A2

( $k = 1$ )

$$\begin{aligned} \tilde{u}_1 &= w_1 \\ \|\tilde{u}_1\|^2 &= 3 \\ \|\tilde{u}_1\| &= \sqrt{3} \\ u_1 &= \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \end{aligned}$$

$(k = 2)$ 

$$\begin{aligned} \langle w_2, u_1 \rangle &= \frac{1}{\sqrt{3}} \left\langle \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right\rangle = -\frac{1}{\sqrt{3}} \\ \langle w_2, u_1 \rangle u_1 &= -\frac{1}{\sqrt{3}} \cdot \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} = -\frac{1}{3} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \\ \tilde{u}_2 &= w_2 - \langle w_2, u_1 \rangle u_1 = \begin{pmatrix} 0 \\ -1 \\ 2 \\ 0 \end{pmatrix} + \frac{1}{3} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 \\ -2 \\ 6 \\ 1 \end{pmatrix} \\ \|\tilde{u}_2\|^2 &= \frac{1}{9}(1 + 4 + 36 + 1) = \frac{14}{3} \\ \|\tilde{u}_2\| &= \sqrt{\frac{14}{3}} \\ u_2 &= \frac{\tilde{u}_2}{\|\tilde{u}_2\|} = \sqrt{\frac{3}{14}} \cdot \frac{1}{3} \begin{pmatrix} 1 \\ -2 \\ 6 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{42}} \begin{pmatrix} 1 \\ -2 \\ 6 \\ 1 \end{pmatrix} \end{aligned}$$

$(k = 3)$  Man muss sich jetzt einen Vektor  $w_3$  selbst suchen. Günstig ist einer, der auf  $u_1, u_2$  schon senkrecht steht. Daher wählt man beispielsweise

$$w_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}$$

Dieser Vektor erfüllt tatsächlich  $\langle w_3, u_1 \rangle = \langle w_3, u_2 \rangle = 0$ , er steht also schon auf  $U$  senkrecht. Setze dann

$$\begin{aligned} \tilde{u}_3 &= w_3 \\ \|\tilde{u}_3\|^2 &= 2 \\ \|\tilde{u}_3\| &= \sqrt{2} \\ u_3 &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix} \end{aligned}$$

 $(k = 4)$  Wähle jetzt

$$w_4 = \begin{pmatrix} 0 \\ 3 \\ 1 \\ 0 \end{pmatrix}$$

Es gilt dann  $w_4 \notin \text{span}\{u_1, u_2, u_3\}$ , da

$$\begin{aligned} \langle w_4, u_2 \rangle &= 0 \\ \langle w_4, u_3 \rangle &= 0 \\ \text{span}\{w_4\} &\neq \text{span}\{u_1\}. \end{aligned}$$

Weiter ist

$$\begin{aligned} \langle w_4, u_1 \rangle &= \frac{1}{\sqrt{3}} \left\langle \begin{pmatrix} 0 \\ 3 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right\rangle = \sqrt{3} \\ \langle w_4, u_1 \rangle u_1 &= \sqrt{3} \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} \\ \tilde{u}_4 &= w_4 - \langle w_4, u_1 \rangle u_1 = \begin{pmatrix} 0 \\ 3 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ 1 \\ -1 \end{pmatrix} \\ \|\tilde{u}_4\|^2 &= 7 \\ \|\tilde{u}_4\| &= \sqrt{7} \\ u_4 &= \frac{1}{\sqrt{7}} \begin{pmatrix} -1 \\ 2 \\ 1 \\ -1 \end{pmatrix} \end{aligned}$$

Insgesamt haben wir die ONB

$$\{u_1, u_2, u_3, u_4\} \quad \text{mit} \quad U = \text{span}\{u_1, u_2\} \quad U^\perp = \text{span}\{u_3, u_4\}$$

gefunden.

## 11.6 Der unorientierte Winkel

### 11.6.1 Die Kosinusfunktion

In der Analysis werden die trigonometrischen Funktionen Kosinus und Sinus

$$\begin{aligned} \cos : \begin{cases} \mathbb{R} & \rightarrow [-1, +1] \\ \alpha & \mapsto \cos \alpha = 1 - \frac{\alpha^2}{2!} + \frac{\alpha^4}{4!} \pm \dots \end{cases} \\ \sin : \begin{cases} \mathbb{R} & \rightarrow [-1, +1] \\ \alpha & \mapsto \sin \alpha = \alpha - \frac{\alpha^3}{3!} + \frac{\alpha^5}{5!} \pm \dots \end{cases} \end{aligned}$$

über Potenzreihen definiert.

Daraus lässt sich dann das Kapitel einer Formelsammlung über trigonometrische Funktionen herleiten. Im Mittelpunkt stehen dabei die so genannten Additionstheoreme. Für beliebige  $\alpha, \beta \in \mathbb{R}$  gilt:

$$\begin{aligned} \cos(\alpha - \beta) &= \cos \alpha \cos \beta + \sin \alpha \sin \beta \\ \cos(\alpha + \beta) &= \cos \alpha \cos \beta - \sin \alpha \sin \beta \\ \sin(\alpha + \beta) &= \sin \alpha \cos \beta + \cos \alpha \sin \beta \\ \sin(\alpha - \beta) &= \sin \alpha \cos \beta - \cos \alpha \sin \beta \end{aligned}$$

### 11.6.2 Die Arcuscosinus-Funktion

Aus der Analysis ist bekannt, dass die Einschränkung des Kosinus auf das Intervall  $[0, \pi]$

$$\cos|_{[0, \pi]} : \begin{cases} [0, \pi] & \rightarrow [-1, +1] \\ \alpha & \mapsto \cos \alpha \end{cases}$$

bijektiv ist. Nach Satz 3.7.9 existiert die Umkehrfunktion

$$\arccos : \begin{cases} [-1, +1] & \rightarrow [0, \pi] \\ r & \mapsto \arccos r. \end{cases}$$

Es gibt viele andere Intervalle in  $\mathbb{R}$ , so dass die Einschränkung des Kosinus darauf bijektiv wird. Da die Definition des arccos auf dieser Einschränkung beruht, gibt es viele „Arcus-Kosi-Nüsse“.

### 11.6.3 Die Euklidische Ebene

Es sei auf  $V = \mathbb{R}^2$  das durch

$$\left\langle \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \right\rangle := v_1 w_1 + v_2 w_2$$

definierte Standard-Skalarprodukt gegeben. Wir bezeichnen diesen Vektorraum als *euklidische Ebene*. Als ONB wählen wir die Basis  $B = \{e_1, e_2\}$  der kanonischen Einheitsvektoren.

### 11.6.4 Definition: Der unorientierte Winkel

Für zwei Vektoren  $v, w \in V \setminus \{\vec{0}\}$  hat die reelle Zahl

$$c(v, w) := \frac{\langle v, w \rangle}{\|v\| \cdot \|w\|}$$

aufgrund der Cauchy-Schwarz-Ungleichung einen Betrag

$$|c(v, w)| = \left| \frac{\langle v, w \rangle}{\|v\| \cdot \|w\|} \right| \leq \left| \frac{\|v\| \cdot \|w\|}{\|v\| \cdot \|w\|} \right| = 1.$$

Diese Betragsabschätzung für eine reelle Zahl kann in eine Doppelungleichung übersetzt werden:

$$-1 \leq c(v, w) \leq +1.$$

Das aber bedeutet, dass sie in den Arcuscosinus aus Abschnitt 11.6.2 eingesetzt werden kann.

Dies mündet in die Definition des *unorientierten Winkels* zwischen zwei Vektoren  $v, w$  in der euklidischen Ebene

$$\angle(v, w) := \arccos c(v, w) = \arccos \frac{\langle v, w \rangle}{\|v\| \cdot \|w\|}.$$

### 11.6.5 Satz: Der unorientierte Winkel

Es seien  $u, v, w \in V \setminus \{\vec{0}\}$  drei Vektoren der euklidischen Ebene,  $\lambda \in \mathbb{R} \setminus \{0\}$ . Dann gilt:

- (i)  $\angle(v, v) = 0$
- (ii)  $\angle(v, w) = \frac{\pi}{2} \iff \langle v, w \rangle = 0$
- (iii)  $\angle(v, w) = \angle(w, v)$
- (iv)  $\angle(\lambda v, w) = \angle(v, \lambda w) = \begin{cases} \angle(v, w), & \text{falls } \lambda > 0 \\ \pi - \angle(v, w), & \text{falls } \lambda < 0 \end{cases}$
- (v) (nur unter geeigneten Zusatzvoraussetzungen)

$$\angle(v, w) = \angle(v, u) + \angle(u, w)$$

### 11.6.6 Bemerkungen

(1) Damit erfüllt der unorientierte Winkel fast genau die Eigenschaften, die aus der Winkelmessung der Alltagsgeometrie vertraut sind. Als Winkelskala tritt hier lediglich das so genannte Bogenmaß anstelle der klassischen Gradskala  $[0^\circ, 180^\circ]$  in Erscheinung. Diese Vorgehensweise ist — mathematisch — viel natürlicher, leider in Alltags- und Schulkontexten vermeintlich unanschaulich.

(2) Wir haben mit Hilfe des Skalarprodukts und der trigonometrischen Funktionen die Winkelmessung eingeführt. Historisch und schulisch war/ist es gerade anders herum: Auf der Grundlage der Winkelmessung werden die trigonometrischen Funktionen — und evtl. das Skalarprodukt — eingeführt.

(3) Die reichlich seltsame Bedingung in (v) lässt sich mathematisch befriedigend auflösen, wenn man

für die ONB  $\{e_1, e_2\}$  eine so genannte Orientierung einführt und die Willkürlichkeit bei der Definition der Arcus-Kosinus-Funktion mit Hilfe einer Äquivalenzrelation  $\text{mod } 2\pi$  auf der Bogenmaßskala  $\mathbb{R}$  beseitigt.

### 11.6.7 Beweis

Mit

$$\angle(v, v) = \arccos \frac{\langle v, v \rangle}{\|v\| \cdot \|v\|} = \arccos 1 = 0$$

ist (i) bewiesen.

Durch Zerlegung in Einzelschritte sehen wir die Äquivalenz (ii) ein:

$$\angle(v, w) = \frac{\pi}{2} \iff \arccos \frac{\langle v, w \rangle}{\|v\| \cdot \|w\|} = \frac{\pi}{2} \iff \frac{\langle v, w \rangle}{\|v\| \cdot \|w\|} = 0 \iff \langle v, w \rangle = 0$$

Die Aussage (iii) folgt sofort aus der Symmetrie des Skalarprodukts.

Weiter ist mit

$$\begin{aligned} \angle(\lambda v, w) &= \arccos \frac{\langle \lambda v, w \rangle}{\|\lambda v\| \cdot \|w\|} = \arccos \frac{\lambda \langle v, w \rangle}{|\lambda| \|v\| \cdot \|w\|} \\ &= \begin{cases} \arccos \frac{\langle v, w \rangle}{\|v\| \cdot \|w\|} = \angle(v, w), & \text{falls } \lambda > 0, \\ \arccos \frac{-\langle v, w \rangle}{\|v\| \cdot \|w\|} = \pi - \arccos \frac{\langle v, w \rangle}{\|v\| \cdot \|w\|} = \pi - \angle(v, w), & \text{falls } \lambda < 0. \end{cases} \end{aligned}$$

die Aussage (iv) gezeigt.

Zum Beweis von (v). Mit (iv) folgt für alle  $v, w \in V \setminus \{\vec{0}\}$

$$\angle(v, w) = \angle\left(\frac{v}{\|v\|}, \frac{w}{\|w\|}\right),$$

so dass wir in (v) annehmen können, dass

$$\|v\|^2 = v_1^2 + v_2^2 = 1, \quad \|w\|^2 = w_1^2 + w_2^2 = 1, \quad \|u\|^2 = u_1^2 + u_2^2 = 1.$$

Daraus folgt, dass es Zahlen  $\alpha, \beta, \gamma \in [0, \pi]$  mit

$$v = \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix}, \quad w = \begin{pmatrix} \cos \beta \\ \sin \beta \end{pmatrix}, \quad u = \begin{pmatrix} \cos \gamma \\ \sin \gamma \end{pmatrix}.$$

Es ist dann weiter

$$\begin{aligned} \langle v, u \rangle &= v_1 u_1 + v_2 u_2 = \cos \alpha \cos \gamma + \sin \alpha \sin \gamma = \cos(\alpha - \gamma) \\ \langle u, w \rangle &= u_1 w_1 + u_2 w_2 = \cos \gamma \cos \beta + \sin \gamma \sin \beta = \cos(\gamma - \beta) \\ \langle v, w \rangle &= v_1 w_1 + v_2 w_2 = \cos \alpha \cos \beta + \sin \alpha \sin \beta = \cos(\alpha - \beta) \end{aligned}$$



Ist jetzt die geometrische Situation so eingerichtet, dass „unsere“ Arcus-Kosinus-Funktion gerade die Werte

$$\arccos \langle v, u \rangle = \alpha - \gamma$$

$$\arccos \langle u, w \rangle = \gamma - \beta$$

$$\arccos \langle v, w \rangle = \alpha - \beta$$

herausfischt, so gilt

$$\angle(v, w) = \alpha - \beta = (\alpha - \gamma) + (\gamma - \beta) = \angle(v, u) + \angle(u, w).$$

## 11.7 Das Vektorprodukt im euklidischen $\mathbb{R}^3$

Wir betrachten den euklidischen Vektorraum  $V = \mathbb{R}^3$  mit der kanonischen Basis  $\{e_1, e_2, e_3\}$ , die bezüglich des kanonischen Skalarprodukts eine ONB ist.

### 11.7.1 Definition

Wir definieren auf  $V$  eine zweistellige Verknüpfung, das *Vektorprodukt*

$$\times : \begin{cases} V \times V & \rightarrow V \\ (v, w) & \mapsto v \times w \end{cases}$$

durch

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

Anders als beim Skalarprodukt ist das Ergebnis wieder ein Vektor. Beachte, dass das Vektorprodukt nicht assoziativ ist:

$$\begin{aligned} (e_1 \times e_2) \times e_2 &= e_3 \times e_2 = -e_1 \\ e_1 \times (e_2 \times e_2) &= e_1 \times \vec{0} = \vec{0} \end{aligned}$$

### 11.7.2 Satz: Eigenschaften des Vektorprodukts

(i) Das Vektorprodukt ist bilinear, d.h. für alle  $v, \tilde{v}, w, \tilde{w} \in V$  und  $\alpha, \tilde{\alpha} \in \mathbb{R}$  gilt

$$\begin{aligned} (\alpha v + \beta \tilde{v}) \times w &= \alpha v \times w + \beta \tilde{v} \times w \\ v \times (\alpha w + \beta \tilde{w}) &= \alpha v \times w + \beta v \times \tilde{w}. \end{aligned}$$

(ii) Das Vektorprodukt ist schiefssymmetrisch, d.h. für alle  $v, w \in V$  gilt

$$w \times v = -v \times w.$$

(iii) Für  $u, v, w \in V$  ist das so genannte *Spatprodukt* gegeben durch

$$\langle u \times v, w \rangle = \det(u, v, w).$$

Es kann — bei geeigneten weiteren Definitionen — als Volumen des von den Vektoren  $u, v, w$  aufgespannten *Parallelepipeds* angesehen werden.

(iv) Für  $v, w \in V$  gilt:

$$\|v \times w\|^2 = \|v\|^2 \|w\|^2 - \langle v, w \rangle^2.$$

(v)  $v$  und  $w$  sind linear abhängig genau dann, wenn  $v \times w = \vec{0}$ .

### 11.7.3 Beweis

(i) – (iv) sind einfach nachzurechnen. (v) zeigt man mit Hilfe von (iv) und der Cauchy–Schwarz–Ungleichung.

### 11.7.4 Bemerkung

Aus der Gleichung (iv) und mit Blick auf Abschnitt 11.6.4 folgt, dass

$$\begin{aligned}\|v \times w\|^2 &= \|v\|^2 \|w\|^2 - \langle v, w \rangle^2 = \|v\|^2 \|w\|^2 \cdot [1 - \cos^2 \angle(v, w)] \\ &= \|v\|^2 \|w\|^2 \cdot \sin^2 \angle(v, w),\end{aligned}$$

und damit weiter

$$\|v \times w\| = \|v\| \|w\| \cdot |\sin \angle(v, w)|.$$

### 11.7.5 Frage und Antwort

Warum gibt es ein Vektorprodukt nur bei Dimension 3 ?

Es sei  $V$  ein  $\mathbb{K}$ -Vektorraum der Dimension  $n \in \mathbb{N}$ . Man kann ihm weitere Vektorräume

$$\bigwedge^p V \quad \text{mit} \quad \dim(\bigwedge^p V) = \binom{n}{p}$$

zuordnen. Die Elemente heißen  $p$ -Vektoren.

Es ist  $\bigwedge^1 V = V$ . Wenn auf  $V$  ein Skalarprodukt und eine Orientierung definiert sind, so lässt sich ein so genannter *Hodge-Isomorphismus*

$$\bigwedge^p V \simeq \bigwedge^{n-p} V$$

herstellen.

Ein  $p$ -Vektor und ein  $q$ -Vektor lassen sich multiplizieren, Grundlage ist das so genannte (bilineare, assoziative, aber nicht kommutative) *Keilprodukt*

$$\wedge : \begin{cases} \bigwedge^p V \times \bigwedge^q V & \rightarrow \bigwedge^{p+q} V \\ (x, y) & \mapsto x \wedge y. \end{cases}$$

Nur im Fall  $n = 3$  resultiert daraus eine Verknüpfung auf  $V$

$$\wedge : \begin{cases} V \times V & \rightarrow \bigwedge^2 V \simeq V \\ (x, y) & \mapsto x \wedge y, \end{cases}$$

eben das Vektorprodukt.

## 11.8 Allgemeinere Normen

### 11.8.1 Definition

Eine Abbildung  $\|\cdot\|$  auf einem beliebigen (evtl. nicht mit einem Skalarprodukt versehenen)  $\mathbb{R}$ -Vektorraum oder  $\mathbb{C}$ -Vektorraum heißt (allgemeine) *Norm*, wenn sie die folgenden Eigenschaften hat.

Für  $v, w \in V$  und  $\alpha \in \mathbb{K}$  gilt:

Positiv-Definitheit: Für  $v \in V$  gilt:  $\|v\| = 0 \implies v = \vec{0}$ .

Homogenität: Für  $v \in V, \alpha \in \mathbb{K}$  gilt:  $\|\alpha \cdot v\| = |\alpha| \cdot \|v\|$

Dreiecksungleichung: Für  $v, w \in V$  gilt:  $\|v + w\| \leq \|v\| + \|w\|$

Das sind gerade die Eigenschaften (i),(ii) und (v) aus Satz 11.4.2.

### 11.8.2 Beispiele

Die beiden Abbildungen

$$\|\cdot\|_1 : \begin{cases} \mathbb{R}^n \rightarrow \mathbb{R}_0^+ \\ \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \mapsto |v_1| + \dots + |v_n| \end{cases}$$

$$\|\cdot\|_\infty : \begin{cases} \mathbb{R}^n \rightarrow \mathbb{R}_0^+ \\ \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \mapsto \max\{|v_1|, \dots, |v_n|\} \end{cases}$$

sind Normen, die im Fall  $n \geq 2$  nicht von Skalarprodukten induziert werden.

### 11.8.3 Definition: Metrik

Es sei  $M$  eine beliebige Menge. Eine Vektorraumstruktur, ein Skalarprodukt oder eine Norm auf  $M$  sind nicht nötig. Eine Abbildung

$$d : \begin{cases} M \times M \rightarrow \mathbb{R}_0^+ \\ (x, y) \mapsto d(x, y) \end{cases}$$

heißt *Metrik* auf  $M$ , wenn sie die folgenden drei Eigenschaften erfüllt:

Positiv-Definitheit: Für alle  $x, y \in M$  gilt  $d(x, y) = 0 \iff x = y$ .

Symmetrie: Für alle  $x, y \in M$  gilt  $d(x, y) = d(y, x)$ .

Dreiecksungleichung: Für alle  $x, y, z \in M$  gilt  $d(x, y) \leq d(x, z) + d(z, y)$ .

Man kann die Zahl  $d(x, y)$  als Abstand zwischen  $x$  und  $y$  ansehen. Man beachte aber, dass die Eigenschaften viel allgemeiner gehalten sind als die durch die Norm bzw. das Skalarprodukt auf einem Vektorraum nahegelegten.

Wir werden in der Linearen Algebra diesen Begriff nicht weiter einbeziehen. Der Begriff der Metrik ist nichtsdestoweniger ein fundamentaler Begriff des fundamentalen Teilgebiets der Topologie und durchzieht deshalb verschiedenste Teilgebiete der Mathematik.

#### 11.8.4 Norm induziert Metrik

Ist auf einem euklidischen oder unitären Vektorraum  $V$  eine Norm  $\|\cdot\|$  definiert, so kann man mit ihrer Hilfe eine Abbildung

$$d : \begin{cases} V \times V & \rightarrow \mathbb{R}_0^+ \\ (v, w) & \mapsto d(v, w) := \|v - w\| \end{cases}$$

definieren. Sie hat dann die Eigenschaften einer Metrik wie gerade beschrieben.

## 12 Basiswechsel

### 12.1 Vektoren beim Basiswechsel

#### 12.1.1 Definition: Geordnete Basis

Ist  $V$  ein  $n$ -dimensionaler Vektorraum, so bezeichnen wir ein  $n$ -Tupel

$$B = (v_1, \dots, v_n) \in V^n$$

als *geordnete Basis*, wenn die Menge

$$\{v_1, \dots, v_n\} \in \mathcal{P}(V)$$

eine Basis (im bisherigen Sinne) ist. Die zusätzliche Eigenschaft „geordnet“ bedeutet also, dass die Basisvektoren ein-eindeutig mit  $1, \dots, n$  durchnummeriert sind — und nicht nur in einer Menge zusammengefasst sind.

#### 12.1.2 Definition: Basis-Entwicklung eines Vektors

Ist für einen  $n$ -dimensionalen Vektorraum  $V$  über dem Körper  $\mathbb{K}$  eine geordnete Basis  $B_V = (v_1, \dots, v_n)$  vorgegeben, so kann damit eine bijektive lineare Abbildung

$$\eta_B : \begin{cases} V \rightarrow \mathbb{K}^n \\ v \mapsto v_B = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \end{cases}$$

dadurch definiert werden, dass der Vektor  $(\alpha_1, \dots, \alpha_n)^T$  gerade die eindeutige Basisentwicklung von  $v$  bzgl.  $B$  ist:

$$v = \alpha_1 v_1 + \dots + \alpha_n v_n.$$

Wir nennen  $v_B = \eta_B(v)$  die *B-Entwicklung* ( $\eta \sim \text{eta} \sim \mathbf{E}$ ntwicklung) oder *B-Darstellung* (von  $v$  bzgl. der geordneten Basis  $B$ ).

Wie man sofort sehen kann, handelt es sich bei  $\eta_B$  um die nach Satz 7.2.1 eindeutig bestimmte Abbildung, die die Basisvektoren  $v_i \in V$  auf die kanonischen Einheitsvektoren  $e_i \in \mathbb{K}^n$  abbildet:

$$\eta_B(v_i) = e_i.$$

#### 12.1.3 Quintessenz

Alles, was man in  $V$  oder mit  $V$  machen kann, kann man stattdessen auch in  $\mathbb{K}^n$  oder mit dem  $\mathbb{K}^n$  machen. Der Vorteil besteht darin, dass man im  $\mathbb{K}^n$ , also mit Zahlen aus  $\mathbb{K}$ , rechnen kann. Von Nachteil ist, dass diese Berechnungen immer von der aktuell festgesetzten geordneten Basis abhängen.

### 12.1.4 Vektoren beim Basiswechsel

Wir betrachten jetzt für einen fixierten Vektorraum  $V$  zwei geordnete Basen  $B = \{v_1, \dots, v_n\}$  und  $\tilde{B} = \{\tilde{v}_1, \dots, \tilde{v}_n\}$  mit den zugehörigen Abbildungen  $\eta_B$  und  $\eta_{\tilde{B}}$



In den Diagrammen ist erkennbar, dass die bijektive lineare Abbildung  $\Theta_{\tilde{B}B}$  definiert ist durch

$$\Theta_{\tilde{B}B} := \eta_{\tilde{B}} \circ \eta_B^{-1} : \mathbb{K}^n \rightarrow \mathbb{K}^n.$$

Sie kann als invertierbare  $n \times n$ -Matrix aufgefasst werden. Sie heißt *Übergangsmatrix* (*transition matrix*) für den Basiswechsel  $\tilde{B} \leftarrow B$ . ( $\Theta \sim$  Theta  $\sim$  Transition)

### 12.1.5 Satz über Übergangsmatrizen

- (i) Für zwei geordnete Basen  $B, \tilde{B}$  sind die beiden Übergangsmatrizen invers zueinander:

$$\Theta_{\tilde{B}B} \cdot \Theta_{B\tilde{B}} = I.$$

- (ii) Für drei geordnete Basen  $B, \tilde{B}, \hat{B}$  gilt:

$$\Theta_{\hat{B}B} = \Theta_{\hat{B}\tilde{B}} \cdot \Theta_{\tilde{B}B}.$$

- (iii) Hat ein Vektor  $v \in V$  bzgl. zweier geordneter Basen  $B$  und  $\tilde{B}$  die Entwicklungen  $v_B$  bzw.  $v_{\tilde{B}}$ , so gilt

$$v_{\tilde{B}} = \Theta_{\tilde{B}B} v_B \quad \text{bzw.} \quad v_B = \Theta_{\tilde{B}B}^{-1} v_{\tilde{B}}.$$

- (iv) In der  $j$ -ten Spalte von

$$\Theta_{\tilde{B}B} = \begin{pmatrix} \vartheta_{11} & \cdots & \vartheta_{1j} & \cdots & \vartheta_{1n} \\ \vdots & & \vdots & & \vdots \\ \vartheta_{n1} & \cdots & \vartheta_{nj} & \cdots & \vartheta_{nn} \end{pmatrix}$$

stehen die Entwicklungskoeffizienten des  $B$ -Basisvektors  $v_j$  bzgl. der geordneten Basis  $\tilde{B}$

$$v_j = \vartheta_{1j} \tilde{v}_1 + \cdots + \vartheta_{nj} \tilde{v}_n.$$

- (v) In der  $j$ -ten Spalte von  $\Theta_{\tilde{B}B}^{-1}$  stehen die Entwicklungskoeffizienten des  $\tilde{B}$ -Basisvektors  $\tilde{v}_j$  bzgl. der geordneten Basis  $B$ , das heißt

$$\tilde{v}_j = (\Theta_{\tilde{B}B}^{-1})_{1j} v_1 + \cdots + (\Theta_{\tilde{B}B}^{-1})_{nj} v_n.$$

### 12.1.6 Beweis

Die ersten beiden Aussagen ergeben sich aus der Definition der Übergangsmatrix. Für (iv) schauen wir uns das folgende Diagramm an:

$$\sum_{k=1}^n (\Theta_{\tilde{B}B})_{kj} \tilde{v}_k \xrightarrow{\eta_{\tilde{B}}} \Theta_{\tilde{B}B}(e_j) \xrightarrow{\Theta_{\tilde{B}B}^{-1}} e_j \xrightarrow{\eta_B^{-1}} v_j.$$

Wegen  $\Theta_{\tilde{B}B}^{-1} = \eta_B \circ \eta_{\tilde{B}}^{-1}$  ist die gesamte Abbildung die Identität, es besteht also zwischen den Ausdrücken ganz links und ganz rechts Gleichheit.

(v) ergibt sich entsprechend aus einem Diagramm.



## 12.2 Homomorphismen beim Basiswechsel

### 12.2.1 Basis-Entwicklung einer linearen Abbildung

Ist  $B = (v_1, \dots, v_n)$  eine geordnete Basis für  $V$  und  $C = (w_1, \dots, w_m)$  eine geordnete Basis für  $W$ , so ergibt sich das folgende Diagramm von Abbildungen:

$$\mathbb{K}^n \begin{array}{c} \xleftarrow{\eta_B^{-1}} \\ \xrightarrow{\eta_B} \end{array} V \xrightarrow{f} W \begin{array}{c} \xleftarrow{\eta_C} \\ \xrightarrow{\eta_C^{-1}} \end{array} \mathbb{K}^m$$

Die gesamte Abbildung

$$f_{CB} := \eta_C \circ f \circ \eta_B^{-1} : \mathbb{K}^n \rightarrow \mathbb{K}^m$$

ist linear, sie kann also gemäß Abschnitt 7.1.7 durch eine  $m \times n$ -Matrix  $f_{CB} \in \mathbb{K}^{m \times n}$  dargestellt werden.

Beachte, dass diese *Matrixdarstellung* von den geordneten Basen  $C$  und  $B$  abhängt.

### 12.2.2 Satz über die Matrix $f_{CB}$

In der  $j$ -ten Spalte von  $f_{CB}$  stehen die Entwicklungskoeffizienten von  $f(v_j)$  bezüglich der geordneten Basis  $C$ .

### 12.2.3 Begründung

Zur Begründung muss man diese Behauptung etwas abstrakter aufschreiben und die Definition von  $f_{CB}$  einsetzen:

$$\begin{pmatrix} (f_{CB})_{1j} \\ \vdots \\ (f_{CB})_{mj} \end{pmatrix} = f_{CB}(e_j) \stackrel{\text{def}}{=} \eta_C \circ f \circ \underbrace{\eta_B^{-1}(e_j)}_{v_j} = \eta_C(f(v_j)).$$

In diesem Abschnitt seien durchgängig  $V$  und  $W$  zwei  $\mathbb{K}$ -Vektorräume mit  $\dim V = n$ ,  $\dim W = m$ . Weiter sei  $f : V \rightarrow W$  ein Homomorphismus (= lineare Abbildung).

### 12.2.4 Satz: Normalform eines Homomorphismus

Es sei  $f : V \rightarrow W$  wie oben.

Dann gibt es eine geordnete Basis  $B$  für  $V$  und eine geordnete Basis  $C$  für  $W$ , so dass die der linearen Abbildung  $f$  zugeordnete Matrix  $f_{CB}$  bzgl. der geordneten Basen  $B$  und  $C$  die Form

$$f_{CB} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots & \vdots & & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & & \vdots \\ 0 & \cdots & 0 & 1 & \vdots & & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & & \vdots \\ \vdots & & & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \end{pmatrix} \quad \text{mit } r \text{ Einsen, } r = \text{Rang } f,$$

annimmt.

### 12.2.5 Beweis

(1) Der Kern von  $f$  hat nach Satz 7.6.1 die Dimension

$$\dim(\ker f) = n - r$$

Er besitzt also eine geordnete Basis  $(v_{r+1}, \dots, v_n)$ .

(2) Wir ergänzen die Basis des Kerns durch  $r$  Vektoren  $v_1, \dots, v_r$  zu einer geordneten Basis  $B = (v_1, \dots, v_r, v_{r+1}, \dots, v_n)$  von  $V$ .

(3) Es sei jetzt  $w_k := f(v_k)$  für  $k = 1, \dots, r$ .

Wir zeigen, dass die Abbildung

$$\widehat{f}: \langle v_1, \dots, v_r \rangle \rightarrow \langle w_1, \dots, w_r \rangle$$

injektiv ist.

Anderenfalls würde ein Vektor  $v \neq \vec{0}$  existieren mit

$$\begin{aligned} v &\in \langle v_1, \dots, v_r \rangle \\ v &\in \ker f = \langle v_{r+1}, \dots, v_n \rangle \end{aligned}$$

$v$  könnte also als nicht-triviale Linearkombination sowohl von  $v_1, \dots, v_r$  wie von  $v_{r+1}, \dots, v_n$  dargestellt werden. Das steht im Widerspruch dazu, dass  $B = \{v_1, \dots, v_n\}$  Basis von  $V$  ist.

(4) Nach Satz 7.4.2 (i) ist  $\{w_1, \dots, w_r\}$  linear unabhängig. Gemäß Satz 5.7.1 ergänzen wir diese linear unabhängige Menge zu einer geordneten Basis  $C = (w_1, \dots, w_m)$  von  $W$ .

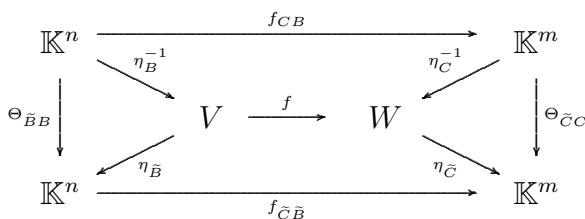
Es gilt dann

$$f(v_k) = \begin{cases} w_k, & \text{falls } k \in \{1, \dots, r\}, \\ \vec{0}, & \text{falls } k \in \{r + 1, \dots, n\}. \end{cases}$$

(5) Das aber bedeutet gerade, dass  $f_{CB}$  die im Satz angegebene Gestalt hat.

### 12.2.6 Homomorphismus beim Basiswechsel

Es seien jetzt jeweils zwei geordnete Basen  $B, \tilde{B}$  für  $V$  und  $C, \tilde{C}$  für  $W$  gegeben. Wir stellen die Frage, was mit der Matrix  $f_{CB}$  bei einem Basiswechsel passiert. Die Situation ergibt sich aus dem folgenden Diagramm:



Man kann ablesen, dass gilt:

$$f_{\tilde{C}\tilde{B}} = \Theta_{\tilde{C}C} \cdot f_{CB} \cdot \Theta_{B\tilde{B}} = \Theta_{\tilde{C}C} \cdot f_{CB} \cdot \Theta_{\tilde{B}B}^{-1}.$$

Es gibt also verschiedene Matrizen für die gleiche lineare Abbildung. Gehören zwei Matrizen zu der gleichen linearen Abbildung, so stehen sie in einer bestimmten Beziehung zueinander, die wir im nächsten Abschnitt definieren und studieren.

## 12.3 Äquivalenz von Matrizen

### 12.3.1 Definition: Äquivalenz von Matrizen

Zwei Matrizen  $F, \tilde{F} \in \mathbb{K}^{m \times n}$  heißen *äquivalent* zueinander, wenn es eine reguläre  $n \times n$ -Matrix  $T$  und eine reguläre  $m \times m$ -Matrix  $S$  gibt, so dass

$$\tilde{F} = S \cdot F \cdot T^{-1}.$$

**12.3.2 Satz** Die Äquivalenz von Matrizen aus  $\mathbb{K}^{m \times n}$  ist eine Äquivalenzrelation auf dieser Menge.

### 12.3.3 Beweis

(1) Reflexivität: Jede Matrix  $F$  ist zu sich selbst äquivalent, man setze in der Definition  $S = I_m$  und  $T = I_n$ .

(2) Symmetrie: Ist die Matrix  $F$  zu  $\tilde{F}$  äquivalent, so gibt es zwei reguläre Matrizen  $S \in \mathbf{GL}(m, \mathbb{K})$ ,  $T \in \mathbf{GL}(n, \mathbb{K})$  mit

$$\tilde{F} = S \cdot F \cdot T^{-1}.$$

Durch Multiplikation dieser Gleichung mit  $S^{-1}$  von links und  $T$  von rechts erhält man

$$S^{-1} \cdot \tilde{F} \cdot T = F,$$

also die Äquivalenz von  $\tilde{F}$  zu  $F$ .

(3) Transitivität: Sind drei Matrizen  $F, \tilde{F}, \hat{F}$  gegeben mit „ $F$  äquivalent zu  $\tilde{F}$ “ und „ $\tilde{F}$  äquivalent zu  $\hat{F}$ “, so gibt es Matrizen  $S, S' \in \mathbf{GL}(m, \mathbb{K})$ ,  $T, T' \in \mathbf{GL}(n, \mathbb{K})$  mit

$$\tilde{F} = S \cdot F \cdot T^{-1} \quad \text{und} \quad \hat{F} = S' \cdot \tilde{F} \cdot (T')^{-1}.$$

Daraus folgt aber

$$\begin{aligned} \hat{F} &= S' \cdot \tilde{F} \cdot (T')^{-1} = S' \cdot S \cdot F \cdot T^{-1} \cdot (T')^{-1} \\ &= (S' \cdot S) \cdot F \cdot (T' \cdot T)^{-1}. \end{aligned}$$

Demzufolge ist auch  $\hat{F}$  äquivalent zu  $F$ .

### 12.3.4 Folgerung: Äquivalenz von Matrizen

- (i) Zwei  $m \times n$ -Matrizen  $F$  und  $\tilde{F}$  sind genau dann äquivalent, wenn sie den gleichen Rang  $r$  haben.
- (ii) In der durch den Rang  $r$  charakterisierten Äquivalenzklasse gibt es einen besonderen Repräsentanten, nämlich die Matrix

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots & \vdots & & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & & \vdots \\ 0 & \cdots & 0 & 1 & \vdots & & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & & \vdots \\ \vdots & & & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \end{pmatrix}$$

mit  $r$  Einsen.

(iii) Die Zahl der Äquivalenzklassen in  $\mathbb{K}^{m \times n}$  ist  $\min\{m, n\} + 1$ .

### 12.3.5 Beweis

(i) und (ii) folgen sofort aus dem Satz 12.2.4.

(iii) Wir wissen bereits, dass für den Rang einer Matrix  $A \in \mathbb{K}^{m \times n}$  gilt:

$$0 \leq \operatorname{rang} A \leq \min\{m, n\}.$$

## 12.4 Beispielaufgaben: Orthogonale Polynome

### 12.4.1 Einstieg

Es sei  $\mathcal{P}_n$  der  $(n+1)$ -dimensionale  $\mathbb{R}$ -Vektorraum der Polynomfunktionen vom Grad  $\leq n$ . Wir beschäftigen uns im folgenden mit  $\mathcal{P}_3$ .

Wir wissen bereits, dass die durch

$$\begin{array}{ll} e_0(x) = 1 & \text{bzw.} \quad f_0(x) = 1 \\ e_1(x) = x & f_1(x) = 1 + x \\ e_2(x) = x^2 & f_2(x) = 1 + x + x^2 \\ e_3(x) = x^3 & f_3(x) = 1 + x + x^2 + x^3 \end{array}$$

gegebenen Polynomfunktionen Basen des Vektorraums  $\mathcal{P}_3$  bilden. Wir nennen im folgenden Polynomfunktionen kurz Polynome, wenn auch wir später diese feine Unterscheidung noch einmal thematisieren werden.

### 12.4.2 Orthogonale Polynome

Die folgenden Polynome sind Beispiele für *orthogonale* Polynome, die in weiten Teilen der Analysis, numerischen Mathematik und mathematischen Physik eine wichtige Rolle spielen.

Hermite-Polynome	Legendre-Polynome	Laguerre-Polynome
$h_0(x) = 1$	$j_0(x) = 1$	$\ell_0(x) = 1$
$h_1(x) = 2x$	$j_1(x) = x$	$\ell_1(x) = -x + 1$
$h_2(x) = 4x^2 - 2$	$j_2(x) = \frac{3}{2}x^2 - \frac{1}{2}$	$\ell_2(x) = \frac{1}{2}(x^2 - 4x + 2)$
$h_3(x) = 8x^3 - 12x$	$j_3(x) = \frac{5}{2}x^3 - \frac{3}{2}x$	$\ell_3(x) = \frac{1}{6}(-x^3 + 9x^2 - 18x + 6)$

Wir bezeichnen die Basen von  $\mathcal{P}_3$ , die aus diesen jeweils vier Polynomen gebildet werden, mit  $E, F, H, J$  bzw.  $L$ .

### 12.4.3 Aufgabe A

Berechnen Sie  $\eta_B(y)$  für  $B = E, H, J, L$  und  $y = \dots$

$$\begin{array}{l} p(x) = 4 \\ q(x) = x + 2 \\ r(x) = x^2 - 3x + 2 \end{array}$$

### 12.4.4 Lösung A

$$\eta_E(p) = \eta_H(p) = \eta_J(p) = \eta_L(p) = \begin{pmatrix} 4 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\eta_E(q) = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \eta_H(q) = \begin{pmatrix} 2 \\ \frac{1}{2} \\ 0 \\ 0 \end{pmatrix}, \quad \eta_J(q) = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \eta_L(q) = \begin{pmatrix} 3 \\ -1 \\ 0 \\ 0 \end{pmatrix}$$

$$\eta_E(r) = \begin{pmatrix} 2 \\ -3 \\ 1 \\ 0 \end{pmatrix}, \quad \eta_H(r) = \begin{pmatrix} \frac{5}{2} \\ \frac{3}{2} \\ \frac{1}{4} \\ 0 \end{pmatrix}, \quad \eta_J(r) = \begin{pmatrix} \frac{7}{3} \\ -3 \\ \frac{2}{3} \\ 0 \end{pmatrix}, \quad \eta_L(r) = \begin{pmatrix} 1 \\ -1 \\ 2 \\ 0 \end{pmatrix}$$

### 12.4.5 Aufgabe B

Berechnen Sie die Übergangsmatrizen  $\Theta_{EH}$ ,  $\Theta_{EJ}$ ,  $\Theta_{EL}$  durch Entwicklung der Basisvektoren von  $H$  nach denen von  $E$  usw.

### 12.4.6 Lösung B

$$\text{(Hermite)} \quad \Theta_{EH} = \begin{pmatrix} 1 & 0 & -2 & 0 \\ 0 & 2 & 0 & -12 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 8 \end{pmatrix}$$

$$\text{(Legendre)} \quad \Theta_{EJ} = \begin{pmatrix} 1 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & -\frac{3}{2} \\ 0 & 0 & \frac{3}{2} & 0 \\ 0 & 0 & 0 & \frac{5}{2} \end{pmatrix}$$

$$\text{(Laguerre)} \quad \Theta_{EL} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & -1 & -2 & -3 \\ 0 & 0 & \frac{1}{2} & \frac{3}{2} \\ 0 & 0 & 0 & -\frac{1}{6} \end{pmatrix}$$

### 12.4.7 Aufgabe C

Berechnen Sie die Übergangsmatrizen  $\Theta_{HE}$ ,  $\Theta_{JE}$ ,  $\Theta_{LE}$  durch

- Entwicklung der Basisvektoren von  $E$  nach denen von  $H$  usw. oder
- durch Invertierung der Matrizen aus B.

**12.4.8 Lösung** C

$$\begin{aligned}
\text{(Hermite)} \quad \Theta_{HE} &= \Theta_{EH}^{-1} = \begin{pmatrix} 1 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{3}{4} \\ 0 & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{8} \end{pmatrix} \\
\text{(Legendre)} \quad \Theta_{JE} &= \Theta_{EJ}^{-1} = \begin{pmatrix} 1 & 0 & \frac{1}{3} & 0 \\ 0 & 1 & 0 & \frac{3}{5} \\ 0 & 0 & \frac{2}{3} & 0 \\ 0 & 0 & 0 & \frac{2}{5} \end{pmatrix} \\
\text{(Laguerre)} \quad \Theta_{LE} &= \Theta_{EL}^{-1} = \begin{pmatrix} 1 & 1 & 2 & 6 \\ 0 & -1 & -4 & -18 \\ 0 & 0 & 2 & 18 \\ 0 & 0 & 0 & -6 \end{pmatrix}
\end{aligned}$$

**12.4.9 Aufgabe** D

Berechnen Sie mit Hilfe der Gleichung  $\Theta_{\hat{B}B} = \Theta_{\hat{B}\tilde{B}} \cdot \Theta_{\tilde{B}\hat{B}}$  weitere Übergangsmatrizen

**12.4.10 Lösungsbeispiel** D

$$\begin{aligned}
\Theta_{LJ} &= \Theta_{LE} \cdot \Theta_{EJ} = \begin{pmatrix} 1 & 1 & 2 & 6 \\ 0 & -1 & -4 & -18 \\ 0 & 0 & 2 & 18 \\ 0 & 0 & 0 & -6 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & -\frac{3}{2} \\ 0 & 0 & \frac{3}{2} & 0 \\ 0 & 0 & 0 & \frac{5}{2} \end{pmatrix} \\
&= \begin{pmatrix} 1 & 1 & \frac{5}{2} & \frac{27}{2} \\ 0 & -1 & -6 & -\frac{87}{2} \\ 0 & 0 & 3 & 45 \\ 0 & 0 & 0 & -15 \end{pmatrix}
\end{aligned}$$

**12.4.11 Aufgabe** E

Überprüfen Sie die Gleichung  $\eta_{\tilde{B}}(y) = \Theta_{\tilde{B}B} \eta_B(y)$  für verschiedene Basen  $B$ ,  $\tilde{B} = E, H, J, L$  und Polynome  $y$  aus A.

**12.4.12 Lösungsbeispiel** E

Für  $y = r = x^2 - 3x + 2$  gilt:

$$\Theta_{LJ} \eta_J(r) = \begin{pmatrix} 1 & 1 & \frac{5}{2} & \frac{27}{2} \\ 0 & -1 & -6 & -\frac{87}{2} \\ 0 & 0 & 3 & 45 \\ 0 & 0 & 0 & -15 \end{pmatrix} \cdot \begin{pmatrix} \frac{7}{3} \\ -3 \\ \frac{2}{3} \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 2 \\ 0 \end{pmatrix}$$

**12.4.13 Aufgabe F**

Die Differentiation  $D$  kann als lineare Abbildung

$$D : \begin{cases} \mathcal{P}_3 & \rightarrow \mathcal{P}_3 \\ \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3 & \mapsto \alpha_1 + 2\alpha_2 x + 3\alpha_3 x^2 \end{cases}$$

aufgefasst werden. Überlegen Sie, dass bezüglich der kanonischen Basis  $E$  für  $\mathcal{P}_3$  dann gilt

$$D_{EE} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Berechnen Sie  $D_{BB}$  oder  $D_{\tilde{B}\tilde{B}}$  für andere Basen  $B, \tilde{B} = E, H, J, L$ .

**12.4.14 Lösungsbeispiel F**

$$\begin{aligned} D_{HH} &= \Theta_{HE} D_{EE} \Theta_{EH} \\ &= \begin{pmatrix} 1 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{3}{4} \\ 0 & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{8} \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -2 & 0 \\ 0 & 2 & 0 & -12 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 8 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 2 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Testen Sie diese Matrix für die Polynome  $r = x^2 - 3x + 2$  und  $Dr = r' = 2x - 3$ .

Lösung: Es gilt

$$\eta_H(r) = \begin{pmatrix} \frac{5}{2} \\ -\frac{3}{2} \\ \frac{1}{4} \\ 0 \end{pmatrix} \quad \text{und} \quad \eta_H(r') = \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

Tatsächlich gilt

$$\eta_H(r') = D_{HH} \cdot \eta_H(r).$$

**12.4.15 Aufgabe G**

Die Multiplikation  $X$  mit einer Variablen kann als lineare Abbildung

$$X : \begin{cases} \mathcal{P}_2 & \rightarrow \mathcal{P}_3 \\ \alpha_0 + \alpha_1 x + \alpha_2 x^2 & \mapsto \alpha_0 x + \alpha_1 x^2 + \alpha_2 x^3 \end{cases}$$

aufgefasst werden. Bezüglich der kanonischen Basen  $\tilde{E}$  für  $\mathcal{P}_2$  und  $E$  für  $\mathcal{P}_3$  gilt dann

$$X_{E\tilde{E}} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Berechnen Sie  $X_{B\tilde{B}}$  für andere Basen  $B$  und  $\tilde{B} = H, J, L$ .



### 12.4.16 Lösungsbeispiel G

$$\begin{aligned}
 X_{J\tilde{J}} &= \Theta_{JE} X_{E\tilde{E}} \Theta_{\tilde{E}\tilde{J}} \\
 &= \begin{pmatrix} 1 & 0 & \frac{1}{3} & 0 \\ 0 & 1 & 0 & \frac{2}{5} \\ 0 & 0 & \frac{2}{3} & 0 \\ 0 & 0 & 0 & \frac{1}{5} \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & \frac{3}{2} \end{pmatrix} \\
 &= \begin{pmatrix} 0 & \frac{1}{3} & 0 \\ 1 & 0 & \frac{2}{5} \\ 0 & \frac{2}{3} & 0 \\ 0 & 0 & \frac{3}{5} \end{pmatrix}
 \end{aligned}$$

Testen Sie diese Matrix für die Polynome  $r = x^2 - 3x + 2$  und  $Xr = x^3 - 3x^2 + 2x$ .

Lösung: Es gilt

$$\eta_{\tilde{J}}(r) = \begin{pmatrix} \frac{7}{3} \\ -3 \\ \frac{2}{3} \end{pmatrix} \quad \text{und} \quad \eta_J(Xr) = \begin{pmatrix} -1 \\ \frac{13}{5} \\ -2 \\ \frac{2}{5} \end{pmatrix}$$

Tatsächlich gilt

$$\eta_J(Xr) = X_{J\tilde{J}} \cdot \eta_{\tilde{J}}(r).$$

## 13 Polynomringe

### 13.1 Polynome

#### 13.1.1 Finite Folgen

Es sei  $\mathbb{K}$  ein Körper. Wir betrachten den Vektorraum  $\mathbb{K}_0^{\mathbb{N}}$  der finiten Folgen mit Elementen aus  $\mathbb{K}$ .

Zur Erinnerung: Eine Folge  $a : \mathbb{N} \rightarrow \mathbb{K}$  heißt *finit*, wenn es ein  $N \in \mathbb{N}$  gibt, so dass

$$a_n = a(n) = 0 \quad \text{für alle } n > N.$$

Anders ausgedrückt: Die Folge enthält nur endlich viele Glieder ungleich Null.

#### 13.1.2 Multiplikation zweier finiter Folgen

Sind nun zwei finite Folgen

$$p = (p_0, p_1, p_2, \dots, p_m, 0, \dots), \quad q = (q_0, q_1, q_2, \dots, q_\ell, 0, \dots)$$

gegeben, so definieren wir ein Produkt dieser beiden Folgen durch

$$(p \cdot q)_n = p_0 \cdot q_n + p_1 \cdot q_{n-1} + \dots + p_{n-1} \cdot q_1 + p_n \cdot q_0.$$

Das bedeutet, dass jeweils alle Glieder von  $p$  und  $q$  mit gleicher Indexsumme  $n$  miteinander multipliziert werden und dann alle Produkte aufsummiert werden. Das Ergebnis steht an der Stelle  $n$  des Produktpolynoms.

Man mache sich klar, dass das Ergebnis wieder eine finite Folge ist.

#### 13.1.3 Beispiel

$$\begin{aligned} & (4, 2, -5, 3, 0, 0, 0, \dots) \cdot (-1, 3, -4, 0, 0, 0, \dots) \\ &= (4 \cdot (-1), 4 \cdot 3 + 2 \cdot (-1), 4 \cdot (-4) + 2 \cdot 3 + (-5) \cdot (-1), \\ &\quad 2 \cdot (-4) + (-5) \cdot 3 + 3 \cdot (-1), (-5) \cdot (-4) + 3 \cdot 3, 3 \cdot (-4), 0, 0, \dots) \\ &= (-4, 10, -5, -26, 29, -12, 0, 0, \dots) \end{aligned}$$

#### 13.1.4 Definitionen: Polynom als finite Folge

- (1) Im Zusammenhang mit dieser Multiplikation heißt eine finite Folge  $(p_n)_{n \in \mathbb{N}}$  in  $\mathbb{K}$  auch (*abstraktes*) *Polynom über  $\mathbb{K}$* .
- (2) Ist  $(p_n)_{n \in \mathbb{N}}$  nicht die Nullfolge, so heißt die eindeutig bestimmte Zahl  $m \in \mathbb{N}$  mit

$$p_m \neq 0, \quad p_n = 0 \quad \text{für alle } n > m,$$

der *Grad* des Polynoms. Die Nullfolge hat per definitionem den Grad gleich  $-\infty$ .

- (3) Ist  $m$  der Grad eines Polynoms, so nennt man  $p_m$  den *Leitkoeffizienten*.
- (4) Ein Polynom heißt *normiert*, wenn der Leitkoeffizient gleich 1 ist.

Die Polynome höchstens  $m$ -ten Grades bilden einen  $\mathbb{K}$ -Vektorraum der Dimension  $m + 1$ .

### 13.1.5 Gradformel

Sind  $p, q \in \mathbb{K}[x]$  zwei Polynome, so gilt:

$$\text{grad}(p \cdot q) = \text{grad } p + \text{grad } q.$$

Der Fall des Nullpolynoms ist hier enthalten, wenn man Rechenregeln für  $-\infty$  einbezieht. Wenn das nicht gefällt, sollte man diesen Fall für die Gradformel ausschließen.

### 13.1.6 Nebenbemerkung

Gemäß der obigen Konstruktion kann für jeden Ring der Polynomring erstellt werden. Ist eine Abbildung zwischen Ringen gegeben, so lässt sie sich zu einer Abbildung zwischen den zugehörigen Polynomringen fortsetzen (Funktor).

Die Gradformel stimmt nicht mehr, wenn man Polynome über allgemeineren (Nullteiler enthaltenden) Ringen betrachtet: In  $\mathbb{Z}_6$  ist

$$(3x - 1) \cdot (2x - 1) = 6x^2 - 5x + 1 = x + 1 \pmod{6}.$$

### 13.1.7 Der Polynomring $\mathbb{K}[x]$

Auf der Menge der abstrakten Polynome sind eine Addition und eine Multiplikation erklärt, so dass die Axiome für einen Ring mit Eins (vgl. Abschnitt 4.3.4) erfüllt sind.

Diese Menge wird deshalb als *Polynomring* bezeichnet. Als Symbol hat sich international  $\mathbb{K}[x]$  als Standard etabliert. Das ist ein bisschen seltsam, da der Buchstabe  $x$  eigentlich ohne Bedeutung ist.

Da auf  $\mathbb{K}[x]$  auch noch die Struktur eines  $\mathbb{K}$ -Vektorraums existiert, spricht man auch von der Polynom- $\mathbb{K}$ -Algebra  $\mathbb{K}[x]$ .

## 13.2 Polynomfunktionen

### 13.2.1 Definition: Polynomfunktion

Es sei  $\mathbb{K}$  ein Körper.

- (1) Eine Abbildung  $p : \mathbb{K} \rightarrow \mathbb{K}$  heißt *Polynomfunktion (über  $\mathbb{K}$ )* wenn es Zahlen  $m \in \mathbb{N}_0$ ,  $p_m, p_{m-1}, \dots, p_1, p_0 \in \mathbb{K}$  gibt, so dass

$$p(x) = p_m x^m + p_{m-1} x^{m-1} + \dots + p_1 x + p_0.$$

- (2) Eine Zahl  $\lambda \in \mathbb{K}$  heißt *Nullstelle* der Polynomfunktion  $p$ , wenn  $p(\lambda) = 0$ .

Polynomfunktionen können multipliziert und nacheinander ausgeführt werden.

### 13.2.2 Zuordnung: Polynom $\rightarrow$ Polynomfunktion

Man kann nun einem Polynom (= finite Folge) eine Polynomfunktion (= Abbildung  $\mathbb{K} \rightarrow \mathbb{K}$ ) zuordnen:

$$(p_0, p_1, p_2, \dots, p_m, 0, \dots) \mapsto p_0 + p_1 x + p_2 x^2 + \dots + p_m x^m \quad (*)$$

Es ist nicht schwer nachzuprüfen, dass die obige abstrakte Multiplikation 13.1.2 bei dieser Zuordnung in die bekannte Polynom-Multiplikation übergeht.

Der Vorteil der abstrakteren Sichtweise aus Kapitel 13.1 besteht darin, dass die algebraische Struktur des Polynomrings unabhängig von der Idee „Polynom als Abbildung“, also auch unabhängig von Variablen-Symbolen wie  $x, y, \dots$ , eingeführt werden kann.

Im Fall  $\mathbb{K} = \mathbb{R}$  oder  $\mathbb{K} = \mathbb{C}$  ist die Zuordnung (\*) ein-eindeutig. Ist eine polynomiale Abbildung gegeben, so können die zugehörigen Polynom-Koeffizienten ermittelt werden — beispielsweise durch mehrfaches Differenzieren.

### 13.2.3 Beispiel

Ist beispielsweise  $\mathbb{K} = \mathbb{Z}_2$  der Körper mit zwei Elementen, so sind die beiden konkreten Polynome

$$p(x) = x^2 + x + 1, \quad q(x) = x^3 + x + 1$$

aufgrund von

$$p(0) = q(0) = 1, \quad p(1) = q(1) = 1$$

als Abbildungen  $\mathbb{Z}_2 \rightarrow \mathbb{Z}_2$  gleich. Als abstrakte Polynome sind sie jedoch verschieden:

$$p = (1, 1, 1, 0, 0, \dots) \neq (1, 1, 0, 1, 0, \dots) = q.$$

Das Beispiel zeigt auch, dass der Grad einer Polynomfunktion nicht eindeutig definiert sein muss.

Nachdem wir den Unterschied zwischen Polynomen und den ihnen zugeordneten Polynomfunktionen einmal herausgestellt haben, wollen wir ihn ab jetzt nicht mehr so stark thematisieren. Wir sprechen einfach nur von Polynomen.

### 13.3 Teilbarkeit in Polynomringen

#### 13.3.1 Definition: Teilbarkeit

Es seien  $p, q \in \mathbb{K}[x]$  zwei Polynome.

Das Polynom  $q$  heißt *Teiler* von  $p$ , wenn es ein weiteres Polynom  $g$  gibt, so dass

$$p = g \cdot q.$$

#### 13.3.2 Satz: Polynomdivision

Es sei  $\mathbb{K}$  ein Körper.

Sind  $p, q \in \mathbb{K}[x]$ , zwei Polynome, dabei  $q \neq 0$ , so gibt es dazu eindeutig bestimmte Polynome  $g, r \in \mathbb{K}[x]$  mit der Eigenschaft:

$$p = g \cdot q + r, \quad \text{grad } r < \text{grad } q.$$

#### 13.3.3 Beweis

(0) Zunächst zur Eindeutigkeit: Angenommen es gibt jeweils zwei Polynome  $g, \tilde{g}$  und  $r, \tilde{r}$  mit den im Satz geforderten Eigenschaften, also

$$\begin{aligned} p &= g \cdot q + r & \text{grad } r < \text{grad } q \\ p &= \tilde{g} \cdot q + \tilde{r} & \text{grad } \tilde{r} < \text{grad } q. \end{aligned}$$

(1) Die Subtraktion der beiden Gleichungen führt auf

$$0 = (g - \tilde{g}) \cdot q + (r - \tilde{r})$$

und dann auf

$$(g - \tilde{g}) \cdot q = \tilde{r} - r.$$

(2) Daraus lässt sich eine Ungleichung folgern

$$\text{grad } q \stackrel{(0)}{>} \text{grad}(\tilde{r} - r) = \text{grad}(g - \tilde{g}) + \text{grad } q,$$

die im Fall  $\tilde{r} \neq r$  oder  $\tilde{g} \neq g$  einen Widerspruch darstellt.

(3) Nun zur Existenz. Hier führen wir einen Induktionsbeweis über  $n = \text{grad } p$  durch.

(4) Induktionsanfang:  $n < \text{grad } q$ . In diesem Fall ist mit  $g = 0$  und  $r = p$  die Aussage des Satzes erfüllt.

(5) Induktionsschluss: Es sei die Aussage für  $n$  gezeigt. Vorgegeben sind nun zwei Polynome  $p, q$  der Gestalt

$$\begin{aligned} p(x) &= p_{n+1}x^{n+1} + p_nx^n + \dots + p_1x + p_0 \\ q(x) &= q_mx^m + p_{m-1}x^{m-1} + \dots + q_1x + q_0. \end{aligned}$$

(6) Ist  $m \geq n + 1$ , so setze einfach  $q = 0$  und  $r = p$ . Es ist dann  $\text{grad } r = \text{grad } p < \text{grad } q$ .

(7) Im anderen Fall  $m < n + 1$  setzen wir

$$\tilde{p}(x) := p(x) - \frac{p_{n+1}}{q_m}x^{n+1-m} \cdot q(x).$$

In diesem Polynom hat der Term mit Grad  $n + 1$  die Form

$$\tilde{p}_{n+1} = p_{n+1}x^{n+1} - \frac{p_{n+1}}{q_m}x^{n+1-m} \cdot q_mx^m = 0,$$

weshalb es einen Grad  $\leq n$  hat.

(8) Nach Induktionsvoraussetzung gibt es zu  $\tilde{p}, q$  zwei Polynome  $\tilde{g}$  und  $\tilde{r}$  mit

$$\tilde{p}(x) = \tilde{g}(x) \cdot q(x) + \tilde{r}(x), \quad \text{grad } \tilde{r} < \text{grad } q.$$

Wir setzen dies in die Definition von  $\tilde{p}$  ein. Nach Umstellung ergibt sich:

$$\begin{aligned} p(x) &= \tilde{p}(x) + \frac{p_{n+1}}{q_m}x^{n+1-m} \cdot q(x) \\ &= \underbrace{\left[ \tilde{g}(x) + \frac{p_{n+1}}{q_m}x^{n+1-m} \right]}_{=:g(x)} \cdot q(x) + \tilde{r}(x). \end{aligned}$$

Der Beweis ist beendet.

### 13.3.4 Beispiel: Polynomdivision

Die beiden Polynome  $g$  und  $r$  aus dem obigen Satz erhält man algorithmisch durch die aus der Schule bekannte Polynomdivision.

Man überzeuge sich mit Hilfe der Polynomdivision

$$(3x^5 - 4x^4 + 2x^3 - x + 8) : (x^3 + 7x^2 - 5)$$

von der Darstellung

$$3x^5 - 4x^4 + 2x^3 - x + 8 = (3x^2 - 25x + 177)(x^3 + 7x^2 - 5) - 1224x^2 - 126x + 893.$$

## 13.4 Nullstellen und Linearfaktoren

### 13.4.1 Satz: Nullstellen und Linearfaktoren

Die beiden folgenden Aussagen über ein Polynom  $p \in \mathbb{K}[x]$   $n$ -ten Grades und eine Zahl  $\lambda \in \mathbb{K}$  sind äquivalent:

- (A)  $\lambda$  ist eine Nullstelle von  $p$ .  
 (B) Es gibt ein Polynom  $q$  vom Grad  $n - 1$ , so dass

$$p(x) = (x - \lambda) \cdot q(x).$$

Man nennt  $x - \lambda$  einen *Linearfaktor* von  $p$ . Man spricht hier auch von der *Abspaltung* des Linearfaktors.

### 13.4.2 Beweis

Die Richtung (B)  $\Rightarrow$  (A) ist trivial. Für die Umkehrung machen wir eine Vorüberlegung: Es gilt für  $m \in \mathbb{N}$

$$\begin{aligned} (x^m - \lambda^m) &= \left[ x^m + x^{m-1}\lambda + x^{m-2}\lambda^2 + \dots + x^2\lambda^{m-2} + x\lambda^{m-1} \right] \\ &\quad - \left[ x^{m-1}\lambda + x^{m-2}\lambda^2 + \dots + x^2\lambda^{m-2} + x\lambda^{m-1} + \lambda^m \right] \\ &= (x - \lambda) \cdot \underbrace{(x^{m-1} + x^{m-2}\lambda + \dots + x\lambda^{m-2} + \lambda^{m-1})}_{=: Q_{m-1}(x, \lambda)} \end{aligned}$$

Jetzt können wir die Umkehrung zeigen:

$$\begin{aligned} p(x) &= p(x) - p(\lambda) \\ &= \left( p_n x^n + p_{n-1} x^{n-1} + \dots + p_1 x + p_0 \right) - \left( p_n \lambda^n + p_{n-1} \lambda^{n-1} + \dots + p_1 \lambda + p_0 \right) \\ &= p_n (x^n - \lambda^n) + p_{n-1} (x^{n-1} - \lambda^{n-1}) + \dots + p_1 (x - \lambda) \\ &= (x - \lambda) \cdot \left( p_n Q_{n-1}(x, \lambda) + p_{n-1} Q_{n-2}(x, \lambda) + \dots + p_1 \right) \\ &= (x - \lambda) \cdot q(x). \end{aligned}$$

Die fortgesetzte Abspaltung von Linearfaktoren führt auf die

### 13.4.3 Folgerung

Zu einem Polynom  $p$  mit Nullstelle  $\lambda$  gibt es eine eindeutig festgelegte Zahl  $k \in \mathbb{N}$  und ein Polynom  $q$  vom Grad  $n - k$ , so dass

$$p(x) = (x - \lambda)^k \cdot q(x), \quad \text{wobei } q(\lambda) \neq 0.$$

Die Zahl  $k$  heißt die *Vielfachheit* der Nullstelle  $\lambda$  im Polynom  $p$ .

### 13.4.4 Definition

Man sagt, ein Polynom *zerfällt (über  $\mathbb{K}$ ) in Linearfaktoren*, wenn es in der Form

$$p(x) = p_n \cdot (x - \lambda_1)^{k_1} \cdot (x - \lambda_2)^{k_2} \cdot \dots \cdot (x - \lambda_m)^{k_m}$$

mit Zahlen  $\lambda_1, \dots, \lambda_m \in \mathbb{K}$  und  $k_1, \dots, k_m \in \mathbb{N}$  dargestellt werden kann.

## 13.5 Polynome über $\mathbb{R}$ und $\mathbb{C}$

### 13.5.1 Beispiel

Das Polynom  $x^2 + 1$  zerfällt über  $\mathbb{R}$  nicht in Linearfaktoren, über  $\mathbb{C}$  schon:

$$x^2 + 1 = (x - i) \cdot (x + i).$$

### 13.5.2 Satz: Fundamentalsatz der Algebra

Es sei  $p$  ein komplexes Polynom mit  $\text{grad } p \geq 1$ .

- (i)  $p$  besitzt eine Nullstelle in  $\mathbb{C}$ .
- (ii)  $p$  zerfällt über  $\mathbb{C}$  vollständig in Linearfaktoren.

### 13.5.3 Beweis

(i) Er muss mit Mitteln der Analysis geführt werden. Er ist allerdings für den Rahmen einer grundlegenden Vorlesung Analysis zu aufwändig, mit Methoden der Funktionentheorie geht es ganz schnell.

(ii) Per Induktion über den Grad  $n$  des Polynoms. Ein komplexes Polynom  $p$  hat nach (i) mindestens eine Nullstelle. Nach Satz 13.4.1 kann das Polynom durch diesen Linearfaktor dividiert werden, man erhält ein Polynom mit einem um 1 geringeren Grad.

### 13.5.4 Satz: Zerlegung reeller Polynome

Es sei  $p \in \mathbb{R}[x]$  ein Polynom mit reellen Koeffizienten. Da reelle Zahlen spezielle komplexe Zahlen sind, kann das Polynom auch als komplexes Polynom, das heißt  $p \in \mathbb{C}[x]$ , aufgefasst werden.

- (i) Ist  $\lambda \in \mathbb{C}$  eine Nullstelle von  $p$ , so ist auch  $\bar{\lambda}$  Nullstelle.
- (ii) Ist  $\lambda \in \mathbb{C} \setminus \mathbb{R}$  eine Nullstelle, so ist das quadratische reelle Polynom

$$(x - \lambda) \cdot (x - \bar{\lambda}) = x^2 - 2 \operatorname{Re} \lambda x + |\lambda|^2$$

ein Teiler von  $p$ .

- (iii) Jedes reelle Polynom zerfällt vollständig in Linearfaktoren und quadratische Faktoren.

### 13.5.5 Beweis

(i) Ist  $\lambda \in \mathbb{C}$  eine Nullstelle von  $p$ , das heißt

$$p_0 + p_1 \lambda + p_2 \lambda^2 + \dots + \lambda_n \lambda^n = 0,$$

so folgt

$$\begin{aligned} p_0 + p_1 \bar{\lambda} + p_2 \bar{\lambda}^2 + \dots + p_n \bar{\lambda}^n &= \overline{p_0 + p_1 \lambda + p_2 \lambda^2 + \dots + p_n \lambda^n} \\ &= \overline{0} = 0. \end{aligned}$$



(ii) Gemäß (i) ist mit  $\lambda$  auch  $\bar{\lambda} \neq \lambda$  eine Nullstelle. Nach Satz 13.4.1 können die beiden Linearfaktoren  $(x - \lambda)$  und  $(x - \bar{\lambda})$  abgespalten werden. Das Produkt dieser beiden Linearfaktoren ist aber das reelle Polynom wie im Satz angegeben.

(iii) Per Induktion über den Grad  $n$  des Polynoms. Ein reelles Polynom  $p$  ist auch ein komplexes Polynom und hat nach Satz 13.5.2 mindestens eine (komplexe) Nullstelle.

Ist diese Nullstelle reell, so kann sie als Linearfaktor gemäß Satz 13.4.1 abgespalten werden.

Ist diese Nullstelle nicht reell, so kann nach (ii) des aktuellen Satzes das zugehörige quadratische Polynom abgespalten werden.

Die Division durch diesen Faktor führt auf ein reelles Polynom mit einem Grad um 1 oder 2 geringer.

## 13.6 Das Minimalpolynom eines Ideals

### 13.6.1 Definition: Ideale

Eine nicht-leere Teilmenge  $\mathcal{I} \subseteq \mathbb{K}[x]$  eines Polynomrings (allgemeiner: eines kommutativen Rings) heißt Ideal, wenn die folgenden drei Eigenschaften erfüllt sind:

- Wenn  $p, q \in \mathcal{I}$ , dann auch  $p + q \in \mathcal{I}$
- Wenn  $p \in \mathcal{I}$ , dann auch  $-p \in \mathcal{I}$
- Wenn  $p \in \mathcal{I}$  und  $q \in \mathbb{K}[x]$ , dann auch  $p \cdot q \in \mathcal{I}$ .

Da ein beliebiger Skalar in  $\mathbb{K}$  zugleich ein konstantes Polynom ist, folgt, dass ein Ideal auch ein  $\mathbb{K}$ -Unterraum ist.

### 13.6.2 Beispiele

(1) Im Ring  $\mathbb{Z}$  der ganzen Zahlen ist für jedes  $n \in \mathbb{N}$  die  $n$ -Vielfachenmenge

$$V_n = \{z \in \mathbb{Z} \mid \exists k \in \mathbb{Z} \text{ mit } z = k \cdot n\}$$

ein Ideal.

(2) Es sei  $\mathcal{F}(\mathbb{R}, \mathbb{R})$  der Ring der Funktionen  $\mathbb{R} \rightarrow \mathbb{R}$ . Für jede feste Teilmenge  $A \subseteq \mathbb{R}$  ist die Menge der Funktionen, die auf  $A$  gleich Null sind,

$$N_A = \{f \in \mathcal{F}(\mathbb{R}, \mathbb{R}) \mid f(a) = 0 \text{ für alle } a \in A\}$$

ein Ideal.

(3) Wir wenden uns nun dem Polynomring  $\mathbb{K}[x]$  zu. Für ein festes Polynom  $q \in \mathbb{K}[x]$  ist wieder die Menge der Vielfachen von  $q$

$$V_q = \{p \in \mathbb{K}[x] \mid \exists s \in \mathbb{K}[x] \text{ mit } p = s \cdot q\}$$

ein Ideal.

Es ist aber auch für jede Teilmenge  $A \subseteq \mathbb{K}$  die Menge der Polynome, die alle  $a \in A$  als Nullstelle haben,

$$N_A = \{p \in \mathbb{K}[x] \mid p(a) = 0 \text{ für alle } a \in A\}$$

ein Ideal.

In Abschnitt 13.4 haben wir gesehen, dass zwischen diesen beiden Typen von Idealen enge Beziehungen bestehen. Dieses Wechselspiel ist der Ausgangspunkt für ein Teilgebiet der Mathematik, die algebraische Geometrie.

### 13.6.3 Satz: Minimalpolynom eines Ideals

Es sei  $\mathcal{I} \neq \{0\}$  ein Ideal in einem Polynomring  $\mathbb{K}[x]$ .

- (i) Für ein Polynom  $m \in \mathcal{I} \setminus \{0\}$  sind die beiden Aussagen äquivalent:
- (A)  $m$  ist Teiler aller Polynome in  $\mathcal{I}$ .
  - (B) Der Grad von  $m$  ist kleiner oder gleich dem aller anderen Polynome in  $\mathcal{I} \setminus \{0\}$ .
- (ii) Es gibt genau ein Polynom in  $\mathcal{I}$ , das die beiden Bedingungen von (i) erfüllt und als Leitkoeffizient die Zahl 1 hat.

Dieses Polynom heißt *Minimalpolynom* des Ideals. Man kann es mit  $m_{\mathcal{I}}$  bezeichnen.

**13.6.4 Beweis** Die Folgerung (A)  $\Rightarrow$  (B) folgt sofort mit der Gradformel 13.1.5.

Umgekehrt habe das Polynom  $m \in \mathcal{I}$  minimalen Grad in  $\mathcal{I}$ . Ist dann  $p \in \mathcal{I}$  irgendein anderes Polynom, so können wir die Polynomdivision „ $p$  durch  $m$ “ nach Satz 13.3.2 ausführen und erhalten zwei Polynome  $q$  und  $r$  mit

$$p = q \cdot m + r \quad \text{mit} \quad \text{grad } r < \text{grad } m.$$

Da  $m$  und  $p$  in dem Ideal  $\mathcal{I}$  enthalten sind, ist auch

$$r = p - q \cdot m$$

in  $\mathcal{I}$  enthalten. Im Fall  $r = 0$  ist die Aussage (A) zu beweisen, im anderen Fall erhalten wir einen Widerspruch dazu, dass  $m$  minimalen Grad in  $\mathcal{I}$  hat.

(ii) Da das Ideal  $\mathcal{I}$  nicht nur das Nullpolynom enthält, enthält es irgendein Polynom. Gibt es noch ein Polynom in  $\mathcal{I}$  mit kleinerem Grad, so liefert fortgesetzte Polynomdivision ein Polynom  $m$  mit der Eigenschaft (B) aus (i).

Dieses Polynom wird dann durch seinen Leitkoeffizienten ungleich Null dividiert. Der neue Leitkoeffizient ist dann 1.

## 14 Die Polynomialgebra einer quadratischen Matrix

### 14.1 Einführung

In diesem gesamten Kapitel lasse man sich von der Vorstellung einer fest gegebenen Matrix  $A \in \mathbb{K}^{n \times n}$  leiten.

#### 14.1.1 Vorbemerkung

Für eine Zahl  $\alpha \in \mathbb{K}$  bezeichnen wir das  $\alpha$ -Vielfache der Einheitsmatrix oft ebenfalls mit  $\alpha$ .

$$\alpha \cong \alpha \cdot I = \begin{pmatrix} \alpha & 0 & \cdots & \cdots & 0 \\ 0 & \alpha & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & \alpha \end{pmatrix}.$$

Die Wirkung auf Vektoren bzw. andere Matrizen stimmen ja überein.

#### 14.1.2 Matrix-Polynome

Es sei also eine feste Matrix  $A \in \mathbb{K}^{n \times n}$  gegeben. Ist weiter ein Polynom  $p \in \mathbb{K}[x]$  vom Grad  $\ell$  gegeben,

$$p \sim (p_0, p_1, \dots, p_\ell, 0, \dots) \sim p_0 + p_1x + p_2x^2 + \dots + p_\ell x^\ell,$$

so können wir daraus eine neue Matrix in  $\mathbb{K}^{n \times n}$  bilden, nämlich

$$\begin{aligned} p(A) &= p_0 + p_1 \cdot A + p_2 \cdot A^2 + \dots + p_\ell \cdot A^\ell \\ &= p_0 I + p_1 \cdot A + p_2 \cdot A^2 + \dots + p_\ell \cdot A^\ell \end{aligned}$$

Man beachte, dass die Auffassung vom Polynom  $p$  als Abbildung  $\mathbb{K} \rightarrow \mathbb{K}$  für diese Definition zu eng ist.

Die Menge aller Matrizen, die auf diese Weise bei fester Matrix  $A$  und beliebigen Polynomen  $p$  gebildet werden können, bezeichnen wir mit

$$\mathbb{K}[A] = \{M \in \mathbb{K}^{n \times n} \mid \text{Es existiert Polynom } p \in \mathbb{K}[x] \text{ mit } M = p(A)\}.$$

Diese Menge ist ein  $\mathbb{K}$ -Unterraum von  $\mathbb{K}^{n \times n}$  und zugleich ein kommutativer Ring mit 1.

## 14.2 Das Minimalpolynom einer quadratischen Matrix

### 14.2.1 Definition: Annihilator

Es sei weiterhin eine feste Matrix  $A \in \mathbb{K}^{n \times n}$  gegeben.

Wir betrachten nun die Teilmenge aller Polynome, die bei Einsetzung von  $A$  die Null-Matrix ergeben.

$$\mathcal{N}_A := \{p \in \mathbb{K}[x] \mid p(A) = 0\}.$$

Diese Teilmenge ist ein Ideal von  $\mathbb{K}[x]$ , da sie einen Unterraum darstellt und außerdem gegenüber beliebigen Multiplikationen mit Polynomen abgeschlossen ist. Es heißt das Annihilator-Ideal (auch *Nullmacher-Ideal* oder nur *Annihilator*) von  $A$ .

### 14.2.2 Beobachtung: $\mathcal{N}_A \neq \{0\}$

In der endlichen Folge  $I, A, A^2, A^3, \dots, A^{n^2}$  von Matrixpotenzen

- sind entweder zwei Matrizen  $A^m$  und  $A^k$ ,  $m \neq k$ , gleich und dann ist  $x^m - x^k$  ein Polynom ungleich Null in  $\mathcal{N}_A$ ,
- oder sind alle  $n^2 + 1$  Matrizen paarweise verschieden und dann wegen  $\dim \mathbb{K}^{n \times n} = n^2$  linear abhängig. Es gibt daher Skalare  $\alpha_0, \alpha_1, \alpha_2, \dots, \alpha_{n^2}$ , nicht alle gleich Null, so dass

$$\alpha_0 + \alpha_1 A + \alpha_2 A^2 + \dots + \alpha_{n^2} A^{(n^2)} = 0$$

Wir haben ein Polynom ungleich Null in  $\mathcal{N}_A$  gefunden.

### 14.2.3 Definition: Minimalpolynom einer quadratischen Matrix

Gemäß Abschnitt 13.6.3 besitzt der Annihilator von  $A$  ein Minimalpolynom. Wir nennen es das *Minimalpolynom*  $\mu_A$  der quadratischen Matrix  $A$ .

### 14.2.4 Satz: Minimalpolynom einer quadratischen Matrix

Das Minimalpolynom  $\mu_A$  der quadratischen Matrix  $A$  ist durch jede der beiden folgenden Aussagen eindeutig gegeben:

- $\mu_A$  ist normiert und Teiler jedes Polynoms  $p$  mit  $p(A) = 0$ .
- $\mu_A$  ist normiert und hat einen Grad, der kleiner-oder-gleich ist dem Grad jedes Polynoms  $p$  mit  $p(A) = 0$ ,  $p \neq 0$ .

### 14.2.5 Beweis

Er wurde bereits allgemeiner durch Satz 13.6.3 erbracht.

### 14.2.6 Beispiel: Vielfache der Einheitsmatrix

Hat das Minimalpolynom einer Matrix  $A \in \mathbb{K}^{n \times n}$  den Grad 1,

$$\mu_A(x) = x - \lambda,$$

wobei  $\lambda \in \mathbb{K}$ , so gilt die Matrixgleichung

$$A - \lambda I = 0, \quad \text{also } A = \lambda I.$$

Eine quadratische Matrix  $A$  hat also genau dann ein Minimalpolynom vom Grad 1, wenn sie ein skalares Vielfaches der Einheitsmatrix ist.

### 14.2.7 Beispiel: Diagonalmatrix

Eine Diagonalmatrix

$$D = \begin{pmatrix} d_{11} & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & d_{nn} \end{pmatrix} \in \mathbb{K}^{n \times n}$$

hat als Minimalpolynom

$$\mu_D(x) = (x - \lambda_1) \cdot \dots \cdot (x - \lambda_m)$$

wobei die Zahlen  $\lambda_1, \dots, \lambda_m$  paarweise verschieden sind und die Diagonalelemente „umfassen“:

$$\begin{aligned} \lambda_i &\neq \lambda_j, \quad \text{falls } i \neq j \\ \{\lambda_1, \dots, \lambda_m\} &= \{d_{11}, \dots, d_{nn}\}. \end{aligned}$$

Für

$$D = \begin{pmatrix} -3 & 0 & \cdots & \cdots & 0 \\ 0 & 5 & \ddots & & \vdots \\ \vdots & \ddots & -3 & \ddots & \vdots \\ \vdots & & \ddots & 5 & 0 \\ 0 & \cdots & \cdots & 0 & 5 \end{pmatrix} \quad \text{ist } \mu_D(x) = (x - 5)(x + 3)$$

### 14.2.8 Beispiel: Jordan-Block

Für gegebenes  $n \in \mathbb{N}$  und  $\lambda \in \mathbb{K}$  heißt die  $n \times n$ -Matrix

$$J = J_n(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & 0 & \lambda \end{pmatrix}$$

der  $n \times n$ -Jordan-Block zu  $\lambda$ . Es ist dann

$$J - \lambda I = J_n(0) = \begin{pmatrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & 0 & 0 \end{pmatrix}$$

und deshalb

$$(J - \lambda I)^k \neq 0 \quad \text{für } k = 1, \dots, n-1$$

$$(J - \lambda I)^n = 0.$$

Aus der unteren Zeile folgt (vgl. Satz: Minimalpolynom), dass das Minimalpolynom  $\mu_J$  ein Teiler von  $(x - \lambda)^n$  ist, die obere Zeile zeigt dann, dass es gleich  $(x - \lambda)^n$  ist:

$$\mu_J(x) = (x - \lambda)^n.$$

## 14.3 Das charakteristische Polynom einer quadratischen Matrix

### 14.3.1 Satz und Definition: Charakteristisches Polynom

Ist  $A$  eine quadratische Matrix, so ist der Ausdruck

$$\chi_A(x) := \det(A - x)$$

ein Polynom  $n$ -ten Grades in  $x$ . Es heißt das *charakteristische Polynom der Matrix  $A$* .  
Genauer gilt:

$$\chi_A(x) = (-1)^n x^n + (-1)^{n-1} (a_{11} + a_{22} + \dots + a_{nn}) x^{n-1} + \dots + \det A.$$

Die Summe der Diagonalelemente einer quadratischen Matrix  $A$  heißt die *Spur der Matrix  $A$* .

$$\text{Spur } A := a_{11} + \dots + a_{nn}.$$

### 14.3.2 Beweis

(0) Wir schreiben die Determinante aus

$$\det(A - x) = \det \begin{pmatrix} a_{11} - x & a_{12} & \cdots & \cdots & a_{1n} \\ a_{21} & a_{22} - x & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & a_{n-1,n} \\ a_{n1} & \cdots & \cdots & a_{n,n-1} & a_{nn} - x \end{pmatrix}$$

und überlegen, wie die Variable  $x$  über die Leibniz-Formel in dem Ausdruck  $\chi_A(x)$  Eingang findet.

(1) Da in der Leibniz-Formel eine (vorzeichenbehaftete) Summe von  $n$ -fachen Produkten der Matrix-Einträge auftritt, ist es klar, dass es sich um ein Polynom höchstens  $n$ -ten Grades handelt.

(2) Die Potenzen  $x^n$  und  $x^{n-1}$  können nur in dem Produkt auftreten, das zu der Permutation  $\pi = \text{id}$  gehört. In den Produkten, die zu den anderen Permutationen  $\pi \in \mathcal{S}_n \setminus \{(1)\}$  gehören, kommen

- mindestens zwei Nicht-Diagonalelemente der Matrix  $A - x$  und deshalb
- höchstens  $n - 2$  Diagonalelemente dieser Matrix

vor. Das heißt aber, dass  $x$  in diesen Produkten höchstens mit dem Exponenten  $n - 2$  auftritt.

(3) Es genügt also, das zur identischen Permutation gehörende Produkt, also das Produkt aller Diagonalelemente, zu betrachten. Dafür gilt aber

$$\begin{aligned} (a_{11} - x) \cdot (a_{22} - x) \cdot \dots \cdot (a_{nn} - x) &= \\ (-x)^n + (a_{11} + a_{22} + \dots + a_{nn})(-x)^{n-1} &+ \text{Glieder niedrigerer Ordnung} \end{aligned}$$

Das sind genau die ersten beiden Glieder im charakteristischen Polynom.

(4) Setzt man  $x = 0$ , so sieht man, dass das konstante Glied im charakteristischen Polynom gerade die Determinante von  $A$  ist.



### 14.3.3 Beispiel: $2 \times 2$ -Matrix

Wir betrachten eine allgemeine  $2 \times 2$ -Matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

und berechnen ihr charakteristisches Polynom zu

$$\begin{aligned} \chi_A(x) &= \det(A - x) = \det \begin{pmatrix} a - x & b \\ c & d - x \end{pmatrix} = (a - x)(d - x) - bc \\ &= x^2 - \underbrace{(a + d)}_{\text{spur } A} x + \underbrace{(ad - bc)}_{\det A}. \end{aligned}$$

Die Matrix  $\begin{pmatrix} 0 & -1 \\ +1 & 0 \end{pmatrix}$  hat das charakteristische Polynom  $x^2 + 1$ .

### 14.3.4 Beispiel: Dreiecksmatrizen

Ist  $T$  eine obere Dreiecksmatrix, so gilt für das charakteristische Polynom

$$\begin{aligned} \det(T - x) &= \det \begin{pmatrix} t_{11} - x & * & \cdots & \cdots & * \\ 0 & t_{22} - x & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & * \\ 0 & \cdots & \cdots & 0 & t_{nn} - x \end{pmatrix} \\ &= (-1)^n \cdot (x - t_{11}) \cdot (x - t_{22}) \cdot \dots \cdot (x - t_{nn}) \end{aligned}$$

Damit sind auch die charakteristischen Polynome von Diagonalmatrizen, Jordan-Blöcken oder unteren Dreiecksmatrizen (siehe Unterkapitel 14.2) bekannt.

## 14.4 Eigenwerte

### 14.4.1 Definition: Eigenwert, algebraische Vielfachheit

Es sei  $A \in \mathbb{K}^{n \times n}$  eine quadratische Matrix.

- Eine Zahl  $\lambda \in \mathbb{K}$  heißt *Eigenwert* von  $A$ , wenn eine der beiden folgenden Bedingungen erfüllt ist:
  - (A)  $\lambda$  ist Nullstelle des charakteristischen Polynoms von  $A$ :  $\chi_A(\lambda) = 0$ .
  - (B) Die Matrix  $A - \lambda$  ist nicht regulär.
- Die Vielfachheit als Nullstelle in  $\chi_A$  heißt die *algebraische Vielfachheit* (des Eigenwerts  $\lambda$  bzgl. der Matrix  $A$ ).
- Die Menge aller Eigenwerte von  $A$  heißt das *Spektrum* von  $A$ :

$$\Sigma_A := \{\lambda \in \mathbb{K} \mid \chi_A(\lambda) = 0\}.$$

Die Äquivalenz  $(A) \Leftrightarrow (B)$  folgt durch Anwendung von Satz 10.1.1 auf die Matrix  $A - \lambda$ .

## 14.5 Der Satz von Cayley-Hamilton

### 14.5.1 Beispiel

Wir kommen auf die skurrile Idee, die  $2 \times 2$ -Matrix  $A$  aus Beispiel 14.3.3 in ihr eigenes charakteristisches Polynom einzusetzen:

$$\begin{aligned} \chi_A(A) &= A^2 - (a+d)A + (ad-bc)I \\ &= \begin{pmatrix} a^2+bc & (a+d)b \\ (a+d)c & d^2+bc \end{pmatrix} - \begin{pmatrix} (a+d)a & (a+d)b \\ (a+d)c & (a+d)d \end{pmatrix} + \begin{pmatrix} ad-bc & 0 \\ 0 & ad-bc \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

Es kommt die Null-Matrix heraus.

### 14.5.2 Satz von Cayley-Hamilton

Setzt man eine quadratische Matrix  $A$  in ihr eigenes charakteristisches Polynom ein, so erhält man die Null-Matrix:

$$\chi_A(A) = 0.$$

### 14.5.3 Beweis

(1) Wir betrachten die „Meta-Matrix“

$$\widehat{A} := \begin{pmatrix} a_{11}I - A & a_{21}I & \cdots & \cdots & a_{n1}I \\ a_{12}I & a_{22}I - A & \cdots & \cdots & a_{n2}I \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ a_{1n}I & a_{2n}I & \cdots & \cdots & a_{nn}I - A \end{pmatrix} \in (\mathbb{K}[A])^{n \times n},$$

deren Einträge Grad-Eins-Polynome in  $A$  sind.  $I$  ist dabei die  $n \times n$ -Einheitsmatrix. Es ist Absicht, dass im Eintrag an der Position  $(j, k)$  der Skalar  $a_{kj}$  auftritt.

Im folgenden wollen wir den Determinanten-Kalkül des Kapitels 9 auf diese Matrix anwenden. Man überzeuge sich davon, dass dies auch möglich ist, wenn der zugrundeliegende Körper  $\mathbb{K}$  durch einen kommutativen Ring, hier eben  $\mathbb{K}[A]$ , ersetzt wird.

(2) Wir schreiben die Matrix  $\widehat{A}$  etwas anders

$$\widehat{A} := \begin{pmatrix} a_{11}I & a_{21}I & \cdots & \cdots & a_{n1}I \\ a_{12}I & a_{22}I & \cdots & \cdots & a_{n2}I \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ a_{1n}I & a_{2n}I & \cdots & \cdots & a_{nn}I \end{pmatrix} - A \begin{pmatrix} I & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & I \end{pmatrix}$$

und sehen daraus, dass

$$\det \widehat{A} = \chi_A(x) \Big|_{x=A}.$$

(3) Wir wenden den Satz 9.5.2 über die Kofaktor-Matrix an und erhalten die folgende Gleichung in  $(\mathbb{K}[A])^{n \times n}$

$$\widehat{A}^{\circ T} \cdot \widehat{A} = \begin{pmatrix} \det \widehat{A} & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \det \widehat{A} & \ddots & & & \vdots \\ \vdots & \ddots & \det \widehat{A} & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & \det \widehat{A} \end{pmatrix}.$$

(4) Wir fassen jetzt die Matrix  $\widehat{A}$  als  $n^2 \times n^2$ -Matrix mit Einträgen aus  $\mathbb{K}$  auf. Wir wenden sie auf den Spaltenvektor

$$e = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} \in \mathbb{K}^{n^2}$$

an, in dem die kanonischen Einheitsvektoren  $e_j \in \mathbb{K}^n$  untereinander stehen. Es gilt dann in  $\mathbb{K}^{(n^2)}$

$$\widehat{A} \cdot e = \begin{pmatrix} a_{11}e_1 - Ae_1 + a_{21}e_2 + \cdots + a_{n1}e_n \\ a_{12}e_1 + a_{22}e_2 - Ae_2 + \cdots + a_{n2}e_n \\ \vdots \\ \vdots \\ a_{1n}e_1 + a_{2n}e_2 + \cdots + a_{nn}e_n - Ae_n \end{pmatrix} = \begin{pmatrix} \vec{0} \\ \vec{0} \\ \vdots \\ \vdots \\ \vec{0} \end{pmatrix}.$$

(5) Jetzt können wir zusammenfassen und die folgende Berechnung im  $\mathbb{K}^{(n^2)}$  durchführen:

$$\begin{pmatrix} \chi_A(A)e_1 \\ \chi_A(A)e_2 \\ \vdots \\ \chi_A(A)e_n \end{pmatrix} \stackrel{(2)}{=} \begin{pmatrix} \det \widehat{A} & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \det \widehat{A} & \ddots & & & \vdots \\ \vdots & \ddots & \det \widehat{A} & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & \det \widehat{A} \end{pmatrix} \cdot e$$

$$\stackrel{(3)}{=} \widehat{A}^{\circ T} \cdot \widehat{A} \cdot e \stackrel{(4)}{=} \widehat{A}^{\circ T} \cdot \begin{pmatrix} \vec{0} \\ \vec{0} \\ \vdots \\ \vec{0} \end{pmatrix} = \begin{pmatrix} \vec{0} \\ \vec{0} \\ \vdots \\ \vec{0} \end{pmatrix}.$$

(6) Die Matrix  $\chi_A(A) \in \mathbb{K}^{n \times n}$  bildet also alle Einheitsvektoren  $e_j \in \mathbb{K}^n$  auf den Nullvektor ab. Das kann aber nur so sein, wenn sie die Nullmatrix ist:

$$\chi_A(A) = 0.$$

**14.5.4 Folgerung** Es sei  $A \in \mathbb{K}^{n \times n}$ .

- (i) Das charakteristische Polynom von  $A$  ist im Annihilator von  $A$  enthalten:  $\chi_A \in \mathcal{N}_A$ .
- (ii) Das Minimalpolynom von  $A$  ist ein Teiler des charakteristischen Polynoms von  $A$ .
- (iii) Das Minimalpolynom von  $A$  hat einen Grad  $\leq n$ .
- (iv) Das charakteristische Polynom von  $A$  ist ein Teiler der  $n$ -ten Potenz des Minimalpolynoms von  $A$ .

Die Aussage (iv) ist keine Folgerung, sie bedarf eines eigenen Beweises, den wir aber hier nicht präsentieren. Siehe [Fischer, S. 257].

## 14.6 Ähnlichkeit von quadratischen Matrizen

### 14.6.1 Definition und Satz: Ähnliche Matrizen

Zwei Matrizen  $A, \tilde{A} \in \mathbb{K}^{n \times n}$  heißen *ähnlich* oder *konjugiert* zueinander, symbolisch  $A \sim \tilde{A}$ , wenn eine der folgenden äquivalenten Bedingungen erfüllt ist:

(A) Es gibt eine reguläre Matrix  $T \in \mathbb{K}^{n \times n}$ , so dass

$$\tilde{A} = T \cdot A \cdot T^{-1}.$$

(B) Es gibt eine reguläre Matrix  $T \in \mathbb{K}^{n \times n}$ , so dass

$$\tilde{A} = T^{-1} \cdot A \cdot T.$$

Sind zwei quadratische Matrizen ähnlich, so sind sie auch äquivalent zueinander, vgl. die Definition in Abschnitt 12.3.

### 14.6.2 Beobachtung

(1) Ist  $p \in \mathbb{K}[x]$  ein Polynom, so überträgt sich die Ähnlichkeit auf Matrix-Polynome:

$$\tilde{A} \sim A \implies p(\tilde{A}) \sim p(A).$$

Es ist nämlich für  $\ell \in \mathbb{N}_0$

$$\begin{aligned} (\tilde{A})^\ell &= (T A T^{-1})^\ell = (T A T^{-1})(T A T^{-1}) \cdot \dots \cdot (T A T^{-1}) \\ &= T A^\ell T^{-1} \end{aligned}$$

und dann

$$\begin{aligned} p(\tilde{A}) &= p_0 + p_1 \tilde{A} + \dots + p_n \tilde{A}^n = \\ &= p_0 + p_1 T A T^{-1} + \dots + p_n T A^n T^{-1} \\ &= T p_0 T^{-1} + T p_1 A T^{-1} + \dots + T p_n A^n T^{-1} \\ &= T p(A) T^{-1}. \end{aligned}$$

(2) Insbesondere ist eine Matrix ähnlich zu einem Vielfachen der Einheitsmatrix genau dann, wenn sie gleich diesem Vielfachen der Einheitsmatrix ist:

$$A \sim \alpha \cdot I \implies A = \alpha I.$$

Unabhängig davon, dass dies aus (1) folgt, ist in diesem Fall nämlich

$$A = T(\alpha I)T^{-1} = \alpha T I T^{-1} = \alpha I.$$

### 14.6.3 Satz: Ähnlichkeit ist eine Äquivalenzrelation auf $\mathbb{K}^{n \times n}$

- (i) Jede Matrix  $A$  ist zu sich selbst ähnlich (Reflexivität).
- (ii) Gilt  $A \sim \tilde{A}$ , so gilt auch  $\tilde{A} \sim A$  (Symmetrie).
- (iii) Es gilt für drei Matrizen  $A, \tilde{A}, \hat{A}$  die Transitivität

$$\text{WENN } A \sim \tilde{A} \text{ UND } \tilde{A} \sim \hat{A}, \quad \text{DANN } A \sim \hat{A}.$$

#### 14.6.4 Beweis

(i) und (ii) sind einfach.

Zu (iii) Es seien  $A \sim \tilde{A}$  und  $\tilde{A} \sim \hat{A}$ , es existieren also reguläre Matrizen  $S, T$  mit

$$\tilde{A} = S A S^{-1} \quad \text{und} \quad \hat{A} = T \tilde{A} T^{-1}.$$

Dann gilt aber

$$\hat{A} = T \tilde{A} T^{-1} = T S A S^{-1} T^{-1} = (TS) A (TS)^{-1}.$$

#### 14.6.5 Satz: Eigenschaften ähnlicher Matrizen I

Sind  $A, \tilde{A}$  zueinander ähnliche Matrizen, so stimmen überein:

- (i) der Rang
- (ii) die Determinante
- (iii) die Spur
- (iv) das charakteristische Polynom
- (v) das Spektrum
- (vi) das Minimalpolynom

#### 14.6.6 Beweis

(i) wurde schon in Abschnitt 12.3.4 gezeigt.

(iv) wird einfach und sorgfältig mit Hilfe der Eigenschaften der Determinante nachgerechnet:

$$\begin{aligned} \chi_{\tilde{A}}(x) &= \det(\tilde{A} - x) = \det T \cdot \det(\tilde{A} - x) \cdot \det T^{-1} \\ &= \det \left( T (\tilde{A} - x) T^{-1} \right) = \det(T \tilde{A} T^{-1} - x) \\ &= \det(A - x) = \chi_A(x). \end{aligned}$$

Da die Determinante als Grad-0-Koeffizient und die Spur als Grad- $(n-1)$ -Koeffizient im charakteristischen Polynom auftreten, folgen auch (ii) und (iii).

(v) folgt sofort aus (iv).

(vi) Aufgrund der Beobachtung in Abschnitt 14.6.2 gilt für ein beliebiges Polynom  $p \in \mathbb{K}[x]$

$$p(\tilde{A}) = 0 \iff p(A) = 0.$$

Deshalb stimmen die Annihilatoren überein,  $\mathcal{N}_{\tilde{A}} = \mathcal{N}_A$ , und dann auch deren Minimalpolynome:  $\mu_{\tilde{A}} = \mu_A$ .

## 14.7 Direkte Summe von quadratischen Matrizen

### 14.7.1 Definition

Es seien  $A' \in \mathbb{K}^{n' \times n'}$  und  $A'' \in \mathbb{K}^{n'' \times n''}$  zwei quadratische Matrizen. Man kann dann die *direkte Summe* der beiden Matrizen  $A'$  und  $A''$

$$A = \begin{pmatrix} A' & 0 \\ 0 & A'' \end{pmatrix} \in \mathbb{K}^{(n'+n'') \times (n'+n'')}$$

bilden. Die zugehörige lineare Abbildung ist

$$\ell_A : \begin{cases} \mathbb{K}^{n'+n''} & \rightarrow \mathbb{K}^{n'+n''} \\ \begin{pmatrix} v' \\ v'' \end{pmatrix} & \mapsto \begin{pmatrix} A'v' \\ A''v'' \end{pmatrix}. \end{cases}$$

### 14.7.2 Beobachtung

Ist  $p \in \mathbb{K}[x]$  ein Polynom, so gilt bei direkter Summe

$$p(A) = p\left(\begin{pmatrix} A' & 0 \\ 0 & A'' \end{pmatrix}\right) = \begin{pmatrix} p(A') & 0 \\ 0 & p(A'') \end{pmatrix}.$$

### 14.7.3 Satz

Ist die quadratische Matrix  $A$  direkte Summe der beiden Matrizen  $A'$  und  $A''$ , so gilt

- (i)  $\text{rang } A = \text{rang } A' + \text{rang } A''$
- (ii)  $\det A = \det A' \cdot \det A''$
- (iii)  $\text{spur } A = \text{spur } A' + \text{spur } A''$
- (iv)  $\chi_A = \chi_{A'} \cdot \chi_{A''}$
- (v)  $\Sigma_A = \Sigma_{A'} \cup \Sigma_{A''}$
- (vi)  $\mu_{A'}, \mu_{A''}$  sind Teiler von  $\mu_A$ .

### 14.7.4 Beweis

(i) Ist  $r'$  bzw.  $r''$  die Anzahl der linear unabhängigen Spaltenvektoren von  $A'$  bzw.  $A''$ , so ist leicht nachzuprüfen, dass  $r' + r''$  die Anzahl der linear unabhängigen Spaltenvektoren von  $A$  ist.

(ii) Setzt man zunächst  $A'' = I$  (Einheitsmatrix), so sieht man mit Hilfe der Determinanten-Entwicklung, dass

$$\det A = 1 \cdot \dots \cdot 1 \cdot \det A' = \det A'.$$

Weiter gilt aufgrund des Determinanten-Multiplikationssatzes im allgemeinen Fall

$$\begin{aligned} \det A &= \det \begin{pmatrix} A' & 0 \\ 0 & A'' \end{pmatrix} = \det \left( \begin{pmatrix} A' & 0 \\ 0 & I \end{pmatrix} \cdot \begin{pmatrix} I & 0 \\ 0 & A'' \end{pmatrix} \right) \\ &= \det \begin{pmatrix} A' & 0 \\ 0 & I \end{pmatrix} \cdot \det \begin{pmatrix} I & 0 \\ 0 & A'' \end{pmatrix} = \det A' \cdot \det A''. \end{aligned}$$



(iii) folgt direkt aus der Definition. Man kann dies aber auch mit Hilfe der Tatsache nachprüfen, dass die Spur der Koeffizient zum Grad  $n - 1$  im charakteristischen Polynom ist und dabei die Eigenschaft (iv) benutzen.

(iv) Es ist wegen (ii)

$$\begin{aligned}\chi_A(x) &= \det \begin{pmatrix} A' - x & 0 \\ 0 & A'' - x \end{pmatrix} \\ &\stackrel{(ii)}{=} \det(A' - x) \cdot \det(A'' - x) = \chi_{A'}(x) \cdot \chi_{A''}(x).\end{aligned}$$

(v) folgt sofort aus (iv).

(vi) Ist ein Polynom  $p \in \mathcal{N}_A$  gegeben, so gilt

$$\begin{pmatrix} p(A') & 0 \\ 0 & p(A'') \end{pmatrix} = p(A) = 0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

also  $p(A') = 0$  und damit  $p \in \mathcal{N}_{A'}$ .

Dies zeigt insgesamt  $\mathcal{N}_A \subseteq \mathcal{N}_{A'}$ .

Das Minimalpolynom  $\mu_{A'}$  teilt jedes Polynom in  $\mathcal{N}_{A'}$ , damit auch jedes Polynom in  $\mathcal{N}_A$ , insbesondere das Minimalpolynom  $\mu_A$ .

## 15 Endomorphismen

### 15.1 Endomorphismen beim Basiswechsel

#### 15.1.1 Beispiel

Es sei  $f : V \rightarrow V$  ein Endomorphismus eines 2-dimensionalen Vektorraums  $V$ . Bzgl. einer geordneten Basis  $B = (v_1, v_2)$  gelte

$$\begin{aligned} f(v_1) &= \frac{16}{3}v_1 - 5v_2 \\ f(v_2) &= \frac{10}{3}v_1 - 3v_2, \end{aligned}$$

so dass  $f$  bzgl. dieser geordneten Basis die Matrix

$$A = \begin{pmatrix} \frac{16}{3} & \frac{10}{3} \\ -5 & -3 \end{pmatrix}.$$

zugeordnet wird.

#### 15.1.2 Fragen

- Wie kann man der Matrix  $A$  die wesentlichen geometrischen Eigenheiten von  $f$  ansehen?
- Wie kann man wesentliche algebraische Manipulationen mit der Matrix  $A$  besser durchführen? Beispiele sind ...
  - Berechnung von Potenzen  $A^2$ ,  $A^5$  oder Matrix-Polynomen  $p(A)$ .
  - Entscheidung über oder Berechnung von (zweiten, dritten,  $n$ -ten) Wurzeln der Matrix  $A$ .
  - Einsetzen von  $A$  in Potenzreihen (Exponentialfunktion, Sinusfunktion, ...) und Berechnen des Wertes.
- Wie kann die Matrix in kleinere oder einfachere „Bestandteile“ zerlegt werden?

Den Rahmen für die Beantwortung dieser Fragen bildet wieder der Basiswechsel, jetzt aber unter einer heftig einschränkenden Bedingung.

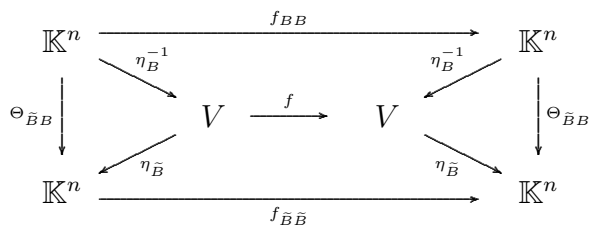
#### 15.1.3 Übereinstimmende Basiswechsel

Es seien  $B$  und  $\tilde{B}$  zwei geordnete Basen des  $n$ -dimensionaler Vektorraums  $V$ .

Die Überlegungen aus Kapitel 12 wollen wir jetzt unter dem sinnvollen, aber einengenden, Gesichtspunkt betrachten, dass die Basiswechsel im Definitionsmenge-Vektorraum und im Wertemenge-Vektorraum übereinstimmen.

### 15.1.4 Diagramm

Das Diagramm aus Abschnitt 12.2.6 hat jetzt die Form:



Bei vorgegebener geordneter Basis  $B = (v_1, \dots, v_n)$  von  $V$  wird der Abbildung  $f$  die quadratische Matrix

$$f_{BB} = \eta_B \circ f \circ \eta_B^{-1}$$

zugeordnet. Dabei umfasst  $\eta_B(v)$  die Koeffizienten, die bei Entwicklung von  $v$  bzgl. der geordneten Basis  $B$  auftreten:

$$\eta_B(v) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, \quad \text{wobei } v = \alpha_1 v_1 + \dots + \alpha_n v_n.$$

### 15.1.5 Matrix bei Basiswechsel

Bei einem durch die Matrix  $\Theta_{\tilde{B}B}$  vermittelten Basiswechsel von  $B$  nach  $\tilde{B}$  geht die Matrix  $f_{BB}$  in die Matrix  $f_{\tilde{B}\tilde{B}}$  über, wobei

$$f_{\tilde{B}\tilde{B}} = \Theta_{\tilde{B}B} \cdot f_{BB} \cdot \Theta_{B\tilde{B}} = \Theta_{\tilde{B}B} \cdot f_{\tilde{B}\tilde{B}} \cdot \Theta_{\tilde{B}B}^{-1}.$$

### 15.1.6 Zur Erinnerung (vgl. Satz 12.1.5):

- In der  $j$ -ten Spalte von

$$\Theta_{\tilde{B}B} = \begin{pmatrix} \vartheta_{11} & \cdots & \vartheta_{1j} & \cdots & \vartheta_{1n} \\ \vdots & & \vdots & & \vdots \\ \vartheta_{n1} & \cdots & \vartheta_{nj} & \cdots & \vartheta_{nn} \end{pmatrix}$$

stehen die Entwicklungskoeffizienten des  $B$ -Basisvektors  $v_j$  bzgl. der Basis  $\tilde{B}$ , das heißt

$$v_j = \vartheta_{1j} \tilde{v}_1 + \cdots + \vartheta_{nj} \tilde{v}_n.$$

- In der  $j$ -ten Spalte von

$$\Theta_{\tilde{B}B}^{-1} = \begin{pmatrix} \theta_{11} & \cdots & \theta_{1j} & \cdots & \theta_{1n} \\ \vdots & & \vdots & & \vdots \\ \theta_{n1} & \cdots & \theta_{nj} & \cdots & \theta_{nn} \end{pmatrix}$$

stehen die Entwicklungskoeffizienten des  $\tilde{B}$ -Basisvektors  $\tilde{v}_j$  bzgl. der Basis  $B$ , das heißt

$$\tilde{v}_j = \theta_{1j} v_1 + \cdots + \theta_{nj} v_n.$$

**15.1.7 Zurück** zum Beispiel 15.1.1 eines Endomorphismus  $f$ , der bzgl. der Basis  $B = \{v_1, v_2\}$  durch die Matrix

$$A = \begin{pmatrix} \frac{16}{3} & \frac{10}{3} \\ -5 & -3 \end{pmatrix}.$$

dargestellt wird.

Wir wählen eine andere geordnete Basis  $\tilde{B} = (\tilde{v}_1, \tilde{v}_2)$  von  $V = \mathbb{K}^2$

$$\tilde{v}_1 = v_1 - v_2, \quad \tilde{v}_2 = -2v_1 + 3v_2,$$

das entspricht einem Basiswechsel mit den Transformationsmatrizen

$$\Theta_{\tilde{B}B}^{-1} = \begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix}, \quad \Theta_{\tilde{B}B} = \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}.$$

Es stellt sich bei Berechnung der Bilder heraus:

$$\begin{aligned} f(\tilde{v}_1) &= f(v_1 - v_2) = \left(\frac{16}{3}v_1 - 5v_2\right) - \left(\frac{10}{3}v_1 - 3v_2\right) = 2v_1 - 2v_2 = 2\tilde{v}_1 \\ f(\tilde{v}_2) &= f(-2v_1 + 3v_2) = -2\left(\frac{16}{3}v_1 - 5v_2\right) + 3\left(\frac{10}{3}v_1 - 3v_2\right) = \\ &= -\frac{2}{3}v_1 + v_2 = \frac{1}{3}(-2v_1 + 3v_2) = \frac{1}{3}\tilde{v}_2 \end{aligned}$$

Bezüglich der neuen Basis wird der Abbildung  $f$  also die Matrix

$$\tilde{A} = f_{\tilde{B}\tilde{B}} = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{3} \end{pmatrix}$$

zugeordnet. Man kann testen, dass

$$\tilde{A} = \Theta_{\tilde{B}B}^{-1} \cdot A \cdot \Theta_{\tilde{B}B} \quad \text{bzw.} \quad A = \Theta_{\tilde{B}B}^{-1} \cdot \tilde{A} \cdot \Theta_{\tilde{B}B}$$

$\tilde{A}$  und  $A$  sind also ähnlich zueinander.

### 15.1.8 Antworten

Jetzt können die in Abschnitt 15.1.2 aufgeworfenen Fragen angegangen werden.

- $f$  ist eine Streckung mit Faktor 2 in  $\tilde{w}_1$ -Richtung, eine Stauchung (mit Faktor  $\frac{1}{3}$ ) in  $\tilde{w}_2$ -Richtung.
- Wir können eine Matrix-Potenz  $\tilde{A}^\ell$  viel leichter ausrechnen:

$$\tilde{A}^\ell = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{3} \end{pmatrix}^\ell = \begin{pmatrix} 2^\ell & 0 \\ 0 & \frac{1}{3^\ell} \end{pmatrix}$$

Dann gilt aber auch

$$\begin{aligned} A^\ell &= (\Theta_{\tilde{B}B}^{-1} \cdot \tilde{A} \cdot \Theta_{\tilde{B}B})^\ell \\ &= (\Theta_{\tilde{B}B}^{-1} \cdot \tilde{A} \cdot \Theta_{\tilde{B}B}) \cdot (\Theta_{\tilde{B}B}^{-1} \cdot \tilde{A} \cdot \Theta_{\tilde{B}B}) \cdot \dots \cdot (\Theta_{\tilde{B}B}^{-1} \cdot \tilde{A} \cdot \Theta_{\tilde{B}B}) \\ &= \Theta_{\tilde{B}B}^{-1} \cdot \tilde{A}^\ell \cdot \Theta_{\tilde{B}B} \\ &= \begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 2^\ell & 0 \\ 0 & \frac{1}{3^\ell} \end{pmatrix} \cdot \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 3 \cdot 2^\ell - 2^{\ell+2} & 2^{\ell+1} - \frac{2}{3^\ell} \\ -3 \cdot 2^\ell + \frac{1}{3^{\ell-1}} & -2^{\ell+1} + \frac{1}{3^{\ell-1}} \end{pmatrix} \end{aligned}$$

- Eine Wurzel der Matrix  $\tilde{A}$  ist gegeben durch

$$\tilde{S} = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \frac{1}{\sqrt{3}} \end{pmatrix}.$$

Es ist dann auch

$$\begin{aligned} S &:= \Theta_{\tilde{B}B}^{-1} \cdot \tilde{S} \cdot \Theta_{\tilde{B}B} \\ &= \Theta_{\tilde{B}B}^{-1} \cdot \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \frac{1}{\sqrt{3}} \end{pmatrix} \cdot \Theta_{\tilde{B}B} = \begin{pmatrix} 3\sqrt{2} - \frac{2}{\sqrt{3}} & 2\sqrt{2} - \frac{2}{\sqrt{3}} \\ -3\sqrt{2} + \sqrt{3} & -2\sqrt{2} + \sqrt{3} \end{pmatrix} \end{aligned}$$

eine Wurzel von  $A$ . Tatsächlich gilt:

$$\begin{aligned} S \cdot S &= \Theta_{\tilde{B}B}^{-1} \cdot \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \frac{1}{\sqrt{3}} \end{pmatrix} \cdot \Theta_{\tilde{B}B} \cdot \Theta_{\tilde{B}B}^{-1} \cdot \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \frac{1}{\sqrt{3}} \end{pmatrix} \cdot \Theta_{\tilde{B}B} \\ &= \Theta_{\tilde{B}B}^{-1} \cdot \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{3} \end{pmatrix} \cdot \Theta_{\tilde{B}B} = A. \end{aligned}$$

**15.1.9 Der Clou** bei den obigen Berechnungen ist, dass die zur Matrix  $A$  ähnliche Matrix  $\tilde{A}$  bzgl. der veränderten Basis  $\tilde{B}$  eine Diagonalmatrix ist.

Diagonalmatrizen sind viel einfacher zu lesen, zu interpretieren, mit ihnen kann man wesentlich besser rechnen. Matrix-Polynome von Diagonalmatrizen lassen sich viel leichter ausrechnen.

In diesem Kapitel wollen wir weitere Fragen in diesem Zusammenhang beantworten:

- Lassen sich alle Endomorphismen bzw. quadratische Matrizen in Diagonalgestalt transformieren?
- Wenn nicht, gibt es abgewandelte „einfache“ Formen von Matrizen, so genannte Normalformen, für die Transformation?
- Wie kann man diese Normalformen finden?
- Wie lässt sich die neue Basis finden, bzgl. derer der Endomorphismus bzw. die quadratische Matrix Normalform annimmt?

## 15.2 Invarianten von Endomorphismen

### 15.2.1 Einstieg

Es sei  $V$  ein endlich-dimensionaler  $\mathbb{K}$ -Vektorraum.

Wir haben gesehen, dass zwei Matrizen  $A, \tilde{A} \in \mathbb{K}^{n \times n}$  genau dann ähnlich sind, wenn sie bei geeigneter Wahl der geordneten Basen  $B$  bzw.  $\tilde{B}$  von  $V$  den gleichen Endomorphismus  $f : V \rightarrow V$  darstellen. Dies führt auf die folgenden

### 15.2.2 Definitionen

Es sei  $f : V \rightarrow V$  ein Endomorphismus,  $B$  eine geordnete Basis von  $V$  und  $A = f_{BB} \in \mathbb{K}^{n \times n}$  eine Matrix, die diesen Endomorphismus bzgl. der geordneten Basis  $B$  repräsentiert.

Wir definieren die folgenden Invarianten des Endomorphismus  $f$ :

$$\begin{aligned} \det f &:= \det A \\ \chi_f &:= \chi_A \\ \Sigma_f &:= \Sigma_A \\ \text{spur } f &:= \text{spur } A \\ \mu_f &:= \mu_A \end{aligned}$$

### 15.2.3 Die Algebra $\mathbb{K}[f]$ eines Endomorphismus

Ist  $f : V \rightarrow V$  ein Endomorphismus, so definieren wir rekursiv die Potenzen von  $f$  durch Iteration

$$f^0 := \text{id}_V, \quad f^1 := f \quad f^\ell := f \circ f^{\ell-1}.$$

Für ein Polynom  $p \in \mathbb{K}[x]$  mit

$$p(x) = p_n x^n + \dots + p_1 x + p_0$$

ist dann auch der Endomorphismus  $p(f) = p_n f^n + \dots + p_1 f + p_0$  durch

$$p(f) : \begin{cases} V & \rightarrow V \\ v & \mapsto p_n f^n(v) + \dots + p_1 f(v) + p_0 v \end{cases}$$

wohldefiniert.

### 15.2.4 Bemerkung

Alternativ können alle diese Invarianten auch ohne Bezugnahme auf die Darstellung durch Matrizen mit Hilfe abstrakterer (multi-)linearer Algebra definiert werden.

Konkreter und naheliegend ist beispielsweise, dass man den Annihilator eines Endomorphismus  $f : V \rightarrow V$  auch definieren kann als

$$\mathcal{N}_f := \left\{ p \in \mathbb{K}[x] \mid p(f) = 0 \right\}.$$

Selbstverständlich ist diese Menge ein Ideal und das Minimalpolynom von  $f$  kann dann als das normierte Polynom in  $\mathcal{N}_f$  definiert werden, das minimalen Grad hat oder/und Teiler aller anderen Polynome in  $\mathcal{N}_f$  ist.

Überlegen Sie, dass die beiden Definitionen von „Minimalpolynom eines Endomorphismus“ auf dieser Seite äquivalent sind.

## 15.3 Direkte Summe von Endomorphismen

### 15.3.1 Definition

Es seien  $V$  und  $V'$  zwei  $\mathbb{K}$ -Vektorräume der Dimension  $n'$  bzw.  $n''$  und  $f' : V' \rightarrow V'$ ,  $f'' : V'' \rightarrow V''$  zwei Endomorphismen. Dann ist durch

$$f' \oplus f'' : \begin{cases} V' \oplus V'' & \rightarrow V' \oplus V'' \\ (v', v'') & \mapsto (f'(v'), f''(v'')) \end{cases}$$

ein Endomorphismus definiert, man nennt ihn die direkte Summe der beiden Endomorphismen  $f'$  und  $f''$ .

Es seien  $B' = (v'_1, \dots, v'_{n'})$  eine geordnete Basis von  $V'$  und  $B'' = (v''_1, \dots, v''_{n''})$  eine geordnete Basis von  $V''$ . Dann ist

$$B = \left( (v'_1, \vec{0}), \dots, (v'_{n'}, \vec{0}), (\vec{0}, v''_1), \dots, (\vec{0}, v''_{n''}) \right)$$

eine geordnete Basis von  $V' \oplus V''$ .

Jetzt seien  $A'$ ,  $A''$  die quadratischen Matrizen, die  $f'$  bzw.  $f''$  bzgl. der Basen  $B'$  bzw.  $B''$  darstellen. Der Endomorphismus  $f' \oplus f''$  hat dann bzgl. der Basis  $B$  die Darstellung

$$(f' \oplus f'')_{BB} = \begin{pmatrix} f'_{B'B'} & 0 \\ 0 & f''_{B''B''} \end{pmatrix}.$$

Auf diese Weise lässt sich der Satz 15.3.2 auch auf direkte Summen von Endomorphismen übertragen.

### 15.3.2 Satz

Ist der Endomorphismus  $f$  direkte Summe der beiden Endomorphismen  $f'$  und  $f''$ , so gilt

- (i)  $\text{rang } f = \text{rang } f' + \text{rang } f''$
- (ii)  $\det f = \det f' \cdot \det f''$
- (iii)  $\text{spur } f = \text{spur } f' + \text{spur } f''$
- (iv)  $\chi_f = \chi_{f'} \cdot \chi_{f''}$
- (v)  $\Sigma_f = \Sigma_{f'} \cup \Sigma_{f''}$
- (vi)  $\mu_{f'}$ ,  $\mu_{f''}$  sind Teiler von  $\mu_f$ .

### 15.3.3 $f$ -invariante Zerlegung

Es sei  $f : V \rightarrow V$  ein Endomorphismus. Ist dann

$$V = V' \oplus V''$$

eine direkte Summe mit der Eigenschaft, dass die Abbildungen

$$f' : \begin{cases} V' & \rightarrow V' \\ v & \mapsto f(v) \end{cases} \quad f'' : \begin{cases} V'' & \rightarrow V'' \\ v & \mapsto f(v) \end{cases}$$

wohldefiniert sind, so nennen wir die innere Summe eine  $f$ -invariante Zerlegung. Es sind dann  $f'$  und  $f''$  wieder Endomorphismen und

$$f = f' \oplus f''.$$

## 16 Nilpotente Endomorphismen

### 16.1 Young-Tableaus

#### 16.1.1 Definition

Wir nennen eine Matrix  $Y = (y_{jk}) \in \mathbb{K}^{M \times N}$ , ein *Young-Tableau*, wenn die folgenden Bedingungen erfüllt sind:

$$\begin{aligned} y_{jk} &\in \{0, 1\} \quad \text{für alle } (j, k) \\ y_{1k} &= 1 \quad \text{für alle } k = 1, \dots, N \\ y_{j1} &= 1 \quad \text{für alle } j = 1, \dots, M \\ y_{jk} &= 1 \implies y_{\tilde{j}\tilde{k}} = 1 \quad \text{für alle } (\tilde{j}, \tilde{k}) \text{ mit } \tilde{j} < j \text{ oder } \tilde{k} < k. \end{aligned}$$

Mit anderen Worten: Das Young-Tableau enthält nur Einsen und Nullen. Erste Zeile und erste Spalte enthalten nur Einsen. Links und oberhalb von Einsen dürfen nur Einsen stehen.

#### 16.1.2 Graphik

Es ist allgemein üblich, ein Young-Tableau nicht in Matrix-Form, sondern als Schema graphisch anzugeben. Am Beispiel wird das deutlich

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \sim \begin{array}{cccccccc} \square & \square & \square & \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square & \square & \square & \square \\ \square & \square & \square & \square & \square & \square & \square & \square \end{array} \sim \begin{pmatrix} 8 \\ 8 \\ 7 \\ 4 \\ 2 \\ 1 \\ 1 \end{pmatrix} \sim (7 \ 5 \ 4 \ 4 \ 3 \ 3 \ 3 \ 2)$$

#### 16.1.3 Zeilenlängen-Vektor

Einem Young-Tableau kann man den Spaltenvektor der Zeilenlängen zuordnen:

$$\begin{pmatrix} n_1 \\ \vdots \\ n_M \end{pmatrix} \quad \text{mit} \quad n_j := \max \{k \mid y_{jk} = 1\}$$

Im Beispiel ist er bereits angegeben. Man überlege, dass  $n_1 = N$ .

#### 16.1.4 Spaltenlängen-Vektor

Dual dazu kann an einem Young-Tableau den Zeilenvektor der Spaltenlängen zuordnen:

$$(m_1 \ \cdots \ m_N) \quad \text{mit} \quad m_k := \max \{j \mid y_{jk} = 1\}$$

Im Beispiel ist er ebenfalls angegeben. Man überlege, dass  $m_1 = M$ .



### 16.1.5 Umwandlung

Das Young-Tableau kann natürlich aus den „Längen-Vektoren“ rückermittelt werden:

$$y_{jk} = 1 \iff k \leq n_j \iff j \leq m_k.$$

Mittels der Formeln

$$\begin{aligned} m_k &= \text{card} \{n_j \mid n_j \geq k\} \\ n_j &= \text{card} \{m_k \mid m_k \geq j\} \end{aligned}$$

können die Vektoren der Zeilenlängen bzw. Spaltenlängen wechselseitig direkt ineinander umgewandelt werden.

### 16.1.6 Die Ordnung

Die Ordnung eines Young-Tableaus ist definiert als die Gesamtzahl der Einsen — und damit gleich der Summe der Zeilenlängen bzw. gleich der Summe der Spaltenlängen:

$$\begin{aligned} L &= \text{card} \left\{ (j, k) \in \{1, \dots, M\} \times \{1, \dots, N\} \mid y_{jk} = 1 \right\} \\ &= n_1 + \dots + n_M \\ &= m_1 + \dots + m_N. \end{aligned}$$

## 16.2 Nilpotente Endomorphismen

In diesem Abschnitt (und in den nächsten) steht jeweils ein Endomorphismus  $f : V \rightarrow V$  im Mittelpunkt. Auch eine quadratische Matrix  $A$  kann als Endomorphismus  $\ell_A : \mathbb{K}^n \rightarrow \mathbb{K}^n$  angesehen werden.

### 16.2.1 Satz und Definition: Nilpotente Endomorphismen

Es sei  $f : V \rightarrow V$  ein Endomorphismus eines  $n$ -dimensionalen Vektorraums. Es sind die folgenden Aussagen äquivalent:

- (A)  $f$  heißt *nilpotent*.
- (B) Es existiert eine Zahl  $\ell \in \mathbb{N}$ , so dass  $f^\ell = 0$ .  
Es sei dann  $m \in \mathbb{N}$  definiert durch  $m := \min\{\ell \in \mathbb{N} \mid f^\ell = 0\}$ .
- (C) Es existiert eine Zahl  $m \in \{1, \dots, n\}$ , so dass  $\mu_f(x) = x^m$ .
- (D) Das charakteristische Polynom ist  $\chi_f(x) = (-1)^n x^n$ .
- (E) Es gibt eine geordnete Basis  $B$ , so dass die Matrix  $f_{BB} \in \mathbb{K}^{n \times n}$  eine streng-obere Dreiecksmatrix ist.
- (F) Es gibt eine geordnete Basis  $B$ , so dass die Matrix  $f_{BB} \in \mathbb{K}^{n \times n}$  eine Matrix der folgenden Form ist:

$$f_{BB} = \begin{pmatrix} J_{m_1}(0) & 0 & \cdots & \cdots & 0 \\ 0 & J_{m_2}(0) & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & J_{m_N}(0) \end{pmatrix}$$

Dabei ist die Größe der Jordan-Blöcke  $J_{m_k}(0)$  monoton abnehmend.

$$m = m_1 \geq m_2 \geq \dots \geq m_N.$$

Die in (B), (C), (F) angegebene Zahl  $m$  heißt der *Nilpotenzgrad* des Endomorphismus  $f$ .

### 16.2.2 Beweis

(B)  $\Rightarrow$  (F): Diese Implikation erfordert einen fünfseitigen Beweis, wie er im nächsten Unterkapitel 16.3 dargelegt ist.

(F)  $\Rightarrow$  (E): Diese Implikation kann der Gestalt der Matrix  $f_{BB}$  in (F) entnommen werden.

(E)  $\Rightarrow$  (D): Es ist aus Beispiel 14.3.4 bekannt, dass das charakteristische Polynom einer streng-oberen Dreiecksmatrix wie in (E) gleich  $(-1)^n x^n$  ist. Das ist dann auch das charakteristische Polynom des Endomorphismus  $f$ .

(D)  $\Rightarrow$  (C): Dies folgt direkt aus 14.5.4 (ii).

(C)  $\Rightarrow$  (B): Dies ist eine direkte Konsequenz der Definition des Minimalpolynoms, vgl. Bemerkung 15.2.4.

### 16.2.3 Beispiel

Eine nilpotente Matrix muss nicht notwendig eine strenge obere oder untere Dreiecksmatrix sein:

$$A = \begin{pmatrix} ab & a^2 \\ -b^2 & -ab \end{pmatrix} \in \mathbb{K}^{2 \times 2}.$$

In (E) steht nur, dass  $A$  ähnlich zu einer streng oberen Dreiecksmatrix sein muss.

## 16.3 Die Potenzkern-Kette

Dieses Unterkapitel ist ganz dem Beweis der Implikation (B)  $\Rightarrow$  (F) aus Satz 16.2.1 gewidmet.

Es sei also der nilpotente Endomorphismus  $f : V \rightarrow V$  mit Nilpotenzgrad  $m$  gegeben.

### 16.3.1 Die Potenzkern-Kette

Im Mittelpunkt des gesamten Beweises steht die durch Inklusion geordnete Kette der „Potenzkerne“

$$\{\vec{0}\} = \ker f^0 \subseteq \ker f^1 \subseteq \ker f^2 \subseteq \dots \subseteq \ker f^m = V.$$

### 16.3.2 Satz mit Algorithmus

Wende das folgende Verfahren an:

1. Ordne den Dimensionen der Potenzkerne so die Zeilenlängen eines Young-Diagramms zu:

$$n_j = \dim \ker f^j - \dim \ker f^{j-1}, \quad j = 1, \dots, M, \quad M := m.$$

2. Berechne die zugehörigen Spaltenlängen des Young-Diagramms

$$m_1, \dots, m_N.$$

3. Dann gibt es eine geordnete Basis  $B$  von  $V$ , bzgl. derer die Matrix  $f_{BB}$  die Form

$$f_{BB} = \begin{pmatrix} J_{m_1}(0) & 0 & \dots & \dots & 0 \\ 0 & J_{m_2}(0) & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & 0 & J_{m_N}(0) \end{pmatrix}$$

hat. Dabei sind  $J_{m_k}(0) \in \mathbb{K}^{m_k \times m_k}$  die Jordan-Blöcke aus Abschnitt 14.2.8.

Damit ist die Aussage (F) aus Satz 16.2.1 erreicht.

Wir beweisen diesen Satz in einer Abfolge von mehreren Lemmata.

### 16.3.3 Lemma: Zerlegung der Potenzkerne

Es existieren Unterräume

$$W_m, W_{m-1}, \dots, W_1$$

von  $V$ , so dass jede Zeile des folgenden Diagramms wohldefiniert ist:

$$\begin{array}{rcl}
 \ker f^1 & = & \ker f^0 \oplus f^{m-1}(W_m) \oplus f^{m-2}(W_{m-1}) \oplus \dots \oplus f^2(W_3) \oplus f(W_2) \oplus W_1 \\
 \uparrow & & \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \\
 \ker f^2 & = & \ker f^1 \oplus f^{m-2}(W_m) \oplus f^{m-3}(W_{m-1}) \oplus \dots \oplus f^1(W_3) \oplus W_2 \\
 \uparrow & & \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \\
 \ker f^3 & = & \ker f^2 \oplus f^{m-3}(W_m) \oplus f^{m-4}(W_{m-1}) \oplus \dots \oplus W_3 \\
 & & \vdots \\
 & & \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \quad \uparrow \\
 \ker f^{m-1} & = & \ker f^{m-2} \oplus f(W_m) \oplus W_{m-1} \\
 \uparrow & & \uparrow \quad \uparrow \quad \uparrow \\
 \ker f^m & = & \ker f^{m-1} \oplus W_m
 \end{array}$$

Der kleine Pfeil oberhalb eines Unterraums deutet an, dass der unterhalb stehende Unterraum durch  $f$  auf den oberhalb stehenden Unterraum abgebildet wird.

Der kleine Pfeil oberhalb eines „direkten-Summe-Zeichens“ deutet an, dass die direkte-Summe-Zerlegung der unterhalb stehenden Unterräume bei Anwendung von  $f$  bestehen bleibt.

Innerhalb der  $j$ -ten Zeile setzen wir zur Abkürzung:

$$\ker f^j = \ker f^{j-1} \oplus \underbrace{f^{m-j}(W_m) \oplus f^{m-j-1}(W_{m-1}) \oplus \dots \oplus f^2(W_{j+2}) \oplus f^1(W_{j+1}) \oplus W_j}_{=:Z_j} \overset{=f(Z_{j+1})}{\quad}$$

### 16.3.4 Beweis

(0) Wir erbringen den Beweis durch Induktion über  $j = m, \dots, 1$ , also bzgl. des Diagramms von unten nach oben.

(1) Induktionsanfang in der  $m$ -ten Zeile. Es sei  $W_m$  ein Komplementärraum zu  $\ker f^{m-1}$  in  $\ker f^m$ . Dann ist die  $m$ -te Zeile wohldefiniert.

(2) Induktionsschritt: Wir betrachten und fixieren jetzt die Zeilennummer  $j \in \{1, \dots, m\}$ . Die Unterräume  $W_m, \dots, W_{j+1}$  seien definiert.

Wir zeigen zunächst — in den Schritten (3a) bis (3d) — die Aussage

$$\ker f^j \supseteq \ker f^{j-1} \oplus f^{m-j}(W_m) \oplus f^{m-j-1}(W_{m-1}) \oplus \dots \oplus f^2(W_{j+2}) \oplus f^1(W_{j+1})$$

(3a) Wir zeigen für ein  $\ell \in \{0, \dots, m - j\}$

$$f^\ell(W_{j+\ell}) \subseteq \ker f^j.$$

Es sei  $v \in f^\ell(W_{j+\ell})$ . Dann gibt ein  $w \in W_{j+\ell} \subseteq \ker f^{j+\ell}$  mit  $v = f^\ell w$  und deshalb

$$f^j v = f^j(f^\ell w) = f^{j+\ell} w = \vec{0},$$

also  $v \in \ker f^j$ .

(3b) Weiter ist

$$\ker f^{j-1} \cap f^\ell(W_{\ell+j}) = \{\vec{0}\}.$$

Für  $v \in \ker f^{j-1} \cap f^\ell(W_{\ell+j})$  gibt es nämlich  $w \in f^{\ell-1}(W_{\ell+j})$  mit  $v = f(w)$  und dann

$$f^j w = f^{j-1} f w = f^{j-1} v = \vec{0}.$$

Das bedeutet  $w \in \ker f^j \cap f^{\ell-1}(W_{\ell+j}) = \{\vec{0}\}$  nach Induktionsvoraussetzung in Zeile  $j + 1$ . Dann ist aber  $w = \vec{0}$  und  $v = f(w) = \vec{0}$ .

(3c) Die Abbildung

$$f : \begin{cases} Z_{j+1} & \rightarrow Z_j \\ v & \mapsto f(v) \end{cases}$$

ist injektiv. Es ist nämlich

$$\ker f|_{Z_{j+1}} = \ker f \cap Z_{j+1} \subseteq \ker f^j \cap Z_{j+1} = \{\vec{0}\}$$

aufgrund der Induktionsvoraussetzung in Zeile  $j + 1$ .

(3d) Wegen (3c) überträgt sich die Direkte-Summe-Zerlegung von

$$Z_{j+1} = f^{m-j-1}(W_m) \oplus f^{m-j-2}(W_{m-1}) \oplus \dots \oplus f^2(W_{j+3}) \oplus f(W_{j+2}) \oplus W_{j+1}$$

innerhalb der Zeile  $j + 1$  auf das Bild

$$f(Z_{j+1}) = f^{m-j}(W_m) \oplus f^{m-j-1}(W_{m-1}) \oplus \dots \oplus f^3(W_{j+3}) \oplus f^2(W_{j+2}) \oplus f(W_{j+1}).$$

innerhalb der Zeile  $j$ .

(4) Es wird jetzt der Unterraum  $W_j$  von  $V$  so gewählt, dass

$$\ker f^j = \ker f^{j-1} \oplus f^{m-j}(W_m) \oplus f^{m-j-1}(W_{m-1}) \oplus \dots \oplus f^2(W_{j+2}) \oplus f(W_{j+1}) \oplus W_j.$$

### 16.3.5 Lemma: Zerlegung des gesamten Vektorraums

Mit den Bezeichnungen aus dem letzten Satz 16.3.3 ist

$$V = \left\{ \begin{array}{l} Z_1 \\ \oplus Z_2 \\ \oplus Z_3 \\ \vdots \\ \oplus Z_{m-1} \\ \oplus Z_m \end{array} \right\} = \left\{ \begin{array}{l} f^{m-1}(W_m) \oplus f^{m-2}(W_{m-1}) \oplus \dots \oplus f^2(W_3) \oplus f(W_2) \oplus W_1 \\ \oplus f^{m-2}(W_m) \oplus f^{m-3}(W_{m-1}) \oplus \dots \oplus f^1(W_3) \oplus W_2 \\ \oplus f^{m-3}(W_m) \oplus f^{m-4}(W_{m-1}) \oplus \dots \oplus W_3 \\ \vdots \\ \oplus f(W_m) \oplus W_{m-1} \\ \oplus W_m \end{array} \right\}$$

### 16.3.6 Beweis

Betrachte das Diagramm in Satz 16.3.3. Setze dann sukzessive von  $j = 1$  bis  $m - 1$  die  $j$ -te Zeile in die  $(j + 1)$ -te Zeile ein.

### 16.3.7 Lemma: Geordnete Basis von $V$

Es existiert eine geordnete Basis

$$B = \left( v_{11}, \dots, v_{M, n_M} \right),$$

deren  $n$  Vektoren  $v_{jk}$  — nach spaltenweiser Zuordnung zu den Eins-Positionen  $(j, k)$  im Young-Tableau —

$$\begin{array}{cccccccccc} v_{11} & v_{12} & v_{13} & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & v_{1, n_1} \\ v_{21} & v_{22} & \cdots & \cdots & \cdots & \cdots & \cdots & v_{2, n_2} & & \\ v_{31} & v_{32} & \cdots & \cdots & \cdots & v_{3, n_3} & & & & \\ \vdots & & & & & & & & & \\ v_{m1} & v_{m2} & \cdots & v_{m, n_m} & & & & & & \end{array}$$

die folgende Eigenschaft haben:

$$f(v_{j,k}) = \begin{cases} v_{j-1,k}, & \text{falls } j \geq 2 \\ 0, & \text{falls } j = 1. \end{cases} \quad (*)$$

### 16.3.8 Beweis

Wir werden die Vektoren zeilenweise rekursiv für  $j = m, \dots, 1$ , also von unten nach oben, angeben. Sei zunächst

$$\{v_{m,1}, \dots, v_{m, n_m}\}$$

irgendeine Basis von  $Z_m$ .

Ist dann für festes  $j$  eine Basis

$$\{v_{j+1,1}, \dots, v_{j+1, n_{j+1}}\}$$

von  $Z_{j+1}$  bereits gegeben, so setzen wir

$$v_{j,k} := f(v_{j+1,k}), \quad k = 1, \dots, n_{j+1}.$$

Wegen der Bijektivität der Abbildung  $f : Z_{j+1} \rightarrow f(Z_{j+1})$  (vgl. Schritt (3c) im Beweis 16.3.4) ist die Menge dieser Vektoren eine Basis von  $f(Z_{j+1})$ .

Wählen wir dann noch eine Basis von  $W_j$ ,

$$\{v_{j, n_{j+1}+1}, \dots, v_{j, n_j}\},$$

so ist

$$\{v_{j,1}, \dots, v_{j, n_{j+1}}, v_{j, n_{j+1}+1}, \dots, v_{j, n_j}\}$$

eine Basis von  $Z_j$ . Die im Satz geforderte Eigenschaft ist erfüllt.

Alle diese Basisvektoren können dann noch — wie im Satz angegeben — geordnet werden.

### 16.3.9 Jordan-Blöcke

Als nächstes definieren wir für jede Spalte  $k$  des Young-Tableaus einen Unterraum  $S_k$  von  $V$ , der von den Vektoren dieser Spalte erzeugt wird:

$$S_k := \langle v_{1,k}, \dots, v_{m_k,k} \rangle, \quad k = 1, \dots, N.$$

Die Gleichung (\*) aus Lemma 16.3.7 zeigt, dass der Endomorphismus  $f : S_k \rightarrow S_k$  wohldefiniert ist, und dass bei Wahl der geordneten Basis  $B_k = (v_{1,k}, \dots, v_{m_k,k})$  diesem Endomorphismus der Jordan-Block

$$f_{B_k, B_k} = J_{m_k}(0) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & 0 \end{pmatrix} \in \mathbb{K}^{m_k \times m_k}.$$

zugeordnet ist.

Damit haben wir eine  $f$ -invariante Zerlegung

$$V = S_1 \oplus \dots \oplus S_N$$

gefunden. Bei Wahl der geordneten Basis  $B$  für  $V$  aus Satz 16.3.7 nimmt der Endomorphismus  $f : V \rightarrow V$  die Gestalt an, die in Aussage (F) des Satzes 16.2.1 beschrieben ist.



## 16.4 Das Lemma von Fitting

### 16.4.1 Lemma von Fitting über die Nilpotenz-Abspaltung

Es sei  $f : V \rightarrow V$  ein Endomorphismus des  $n$ -dimensionalen  $\mathbb{K}$ -Vektorraums  $V$ .

(i) Dann gibt es eine Zahl  $m$  mit  $1 \leq m \leq n$ , so dass

$$V = \ker f^m \oplus \operatorname{im} f^m$$

eine  $f$ -invariante Zerlegung ist.

Die Einschränkung  $f'$  der Abbildung  $f$  auf  $\ker f^m$  ist nilpotent.

Die Einschränkung  $f''$  der Abbildung  $f$  auf  $\operatorname{im} f^m$  ist ein Isomorphismus.

(ii) Es ist  $\dim \ker f^m =$  Vielfachheit der Nullstelle 0 in  $\chi_f$ .

### 16.4.2 Beweis

(0) Ist  $f$  regulär, so ist der Satz mit  $m = 1$ , dann  $\ker f = \{\vec{0}\}$  und  $\operatorname{im} f = V$  trivialerweise erfüllt.

(1) Im Fall „ $f$  nicht regulär“ beruht der gesamte Beweis auf dem folgenden Diagramm:

$$\begin{array}{cccccccccccc} \{\vec{0}\} & \xleftrightarrow{\text{echt}} & \ker f^1 & \xleftrightarrow{\text{echt}} & \ker f^2 & \xleftrightarrow{\text{echt}} & \dots & \xleftrightarrow{\text{echt}} & \ker f^m & \xleftrightarrow{\text{gleich}} & \ker f^{m+1} & \xleftrightarrow{\text{gleich}} & \dots \\ V & \xleftrightarrow{\text{echt}} & \operatorname{im} f^1 & \xleftrightarrow{\text{echt}} & \operatorname{im} f^2 & \xleftrightarrow{\text{echt}} & \dots & \xleftrightarrow{\text{echt}} & \operatorname{im} f^m & \xleftrightarrow{\text{gleich}} & \operatorname{im} f^{m+1} & \xleftrightarrow{\text{gleich}} & \dots \end{array}$$

(2) Die Inklusionen im Diagramm (1) sind wohldefiniert. Für  $\ell \in \mathbb{N}$  gilt :

$$\begin{array}{l} v \in \ker f^\ell \iff f^\ell v = \vec{0} \implies f^{\ell+1} v = \vec{0} \iff v \in \ker f^{\ell+1} \\ v \in \operatorname{im} f^{\ell+1} \iff \exists_{w \in V} f^{\ell+1}(w) = v \iff \exists_{w \in V} f^\ell(f(w)) = v \implies v \in \operatorname{im} f^\ell \end{array}$$

(3) Wegen  $\dim V = n < \infty$  können die Inklusionsketten im Diagramm (1) nicht unbegrenzt anwachsen. Deshalb gibt es ein

$$m := \min\{\ell \in \mathbb{N} \mid \ker f^\ell = \ker f^{\ell+1}\}.$$

(4) Für alle  $\ell < m$  gilt dann — aufgrund der Definition

$$\ker f^\ell \subsetneq \ker f^{\ell+1}$$

(5) Für alle  $\ell \geq m$  gilt

$$\ker f^\ell = \ker f^{\ell+1}$$

Anderenfalls würde es ein  $v \in V$  geben mit

$$f^\ell(v) \neq \vec{0} \quad \text{und} \quad f^{\ell+1}(v) = \vec{0}.$$

Dies kann man umformulieren zu

$$f^m(f^{\ell-m}(v)) \neq \vec{0} \quad \text{und} \quad f^{m+1}(f^{\ell-m}(v)) = \vec{0},$$

was

$$\ker f^m \subsetneq \ker f^{m+1}$$

bedeutet, das ist ein Widerspruch zur Definition von  $m$ .

(6) Wir überlegen die folgende Kette von Äquivalenzen für  $\ell \in \mathbb{N}$ :

$$\begin{aligned} & \operatorname{im} f^\ell = \operatorname{im} f^{\ell+1} \\ \iff & \dim \operatorname{im} f^\ell = \dim \operatorname{im} f^{\ell+1} \\ \stackrel{\text{Satz 7.6.1}}{\iff} & \dim \ker f^\ell = \dim \ker f^{\ell+1} \\ \iff & \ker f^\ell = \ker f^{\ell+1} \end{aligned}$$

(7) Wegen (6) ist

$$\begin{aligned} \operatorname{im} f^\ell & \supsetneq \operatorname{im} f^{\ell+1} & \text{für } \ell < m \\ \operatorname{im} f^\ell & = \operatorname{im} f^{\ell+1} & \text{für } \ell \geq m \end{aligned}$$

Daraus folgt

$$m = \min \{ \ell \in \mathbb{N} \mid \operatorname{im} f^\ell = \operatorname{im} f^{\ell+1} \}.$$

Anders ausgedrückt: Die Gleichheit in den beiden Inklusionsketten des Diagramms (1) tritt jeweils zum ersten Mal genau an der gleichen Stelle zwischen  $m$  und  $m+1$  ein.

(8) Wir zeigen, dass  $\ker f^m \cap \operatorname{im} f^m = \{ \vec{0} \}$ .

Ist  $v \in \ker f^m \cap \operatorname{im} f^m$ , so gibt es  $w \in V$  mit

$$f^m(w) = v, \quad \text{außerdem} \quad f^m(v) = \vec{0}.$$

Daraus folgt

$$f^{2m}(w) = f^m(v) = \vec{0}$$

und deshalb

$$w \in \ker f^{2m} = \ker f^m$$

und weiter

$$v = f^m(w) = \vec{0}.$$

(9) Wegen

$$\begin{aligned} \dim(\ker f^m \oplus \operatorname{im} f^m) &= \dim \ker f^m + \dim \operatorname{im} f^m - \dim(\ker f^m \cap \operatorname{im} f^m) \\ &= n - 0 = n \end{aligned}$$

ist dann

$$\ker f^m \oplus \operatorname{im} f^m = V.$$

(10) Um die  $f$ -Invarianz zu zeigen, muss man nur genau hinschauen:

Ist  $v \in \ker f^m$ , so ist  $f(v) \in \ker f^{m-1} \subseteq \ker f^m$ .

Ist  $v \in \operatorname{im} f^m$ , so existiert  $w \in V$  mit  $f^m w = v$  und dann

$$f(v) = f^{m+1}(w) \in \operatorname{im} f^{m+1} = \operatorname{im} f^m.$$

(11) Ist  $v \in \ker f''$ , so gilt

$$v \in \operatorname{im} f^m \cap \ker f \subseteq \operatorname{im} f^m \cap \ker f^m = \{\vec{0}\}.$$

Damit ist der Endomorphismus  $f''$  injektiv, folglich auch bijektiv.

(12) Wir zeigen noch (ii): Gemäß Satz 16.2.1 (E) ist

$$\chi_{f'}(x) = x^{\dim \ker f^m}.$$

Da  $f''$  ein Isomorphismus ist, gilt  $\chi_{f''}(0) \neq 0$ . Mit  $\chi_f(x) = \chi_{f'}(x) \cdot \chi_{f''}(x)$  folgt dann die Behauptung.

## 17 Die Jordan-Normalform eines Endomorphismus

### 17.1 Die $\lambda$ -Verschiebung eines Endomorphismus

#### 17.1.1 Definition

Ist  $f : V \rightarrow V$  ein Endomorphismus und  $\lambda \in \mathbb{K}$ , so nennen wir den Endomorphismus

$$f - \lambda : \begin{cases} V & \rightarrow V \\ v & \mapsto f(v) - \lambda v \end{cases}$$

die  $\lambda$ -Verschiebung (von  $f$  um  $\lambda$ ).

Mit Hilfe der  $\lambda$ -Verschiebung können die beiden Sätze über nilpotente Endomorphismen „ $\lambda$ -verschoben“ werden.

#### 17.1.2 Satz: Nilpotenter $\lambda$ -verschobener Endomorphismus

Es sei  $f : V \rightarrow V$  ein Endomorphismus eines  $n$ -dimensionalen Vektorraums und  $\lambda \in \mathbb{K}$ . Es sind die folgenden Aussagen äquivalent:

- (A)  $f - \lambda$  ist nilpotent mit Nilpotenzgrad  $m$ .
- (C) Es existiert eine Zahl  $m \in \{1, \dots, n\}$ , so dass  $\mu_f(x) = (x - \lambda)^m$ .
- (D) Das charakteristische Polynom ist  $\chi_f(x) = (-1)^n(x - \lambda)^n$ .
- (E) Es existiert eine geordnete Basis  $B$ , so dass die Matrix  $f_{BB}$  eine obere Dreiecksmatrix ist, mit allen Diagonal-Einträgen gleich  $\lambda$ :

$$f_{BB} = \begin{pmatrix} \lambda & * & \cdots & \cdots & * \\ 0 & \lambda & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & * \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix}$$

- (F) Es existiert eine geordnete Basis  $B$ , so dass die Matrix  $f_{BB}$  die folgende Form hat:

$$f_{BB} = \begin{pmatrix} J_{m_1}(\lambda) & 0 & \cdots & \cdots & 0 \\ 0 & J_{m_2}(\lambda) & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & J_{m_N}(\lambda) \end{pmatrix}$$

Dabei ist die Größe der Jordan-Blöcke  $J_{m_k}(\lambda)$  monoton abnehmend.

$$m = m_1 \geq m_2 \geq \dots \geq m_N$$

Die in (A), (C), (F) genannten Zahlen  $m$  stimmen jeweils überein.

### 17.1.3 Beweis

Es wird einfach der Beweis des entsprechenden Satzes 17.1.2  $\lambda$ -verschoben, d.h. in diesem Satz anstelle des Endomorphismus  $f$  der Endomorphismus  $f - \lambda$  eingesetzt.

### 17.1.4 Satz: Das Lemma von Fitting für den $\lambda$ -verschobenen Endomorphismus

Es sei  $f : V \rightarrow V$  ein Endomorphismus des  $n$ -dimensionalen  $\mathbb{K}$ -Vektorraums  $V$  und  $\lambda \in \mathbb{K}$ .

- (i) Dann gibt es eine Zahl  $m$  mit  $1 \leq m \leq n$ , so dass

$$V = \ker(f - \lambda)^m \oplus \operatorname{im}(f - \lambda)^m$$

eine  $f$ -invariante Zerlegung ist.

Die Einschränkung  $f' - \lambda$  der Abbildung  $f - \lambda$  auf  $\ker(f - \lambda)^m$  ist nilpotent.

Die Einschränkung  $f'' - \lambda$  der Abbildung  $f - \lambda$  auf  $\operatorname{im}(f - \lambda)^m$  ist ein Isomorphismus.

- (ii) Es ist  $\dim \ker(f - \lambda)^m =$  Vielfachheit der Nullstelle  $\lambda$  in  $\chi_f$ .

Die besondere Rolle, die der Unterraum  $\ker(f - \lambda)^m$  in dieser Aussage spielt, wird durch die folgende Definition hervorgehoben:

### 17.1.5 Definition: Hauptraum

Ist  $f : V \rightarrow V$  ein Endomorphismus und  $\lambda \in \mathbb{K}$ , so heißt der in Satz 17.1.4 beschriebene Unterraum

$$\begin{aligned} \operatorname{Hpt}(f, \lambda) &:= \ker(f - \lambda)^m \\ &= \{v \in V \mid \text{Es ex. } \ell \in \mathbb{N} \text{ mit } (f - \lambda)^\ell(v) = \vec{0}\}. \end{aligned}$$

der *Hauptraum* von  $f$  zu  $\lambda$ .

Man überlegt sich leicht, dass der Hauptraum genau dann die Dimension  $\geq 1$  hat, wenn  $\lambda$  Eigenwert von  $f$  ist.

Das „ $\lambda$ -verschobene“ Lemma von Fitting führt dann in Zusammenhang mit diesem neuen Begriff zu der folgenden Aussage.

### 17.1.6 Satz: Hauptraum-Abspaltung

Es sei  $f : V \rightarrow V$  ein Endomorphismus des  $n$ -dimensionalen  $\mathbb{K}$ -Vektorraums  $V$  und  $\lambda \in \mathbb{K}$ .

- (i) Dann gibt es eine  $f$ -invariante Zerlegung von  $V$

$$V = \operatorname{Hpt}(f, \lambda) \oplus W,$$

mit der Eigenschaft, dass die Einschränkung der Abbildung  $f - \lambda$  auf  $W$  ein Isomorphismus ist.

- (ii) Es ist  $\dim \operatorname{Hpt}(f, \lambda) =$  Vielfachheit der Nullstelle  $\lambda$  in  $\chi_f$ .

## 17.2 Die Hauptraumzerlegung

### 17.2.1 Satz: Hauptraumzerlegung

Das charakteristische Polynom des Endomorphismus  $f : V \rightarrow V$  zerfalle über  $\mathbb{K}$  in Linearfaktoren:

$$\chi_f(x) = (-1)^n (x - \lambda_1)^{\ell_1} \cdot \dots \cdot (x - \lambda_r)^{\ell_r},$$

wobei die Eigenwerte  $\lambda_i$  paarweise verschieden sind. Dann ist

$$V = \text{Hpt}(f, \lambda_1) \oplus \dots \oplus \text{Hpt}(f, \lambda_r)$$

eine  $f$ -invariante Zerlegung.

### 17.2.2 Beweis

Wir spalten mit Hilfe von Satz 17.1.6 (i) den ersten Hauptraum ab und erhalten die  $f$ -invariante Zerlegung

$$V = \text{Hpt}(f, \lambda_1) \oplus W_1.$$

Das charakteristische Polynom der Einschränkung  $f_1$  auf  $W$  ist dann

$$\chi_1(x) = (-1)^{\ell_2 + \dots + \ell_r} (x - \lambda_2)^{\ell_2} \cdot \dots \cdot (x - \lambda_r)^{\ell_r}$$

und wir können den nächsten Hauptraum  $\text{Hpt}(f, \lambda_2)$  von  $W_1$  abspalten.

Nach insgesamt  $r$  Hauptraum-Abspaltungen erhalten wir

$$V = \text{Hpt}(f, \lambda_1) \oplus \dots \oplus \text{Hpt}(f, \lambda_r) \oplus W_r$$

Mit Satz 17.1.6(ii) schließen wir weiter

$$\dim \text{Hpt}(f, \lambda_1) + \dots + \dim \text{Hpt}(f, \lambda_r) = \ell_1 + \dots + \ell_r = n,$$

so dass  $W = \{\vec{0}\}$  sein muss.

## 17.3 Die Jordan-Normalform

### 17.3.1 Satz: Jordan-Normalform

Das charakteristische Polynom des Endomorphismus  $f : V \rightarrow V$ ,  $\dim V = n$ , zerfalle über  $\mathbb{K}$  in Linearfaktoren:

$$\chi_f(x) = (-1)^n (x - \lambda_1)^{\ell_1} \cdot \dots \cdot (x - \lambda_r)^{\ell_r}.$$

Die Eigenwerte  $\lambda_i$  seien dabei paarweise verschieden.

Dann gibt es eine geordnete Basis  $B$  von  $V$ , so dass  $f$  diesbezüglich die Darstellung

$$f_{BB} = \begin{pmatrix} \mathbf{J}_1 & 0 & \cdots & \cdots & 0 \\ 0 & \mathbf{J}_2 & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & \mathbf{J}_r \end{pmatrix} \in \mathbb{K}^{n \times n}$$

mit

$$\mathbf{J}_i = \begin{pmatrix} J_{m_{1,i}}(\lambda_i) & 0 & \cdots & \cdots & 0 \\ 0 & J_{m_{2,i}}(\lambda_i) & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & J_{m_{N_i,i}}(\lambda_i) \end{pmatrix} \in \mathbb{K}^{\ell_i \times \ell_i}$$

und dann

$$J_{m_{k,i}}(\lambda_i) = \begin{pmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & \lambda_i \end{pmatrix} \in \mathbb{K}^{m_{k,i} \times m_{k,i}}$$

hat.

### 17.3.2 Beweis

Er besteht in der Zusammenstellung der Sätze über die Hauptraumzerlegung 17.2.1 und der Aussage (F) aus Satz 17.1.2.

## 17.4 Algorithmus zur Transformation in Jordan-Normalform

### 17.4.1 Einführung

Vorgegeben sei ein Endomorphismus des  $\mathbb{K}$ -Vektorraums  $V$  mit  $\dim V = n$ . Es sei eine geordnete Basis  $B = \{w_1, \dots, w_n\}$  vorgegeben, so dass  $f$  durch die Matrix  $A = f_{BB}$  dargestellt ist.

Wir nehmen an, dass das charakteristische Polynom in Linearfaktoren zerfällt.

Ist von vornherein  $V = \mathbb{K}^n$ , so wird ein Endomorphismus  $\ell : \mathbb{K}^n \rightarrow \mathbb{K}^n$  stillschweigend bzgl. der kanonischen Basis  $E$  dargestellt. In diesem Fall werden die lineare Abbildung  $\ell$  und die Darstellungsmatrix  $A = \ell_{EE}$  gar nicht als unterschiedlich wahrgenommen. Meist ist  $A$  vorgegeben.

### 17.4.2 Ziel

Der im folgenden beschriebene Algorithmus liefert ...

- die zugehörige Jordan-Normalform,
- die Konstruktion der Basis  $\tilde{B}$ , bzgl. derer  $J = f_{\tilde{B}\tilde{B}}$  in Jordan-Normalform ist.
- die Berechnung der zugehörigen Transformationsmatrix  $T = \Theta_{\tilde{B}B}$ .

### 17.4.3 Der Algorithmus

**J1** Berechne, falls notwendig, die Matrix  $A = f_{BB}$ : Die Spalten von  $A$  enthalten die Entwicklungskoeffizienten der Entwicklung von  $f(w_j)$  bzgl. der Basis  $B = \{w_1, \dots, w_n\}$ .

**J2** Berechne das charakteristische Polynom  $\chi_f(x) = \chi_A(x)$ .

**J3** Ermittle die  $r$  verschiedenen Nullstellen  $\lambda_i \in \mathbb{K}$ . Folgende Möglichkeiten stehen zur Verfügung:

- Nullstellen erraten oder erspüren,
- Lösungsformel für die Nullstellen einer quadratischen Gleichung (Mitternachtsformel),
- Satz von Vieta,
- Polynomdivision,
- Lösung für biquadratische Gleichungen  $ax^4 + bx^2 + c = 0$  durch Substitution.
- Vielleicht kennt man Eigenwerte (Nullstellen) aufgrund spezieller Eigenschaften des Endomorphismus (Siehe nächste Kapitel: Endomorphismen in euklidischen Vektorräumen).

**J4** Ermittle für  $i = 1, \dots, r$  die algebraischen Vielfachheiten  $\ell_i$  der Nullstellen  $\lambda_i$ .

Das charakteristische Polynom hat dann die Form

$$\chi_f(x) = (-1)^n (x - \lambda_1)^{\ell_1} \cdot \dots \cdot (x - \lambda_r)^{\ell_r}$$



**J5** Für jeden einzelnen Eigenwert  $\lambda := \lambda_i$ :

**J5a** Berechne  $\dim \ker(f - \lambda)^j$ ,  $j = 1, \dots, M$ .

$M = m$  ist dabei das minimale  $j$  so, dass

$$\dim \ker(f - \lambda)^j = \dim \ker(f - \lambda)^{j+1}.$$

**J5b** Stelle das Young-Tableau auf mit den Zeilenlängen

$$n_j = \dim \ker(f - \lambda)^j - \dim \ker(f - \lambda)^{j-1}, \quad j = 1, \dots, M = m$$

**J5c** Entnehme dem Young-Tableau die Block-Größen  $m_1, \dots, m_N$ .

**J5d** Gib die zu  $\lambda_i$  gehörige Teilmatrix  $\mathbf{J}_i$  an. (Siehe Satz 17.1.2(F)).

**J6** Das Minimalpolynom ist

$$\mu_f(x) = (x - \lambda_1)^{M_1} \cdot \dots \cdot (x - \lambda_r)^{M_r}$$

**J7** Stelle aus den einzelnen Blöcken die gesamte Jordan-Normalform zusammen!

**J8** Betrachte für jeden einzelnen Eigenwert  $\lambda = \lambda_i$  das Young-Tableau  $Y$ . Es hat  $M$  Zeilen mit den Zeilenlängen  $n_1, \dots, n_M$ . Wir ordnen nach dem folgenden Verfahren jedem Kästchen des Young-Tableaus einen Vektor zu.

**J8a** Wir ordnen jedem Kästchen der untersten Zeile einen Vektor  $v_{M,k}$  so zu, dass

$$\ker(f - \lambda)^M = \ker(f - \lambda)^{M-1} \oplus \langle v_{M,1}, \dots, v_{M,n_M} \rangle.$$

**J8b** Ist jedem Kästchen der  $j + 1$ -ten Zeile ein Vektor  $v_{j+1,k}$ ,  $k = 1, \dots, n_{j+1}$  zugeordnet, so definieren wir zunächst  $n_{j+1}$  Vektoren der  $j$ -ten Zeile durch

$$v_{jk} = (f - \lambda)(v_{j+1,k}), \quad k = 1, \dots, n_{j+1},$$

und wählen weitere Vektoren  $v_{j,n_{j+1}+1}, \dots, v_{j,n_j}$  so, dass

$$\ker(f - \lambda)^j = \ker(f - \lambda)^{j-1} \oplus \underbrace{\langle v_{j,1}, \dots, v_{j,n_{j+1}} \rangle}_{\text{Bildvektoren}} \underbrace{\langle v_{j,n_{j+1}+1}, \dots, v_{j,n_j} \rangle}_{\text{neu zu wählen}}.$$

**J8c** Stelle alle diese Vektoren — spaltenweise von oben nach unten — zu einer geordneten Basis

$$B^{(i)} = \left( v_{11}, \dots, v_{M,1}, v_{12}, \dots, v_{m_2,2}, \dots, v_{1,n_1}, \dots, v_{m_{n_1},n_1} \right)$$

zusammen.

**J9** Stelle alle diese geordneten Basen zu der  $n \times n$ -Matrix

$$T^{-1} = \left( B^{(1)} \quad B^{(2)} \quad \dots \quad B^{(r)} \right)$$

zusammen. Das „ist“ zugleich die geordnete Basis  $\tilde{B}$ , bzgl. derer  $f_{\tilde{B}\tilde{B}} = J$ .

**J10** Berechne die Matrix  $T$  als die zu  $T^{-1}$  inverse. Es ist dann

$$J = T \cdot A \cdot T^{-1}.$$

## 17.5 Beispiel

### 17.5.1 Bemerkung

In dieser Abstraktheit schaut der obige Algorithmus zunächst ungenießbar aus. Bei der konkreten Anwendung werden viele Schritte kurz, einfacher oder obsolet, wie das folgende Beispiel zeigt.

### 17.5.2 Aufgabe

Wir suchen für die  $3 \times 3$ -Matrix

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

die Jordan-Normalform.

### 17.5.3 Anwendung des Algorithmus

**J2 - J4** Berechne das charakteristische Polynom.

$$\begin{aligned} \chi_A(x) &= \det \begin{pmatrix} 6-x & 2 & 0 \\ -2 & 1-x & 0 \\ 1 & 2 & 2-x \end{pmatrix} \\ &= (2-x)[(6-x)(1-x) + 4] \\ &= (2-x)(x^2 - 7x + 10) \\ &= -(x-2)^2(x-5) \end{aligned}$$

Die Eigenwerte sind  $\lambda_1 = 2$  und  $\lambda_2 = 5$  mit den Vielfachheiten  $\ell_1 = 2$  und  $\ell_2 = 1$ .

**J5** Berechne die Kerne von  $(A - \lambda_i)^j$  und deren Dimensionen. Es ist

$$(A - 2)x = \begin{pmatrix} 4 & 2 & 0 \\ -2 & -1 & 0 \\ 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

und deshalb

$$\ker(A - 2) = \left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle.$$

Weiter ist

$$(A - 2)^2 x = \begin{pmatrix} 12 & 6 & 0 \\ -6 & -3 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

und deshalb

$$\ker(A - 2)^2 = \left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} \right\rangle = \ker(A - 2) \oplus \underbrace{\left\langle \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} \right\rangle}_{=Z_2}.$$

Wir setzen also  $v_{21}^{(1)} = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}$  und dann

$$v_{11}^{(1)} = (A - 2)v_{21}^{(1)} = \begin{pmatrix} 4 & 2 & 0 \\ -2 & -1 & 0 \\ 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -3 \end{pmatrix}.$$

Zuletzt ist

$$(A - 5)x = \begin{pmatrix} 1 & 2 & 0 \\ -2 & -4 & 0 \\ 1 & 2 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

und deshalb

$$\ker(A - 5) = \left\langle \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} \right\rangle.$$

Wir setzen  $v_{11}^{(2)} = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}$ .

**J6** Das Minimalpolynom ist also

$$\mu_A(x) = (x - 2)^2(x - 5)$$

**J7** Demzufolge ist die Jordan-Normalform

$$J = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{pmatrix}.$$

**J8** ist oben in J5 schon erledigt.

**J9 - J10** Wir tragen die Vektoren  $v_{11}^{(1)}, v_{21}^{(1)}, v_{11}^{(2)}$  in die Matrix  $T^{-1}$  ein und berechnen  $T$ :

$$T^{-1} = \begin{pmatrix} 0 & 1 & 2 \\ 0 & -2 & -1 \\ -3 & 0 & 0 \end{pmatrix} \quad T = -\frac{1}{9} \begin{pmatrix} 0 & 0 & 3 \\ 3 & 6 & 0 \\ -6 & -3 & 0 \end{pmatrix}.$$

Man kann testen, dass

$$J = T \cdot A \cdot T^{-1}.$$

### 17.5.4 Hinweis

Gibt man unter <https://www.wolframalpha.com/> den Befehl

`diagonalize {{6,2,0},{-2,1,0},{1,2,2}}`

ein, so wird die soeben von Hand ausgerechnete Jordan-Normalform in kurzer Zeit bestätigt. Beachte, dass die Transformationsmatrix nicht eindeutig ist.

## 17.6 Charakterisierung der Ähnlichkeitsklassen

### 17.6.1 Satz: Charakterisierung der Ähnlichkeitsklassen

Es sei  $\mathbb{K}$  ein algebraisch abgeschlossener Körper, d.h. jedes Polynom aus  $\mathbb{K}[x]$  zerfällt in Linearfaktoren.

Es sei  $V$  ein  $\mathbb{K}$ -Vektorraum mit  $\dim V = n$ .

- (i) Die zu einem Endomorphismus  $f : V \rightarrow V$  gehörige Ähnlichkeitsklasse ist durch die folgenden Daten umkehrbar eindeutig charakterisiert:

- $r$  Eigenwerte  $\lambda^{(1)}, \dots, \lambda^{(r)}$
- Zu jedem  $i = 1, \dots, r$  ein Young-Tableau  $Y^{(i)}$  der Ordnung  $\ell^{(i)}$ , so dass
 
$$\ell^{(1)} + \dots + \ell^{(r)} = n.$$

- (ii) Ein Young-Tableau  $Y^{(i)}$  enthält dabei die folgenden Daten

Zeilenlängen	$n_j^{(i)}$	=	$\dim \ker(f - \lambda^{(i)})^j - \dim \ker(f - \lambda^{(i)})^{j-1}$
Spaltenlängen	$m_k^{(i)}$	=	Länge der Jordanblöcke $J_{m_k}(\lambda^{(i)})$
Ordnung	$\ell^{(i)}$	=	$\dim \text{Hpt}(f, \lambda^{(i)})$ = Vielfachheit der Nullstelle $\lambda^{(i)}$ in $\chi_f$ = algebraische Vielfachheit von $\lambda^{(i)}$ , vgl. 14.4.1.
Zeilenzahl	$M^{(i)}$	=	$m_1^{(i)}$ = Vielfachheit der Nullstelle $\lambda^{(i)}$ in $\mu_f$
Spaltenzahl	$N^{(i)}$	=	$n_1^{(i)}$ = $\dim \text{Eig}(f, \lambda^{(i)})$ = geometrische Vielfachheit von $\lambda^{(i)}$ , vgl. 18.1.1.

## 18 Diagonalisierung

### 18.1 Eigenvektoren

Es sei wieder  $f : V \rightarrow V$  Endomorphismus eines  $n$ -dimensionalen  $\mathbb{K}$ -Vektorraums  $V$ . Weiter sei  $\lambda \in \mathbb{K}$  Eigenwert von  $f$ .

#### 18.1.1 Definitionen: Eigenvektoren, Eigenraum.

Es sei  $f : V \rightarrow V$  Endomorphismus eines  $\mathbb{K}$ -Vektorraums und  $\lambda \in \mathbb{K}$  ein Eigenwert.

- (1) Der Vektor  $v \in V \setminus \{\vec{0}\}$  heißt *Eigenvektor* von  $f$  (zum Eigenwert  $\lambda$ ), wenn die Beziehung  $f(v) = \lambda \cdot v$  gilt.
- (2) Die Menge der Vektoren

$$\text{Eig}(f, \lambda) = \{v \in V \mid f(v) = \lambda \cdot v\} = \ker(f - \lambda)$$

heißt der *Eigenraum* von  $f$  zum Eigenwert  $\lambda$ . Diese Menge besteht aus den Eigenvektoren zum Eigenwert  $\lambda$  und dem Nullvektor.

- (3) Die Dimension des Eigenraums  $\text{Eig}(f, \lambda)$

$$\dim \text{Eig}(f, \lambda) \geq 1$$

heißt die *geometrische Vielfachheit* des Eigenwertes  $\lambda$  bzgl.  $f$ .

## 18.2 Diagonalisierbarkeit

### 18.2.1 Satz und Definition: Diagonalisierbarkeit eines Endomorphismus

Es sei  $f : V \rightarrow V$  ein Endomorphismus. Die folgenden Aussagen sind äquivalent:

- (A)  $f$  heißt *diagonalisierbar*.
- (B) Es gibt eine geordnete Basis  $B$  von  $V$ , bzgl. derer die zugeordnete Matrix  $f_{BB}$  eine Diagonalmatrix ist.
- (C) Es gibt eine geordnete Basis  $B$  von  $V$ , die nur Eigenvektoren enthält.
- (D) Es gibt eine  $f$ -invariante Zerlegung von  $V$  in die Eigenräume ( $\lambda_i \neq \lambda_j$  für  $i \neq j$ ):

$$V = \text{Eig}(f, \lambda_1) \oplus \dots \oplus \text{Eig}(f, \lambda_r).$$

- (E) Für die Summe der geometrischen Vielfachheiten gilt ( $\lambda_i \neq \lambda_j$  für  $i \neq j$ ):

$$\dim \text{Eig}(f, \lambda_1) + \dots + \dim \text{Eig}(f, \lambda_r) = \dim V.$$

- (F) Das charakteristische Polynom zerfällt in Linearfaktoren und

für jeden Eigenwert stimmen

die geometrische Vielfachheit (= Dimension des zugehörigen Eigenraums)

und

die algebraische Vielfachheit (= Vielfachheit als Nullstelle in  $\chi_f$ )

überein.

- (G) Das Minimalpolynom zerfällt in Linearfaktoren mit Vielfachheit 1.
- (H) Die zu den Eigenwerten  $\lambda_1, \dots, \lambda_r$  gehörigen Young-Tableaus haben alle nur eine Zeile.
- (J) Alle Jordan-Blöcke in der Jordan-Normalform haben die Dimension 1.

Das Verfahren zum Auffinden einer geordneten Basis  $B$ , bzgl. der die zugeordnete Matrix  $f_{BB}$  Diagonalmatrix ist, heißt *Diagonalisierung*.

### 18.2.2 Beweis

Die Äquivalenzen (B)  $\Leftrightarrow$  (C)  $\Leftrightarrow$  (D) sind durch Hinschauen einsehbar.

(D)  $\Rightarrow$  (E) ist eine allgemeine Aussage über innere Summen, siehe Satz 6.2.2.

(E)  $\Rightarrow$  (D): Dafür ist nur  $\text{Eig}(f, \lambda_i) \cap \text{Eig}(f, \lambda_j) = \{\vec{0}\}$  für  $i \neq j$  zu zeigen.

Ist  $v \in \text{Eig}(f, \lambda_i) \cap \text{Eig}(f, \lambda_j)$ , so folgt

$$(\lambda_i - \lambda_j)v = \lambda_i v - \lambda_j v = f(v) - f(v) = \vec{0}.$$

Wegen  $\lambda_i \neq \lambda_j$  folgt  $v = \vec{0}$ .

(D)  $\Rightarrow$  (F): Das charakteristische Polynom des auf  $\text{Eig}(f, \lambda_i)$  eingeschränkten Endomorphismus  $f_i : \text{Eig}(f, \lambda_i) \rightarrow \text{Eig}(f, \lambda_i)$  ist

$$\chi_{f_i}(x) = (x - \lambda_i)^{\dim \text{Eig}(f, \lambda_i)}.$$

Demzufolge ist

$$\chi_f(x) = (x - \lambda_1)^{\dim \text{Eig}(f, \lambda_1)} \cdot \dots \cdot (x - \lambda_r)^{\dim \text{Eig}(f, \lambda_r)}.$$

(F)  $\Rightarrow$  (E) folgt sofort wegen  $\text{grad } \chi_f = \dim V$ .

(D)  $\Rightarrow$  (G): Die Zerlegung in Linearfaktoren von  $\mu_f$  ersieht man aus der Folgerung 14.5.4 des Satzes von Cayley-Hamilton.

Ist  $v \in V$  beliebig, so gibt es wegen (D) eine eindeutige Summendarstellung

$$v = v_1 + \dots + v_r \quad \text{mit } v_i \in \text{Eig}(f, \lambda_i)$$

Aufgrund der  $f$ -Invarianz der Zerlegung in (D) gilt

$$\begin{aligned} & (f - \lambda_1) \cdot \dots \cdot (f - \lambda_r)(v_1 + \dots + v_r) \\ &= \sum_{i=1}^r (f - \lambda_1) \cdot \dots \cdot (f - \lambda_i) \cdot \underbrace{(f - \lambda_{i+1}) \cdot \dots \cdot (f - \lambda_r)}_{\in \text{Eig}(f, \lambda_i)} v_i \\ &= \vec{0}. \end{aligned}$$

Also ist  $\mu_f$  ein Teiler von

$$(x - \lambda_1) \cdot \dots \cdot (x - \lambda_r)$$

Jeder dieser Linearfaktoren  $(x - \lambda_i)$  muss aber auch ein Teiler von  $\mu_f$  sein, da anderenfalls ein Vektor aus  $\text{Eig}(f, \lambda_i)$  nicht durch  $\mu_f(f)$  auf den Nullvektor abgebildet werden könnte.

(G)  $\Rightarrow$  (D): Wenn  $\mu_f$  in Linearfaktoren zerfällt, tut dies auch  $\chi_f$ . Siehe Folgerung 14.5.4 des Satzes von Cayley-Hamilton.

Damit liefert Satz 17.2.1

$$V = \text{Hpt}(f, \lambda_1) \oplus \dots \oplus \text{Hpt}(f, \lambda_r)$$

die  $f$ -invariante Zerlegung in die Haupträume. Da das Minimalpolynom  $\mu_{f_i}$  der Einschränkung von  $f$  auf  $\text{Hpt}(f, \lambda_i)$  ein Teiler von  $\mu_f$  sein muss, gilt:  $\mu_{f_i}(x) = (x - \lambda_i)$ . Das aber bedeutet:

$$\text{Hpt}(f, \lambda_i) = \text{Eig}(f, \lambda_i),$$

somit

$$V = \text{Eig}(f, \lambda_1) \oplus \dots \oplus \text{Eig}(f, \lambda_r).$$

Die Äquivalenz all dieser Aussagen zu (H) und (J) ist offensichtlich.

### 18.2.3 Folgerung

Besitzt ein Endomorphismus  $f : V \rightarrow V$  eines  $n$ -dimensionalen Vektorraums  $n$  paarweise verschiedene Eigenwerte  $\lambda_i \in \mathbb{K}$ , so ist  $f$  diagonalisierbar, wobei

$$V = \underbrace{\text{Eig}(f, \lambda_1)}_{\dim=1} \oplus \dots \oplus \underbrace{\text{Eig}(f, \lambda_n)}_{\dim=1}$$
$$\mu_f(x) = \chi_f(x) = (x - \lambda_1) \cdot \dots \cdot (x - \lambda_n)$$

### 18.2.4 Beweis

folgt sofort mit (E) im vorhergehenden Satz wegen  $\dim \text{Eig}(f, \lambda_i) \geq 1$  für alle  $i$ .



## 18.3 Algorithmus zur Diagonalisierung

### 18.3.1 Einführung

Vorgegeben sei ein Endomorphismus des  $\mathbb{K}$ -Vektorraums  $V$  mit  $\dim V = n$ . Es sei eine Basis  $B = \{w_1, \dots, w_n\}$  vorgegeben, so dass  $f$  durch die Matrix  $A = f_{BB}$  dargestellt ist.

Ist von vornherein  $V = \mathbb{K}^n$ , so wird ein Endomorphismus  $f : \mathbb{K}^n \rightarrow \mathbb{K}^n$  stillschweigend bzgl. der kanonischen Basis  $E$  dargestellt. In diesem Fall werden die lineare Abbildung  $f$  und die Darstellungsmatrix  $A = f_{EE}$  gar nicht als unterschiedlich wahrgenommen. Meist ist  $A$  vorgegeben.

**18.3.2 Ziel** Der im folgenden beschriebene Algorithmus liefert ...

- die Entscheidung, ob es eine Basis  $\tilde{B}$  von  $V$  gibt, so dass  $f_{\tilde{B}\tilde{B}}$  eine Diagonalmatrix ist,
- die Diagonalmatrix selbst,
- die Konstruktion der Basis  $\tilde{B}$ ,
- die Berechnung der zugehörigen Transformationsmatrix  $T = \Theta_{\tilde{B}B}$ .

### 18.3.3 Der Algorithmus

**D1** Berechne, falls notwendig, die Matrix  $A = f_{BB}$ : Die Spalten von  $A$  enthalten die Entwicklungskoeffizienten der Entwicklung von  $f(w_j)$  bzgl. der Basis  $B = \{w_1, \dots, w_n\}$ .

**D2** Berechne das charakteristische Polynom  $\chi_f(x) = \chi_A(x)$ .

**D3** Ermittle die  $r$  verschiedenen Nullstellen  $\lambda_i \in \mathbb{K}$ . Folgende Möglichkeiten stehen zur Verfügung:

- Nullstellen erraten oder erspüren,
- Lösungsformel für die Nullstellen einer quadratischen Gleichung (Mitternachtsformel),
- Satz von Vieta,
- Polynomdivision,
- Lösung für biquadratische Gleichungen  $ax^4 + bx^2 + c = 0$  durch Substitution.
- Vielleicht kennt man Eigenwerte (Nullstellen) aufgrund spezieller Eigenschaften des Endomorphismus (Siehe nächste Kapitel: Endomorphismen in euklidischen Vektorräumen).

**D4** Ermittle für  $i = 1, \dots, r$  die algebraischen Vielfachheiten  $\ell_i$  der Nullstellen  $\lambda_i$ .

**D5a**  $\ell_1 + \dots + \ell_r = n$  ? Falls NEIN:  
 $\chi_f$  zerfällt nicht vollständig in Linearfaktoren über  $\mathbb{K}$ , d.h.:  
 $f$  ist nicht über  $\mathbb{K}$  diagonalisierbar.  $\ominus$

(Beachte, dass dies im Fall  $\mathbb{K} = \mathbb{C}$  nicht passieren kann)

- D5b**  $\ell_1 + \dots + \ell_r = n$  ? Falls JA:  
 $\chi_f$  zerfällt vollständig in Linearfaktoren über  $\mathbb{K}$ :  
 Schreibe  $\chi_f$  als Produkt der Linearfaktoren:

$$\chi_f(x) = (x - \lambda_1)^{\ell_1} \cdot \dots \cdot (x - \lambda_r)^{\ell_r}.$$

- D6** Ermittle die geometrischen Vielfachheiten  $N_i$  der Eigenwerte  $\lambda_i$

$$N_i = \dim \ker(f - \lambda_i) = n - \text{Rang}(f - \lambda_i).$$

Beachte, dass  $1 \leq N_i \leq \ell_i$  gelten muss.

- D7a** Stimmen für eine Nullstelle die beiden Vielfachheiten  $N_i$  und  $\ell_i$  nicht überein, so ist  $f$  nicht diagonalisierbar.  $\ominus$

- D7b** Stimmen für alle Nullstellen die beiden Vielfachheiten  $N_i$  und  $\ell_i$  überein, so ist  $f$  diagonalisierbar. Es gibt eine Basis  $\tilde{B}$ , bzgl. derer  $f$  die Darstellung

$$\tilde{A} = f_{\tilde{B}\tilde{B}} = \begin{pmatrix} \lambda_1 & & & & \\ & \ddots & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \lambda_r \end{pmatrix}$$

hat. Dabei kommt der Eigenwert  $\lambda_i$  genau  $\ell_i$ -mal vor.

Ermittle, wenn gewünscht, die neue Basis  $\tilde{B}$  und die zugehörige Übergangsmatrix  $T = \Theta_{\tilde{B}B}$  gemäß der folgenden Schritte.

- D8** Für jede Nullstelle  $\lambda_i$ : Bestimme  $\ell_i$  linear unabhängige Vektoren

$$x_k^{(i)} \in \mathbb{K}^n, \quad k = 1, \dots, \ell_i$$

durch Lösen des Gleichungssystems (vielleicht schon in Schritt **D6** erledigt)

$$(A - \lambda_i) x^{(i)} = \vec{0}.$$

- D9** Bilde die Matrix

$$T^{-1} = \Theta_{\tilde{B}B}^{-1} := \left( x_1^{(1)}, \dots, x_{\ell_1}^{(1)}, x_1^{(2)}, \dots, x_{\ell_2}^{(2)}, \dots, x_1^{(r)}, \dots, x_{\ell_r}^{(r)} \right)$$

- D10** Berechne die Matrix  $T$ .

- D11** Dann hat die Matrix

$$\tilde{A} = T \cdot A \cdot T^{-1}$$

bezüglich der Basis  $\tilde{B} = \{\tilde{w}_1, \dots, \tilde{w}_n\}$  mit

$$\tilde{w}_i = \sum_{k=1}^n (T^{-1})_{kj} w_k$$

die Diagonalgestalt wie in Schritt **D7b** angegeben.

## 18.4 Beispiele

### 18.4.1 Beispiel 1

Wir betrachten das Beispiel vom Anfang des Kapitels 15. Es sei  $V$  ein zweidimensionaler reeller Vektorraum und der Endomorphismus  $f : V \rightarrow V$  definiert durch

$$f(w_1) = \frac{16}{3}w_1 - 5w_2, \quad f(w_2) = \frac{10}{3}w_1 - 3w_2.$$

D1

$$A = f_{BB} = \begin{pmatrix} \frac{16}{3} & \frac{10}{3} \\ -5 & -3 \end{pmatrix}$$

D2

$$\chi_f(x) = \begin{vmatrix} \frac{16}{3} - x & \frac{10}{3} \\ -5 & -3 - x \end{vmatrix} = \left(\frac{16}{3} - x\right)(-3 - x) - \frac{10}{3} \cdot (-5) = x^2 - \frac{7}{3}x + \frac{2}{3}$$

D3

$$\lambda_{1,2} = \frac{\frac{7}{3} \pm \sqrt{\frac{49}{9} - \frac{8}{3}}}{2} = \frac{\frac{7}{3} \pm \frac{5}{3}}{2}, \quad \text{also } \lambda_1 = 2, \quad \lambda_2 = \frac{1}{3}.$$

D4 Die algebraischen Vielfachheiten der Eigenwerte sind  $\ell_1 = 1$ ,  $\ell_2 = 1$ .

D5b Die Zerlegung in Linearfaktoren ist  $\chi_f(x) = (x - 2)(x - \frac{1}{3})$ .

D6 Es muss  $N_1 = 1$ ,  $N_2 = 1$  gelten.

D7b Da die Vielfachheiten alle übereinstimmen, ist  $f$  diagonalisierbar.

D8 Wir lösen das LGS für den ersten Eigenwert

$$(A - \lambda_1)x = \begin{pmatrix} \frac{16}{3} - 2 & \frac{10}{3} \\ -5 & -3 - 2 \end{pmatrix}x = \begin{pmatrix} \frac{10}{3} & \frac{10}{3} \\ -5 & -5 \end{pmatrix}x = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Wir finden ganz leicht eine Lösung

$$x_1^{(1)} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Wir lösen das LGS für den zweiten Eigenwert

$$(A - \lambda_2)x = \begin{pmatrix} \frac{16}{3} - 13 & \frac{10}{3} \\ -5 & -3 - \frac{1}{3} \end{pmatrix}x = \begin{pmatrix} 5 & \frac{10}{3} \\ -5 & -\frac{10}{3} \end{pmatrix}x = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Wir sehen ganz leicht

$$x_1^{(2)} = \begin{pmatrix} 2 \\ -3 \end{pmatrix}.$$

**D9** Jetzt kann die Matrix  $T^{-1}$  zusammengestellt werden:

$$T^{-1} = \begin{pmatrix} 1 & 2 \\ -1 & -3 \end{pmatrix}.$$

(Sie unterscheidet sich von der Matrix  $\Theta_{\bar{B}B}^{-1}$  in Abschnitt 15.1.7 durch einen Faktor  $-1$  in der zweiten Spalte. Beachte, dass die Übergangsmatrix  $T$  nicht eindeutig bestimmt ist.)

**D10** Die inverse Matrix ist

$$T = \begin{pmatrix} 3 & 2 \\ -1 & -1 \end{pmatrix}.$$

**D11** Es gilt dann

$$\tilde{A} = T \cdot A \cdot T^{-1} = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{3} \end{pmatrix}.$$

### 18.4.2 Beispiel 2

Es sei  $f$  eine lineare Abbildung in einem 3-dimensionalen reellen Vektorraum. Bezüglich einer Basis  $B$  habe sie die Matrixdarstellung

$$A = f_{BB} = \begin{pmatrix} 3 & 1 & 0 \\ 0 & -1 & c \\ 0 & 0 & 3 \end{pmatrix}$$

**D1** Dieser Schritt erübrigt sich.

**D2** Das charakteristische Polynom ist wegen der Dreiecksgestalt leicht zu berechnen:

$$\chi_f(x) = (3-x)^2 \cdot (-1-x).$$

**D3** Die beiden Eigenwerte sind  $\lambda_1 = 3$ ,  $\lambda_2 = -1$ .

**D4** Die algebraischen Vielfachheiten der Eigenwerte sind  $\ell_1 = 2$ ,  $\ell_2 = 1$ .

**D5** Die Zerlegung in Linearfaktoren liegt bereits vor.

**D6** Es ist

$$A - 3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & -4 & c \\ 0 & 0 & 0 \end{pmatrix},$$

also

$$N_1 = n - \text{rang}(A - 3) = \begin{cases} 3 - 2 = 1 \neq \ell_1, & \text{falls } c \neq 0, \\ 3 - 1 = 2 = \ell_1, & \text{falls } c = 0. \end{cases}$$

Weiter ist  $N_2 = 1$ .

**D7**  $f$  ist genau dann diagonalisierbar (über  $\mathbb{R}$  oder über  $\mathbb{C}$ ), wenn  $c = 0$ . Wir führen für diesen Fall die Diagonalisierung durch.

**D8** Die Gleichung

$$(A - \lambda_1)x = \begin{pmatrix} 0 & 1 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 0 \end{pmatrix} x = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

hat die beiden linear unabhängigen Lösungen

$$x_1^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \text{und} \quad x_2^{(1)} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

Die Gleichung

$$(A - \lambda_2)x = \begin{pmatrix} 4 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 4 \end{pmatrix} x = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

hat die nicht-triviale Lösung

$$x_1^{(2)} = \begin{pmatrix} 1 \\ -4 \\ 0 \end{pmatrix}.$$

**D9** Es wird die Matrix

$$T^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & -4 \\ 1 & -1 & 0 \end{pmatrix}$$

zusammengestellt.

**D10** Die inverse Matrix ist

$$T = \frac{1}{-8} \cdot \begin{pmatrix} -4 & -1 & -4 \\ -4 & -1 & 4 \\ 0 & 2 & 0 \end{pmatrix}$$

**D11** Es gilt dann

$$\tilde{A} = T \cdot A \cdot T^{-1} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

### 18.4.3 Beispiel 3

Es sei  $f$  eine lineare Abbildung in einem 3-dimensionalen Vektorraum. Bezüglich einer Basis habe sie die Matrixdarstellung

$$A = f_{BB} = \begin{pmatrix} -2 & 0 & 1 \\ 0 & 3 & 0 \\ -1 & 0 & -2 \end{pmatrix}$$

Ist  $f$  über  $\mathbb{R}$  oder über  $\mathbb{C}$  diagonalisierbar? Wenn ja, führe das Diagonalisierungsverfahren durch.

**D1** Dieser Schritt erübrigt sich.

**D2** Durch Entwicklung nach der zweiten Zeile erhält man das charakteristische Polynom:

$$\chi_f(x) = (3 - x) \cdot [(-2 - x)^2 + 1] = -(x - 3) \cdot [(x + 2)^2 + 1].$$

**D3** Der Fall  $\mathbb{K} = \mathbb{R}$ : Der zweite Faktor in  $\chi_f$  ist positiv für alle  $x \in \mathbb{R}$ . Es gibt also nur eine reelle Nullstelle  $\lambda_1 = 3$ .  $f$  ist nicht über  $\mathbb{R}$  diagonalisierbar.

**D3** Der Fall  $\mathbb{K} = \mathbb{C}$ : Man kann das charakteristische Polynom in Linearfaktoren zerlegen:

$$\chi_f(x) = -(x - 3) \cdot [(x + 2)^2 + 1] = -(x - 3) \cdot [x - (i - 2)] \cdot [x - (-i - 2)].$$

Die Eigenwerte sind  $\lambda_1 = 3$ ,  $\lambda_2 = i - 2$ ,  $\lambda_3 = -i - 2$ .

**D4** Die algebraischen Vielfachheiten der Eigenwerte sind  $\ell_1 = \ell_2 = \ell_3 = 1$ .

**D5** Bereits erledigt.

**D6** Die geometrischen Vielfachheiten  $N_1 = N_2 = N_3 = 1$ .

**D7**  $f$  ist über  $\mathbb{C}$  diagonalisierbar.

**D8** Die Gleichung

$$(A - \lambda_1)x = \begin{pmatrix} 5 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & -5 \end{pmatrix} x = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

hat die nicht-triviale Lösung

$$x_1^{(1)} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Die Gleichung

$$(A - \lambda_2)x = \begin{pmatrix} -i & 0 & 1 \\ 0 & 5 - i & 0 \\ -1 & 0 & -i \end{pmatrix} x = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

hat die nicht-triviale Lösung

$$x_1^{(2)} = \begin{pmatrix} 1 \\ 0 \\ i \end{pmatrix}.$$

Die Gleichung

$$(A - \lambda_3)x = \begin{pmatrix} +i & 0 & 1 \\ 0 & 5+i & 0 \\ -1 & 0 & +i \end{pmatrix} x = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

hat die nicht-triviale Lösung

$$x_1^{(3)} = \begin{pmatrix} 1 \\ 0 \\ -i \end{pmatrix}.$$

**D9** Wir bilden die Matrix

$$T^{-1} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & i & -i \end{pmatrix}.$$

**D10** Die inverse Matrix ist

$$T = \frac{1}{2i} \cdot \begin{pmatrix} -0 & 2i & 0 \\ i & 0 & 1 \\ i & 0 & -1 \end{pmatrix}$$

**D11** Es gilt dann

$$\tilde{A} = T \cdot A \cdot T^{-1} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & i-2 & 0 \\ 0 & 0 & -i-2 \end{pmatrix}.$$

## 19 Normale Endomorphismen

### 19.1 Vorbemerkung / Vorbereitung

In dem gesamten Kapitel seien alle auftretenden Vektorräume endlich-dimensionale  $\mathbb{K}$ -Vektorräume mit Skalarprodukt, wobei einer der beiden folgenden Fälle zutrifft:

- $\mathbb{K} = \mathbb{R}$   $\langle \cdot, \cdot \rangle$  ist eine positiv definite Bilinearform mit Werten in  $\mathbb{R}$
- $\mathbb{K} = \mathbb{C}$   $\langle \cdot, \cdot \rangle$  ist eine positiv definite Sesquilinearform mit Werten in  $\mathbb{C}$ .

#### 19.1.1 Wichtige Beobachtung

Es seien  $\{v_1, \dots, v_n\}$  eine ONB für den Vektorraum  $V$  und  $\{w_1, \dots, w_m\}$  eine ONB für den Vektorraum  $W$ .

- (i) Ein Vektor  $v \in V$  ist durch die  $n$  Zahlen

$$\langle v_1, v \rangle \quad \dots \quad \langle v_n, v \rangle$$

umkehrbar eindeutig bestimmt.

- (ii) Ein Homomorphismus  $f : V \rightarrow W$  ist durch die  $m \cdot n$  Zahlen

$$\begin{array}{cccc} \langle w_1, f(v_1) \rangle & \dots & \langle w_1, f(v_n) \rangle & \\ \langle w_2, f(v_1) \rangle & \dots & \langle w_2, f(v_n) \rangle & \\ \vdots & & \vdots & \\ \langle w_m, f(v_1) \rangle & \dots & \langle w_m, f(v_n) \rangle & \end{array}$$

umkehrbar eindeutig bestimmt.

#### 19.1.2 Satz: Dimension des orthogonalen Komplements

Das orthogonale Komplement  $U^\perp$  einer Teilmenge  $U \subseteq V$  ist ein Unterraum von  $V$ .

Ist  $U$  ein Unterraum von  $V$ , so gilt

$$\dim U^\perp = \dim V - \dim U$$

#### 19.1.3 Beweis

- (1) Sind  $a, \tilde{a} \in U^\perp$ , so gilt für alle  $u \in U$

$$\langle \alpha a + \tilde{\alpha} \tilde{a}, u \rangle = \alpha \langle a, u \rangle + \tilde{\alpha} \langle \tilde{a}, u \rangle = 0,$$

also  $\alpha a + \tilde{\alpha} \tilde{a} \in U^\perp$ .

- (2) Man betrachte eine Orthonormalbasis

$$B_U = \{v_1, \dots, v_r\} \quad \text{von } U$$



und ergänze sie zu einer Orthonormalbasis

$$B_V = \{v_1, \dots, v_r, v_{r+1}, \dots, v_n\} \quad \text{von } V$$

(3) Wir überzeugen uns nun davon, dass die Teilmenge

$$\tilde{B} := \{v_{r+1}, \dots, v_n\}$$

eine Basis von  $U^\perp$  ist.

(3a) Die lineare Unabhängigkeit ist klar.

(3b) Wegen  $\langle v_j, v_k \rangle = 0$  für  $j = r+1, \dots, n$  und  $k = 1, \dots, r$  ist  $\langle \tilde{B} \rangle \subseteq U^\perp$ .

(3c) Ist  $a \in U^\perp$ , so gibt es eine Basisentwicklung bzgl.  $B_V$

$$a = \alpha_1 v_1 + \dots + \alpha_r v_r + \alpha_{r+1} v_{r+1} + \dots + \alpha_n v_n$$

Bildet man hier nacheinander das Skalarprodukt mit  $v_1, \dots, v_r$  ein, so folgt  $\alpha_1 = \dots = \alpha_r = 0$  und deshalb

$$a = \alpha_{r+1} v_{r+1} + \dots + \alpha_n v_n.$$

Wir haben also  $U^\perp \subseteq \langle \tilde{B} \rangle$  gezeigt.

#### 19.1.4 Satz: Doppel-Orthogonal-Komplement

Für einen Unterraum  $U \subseteq V$  gilt

$$U^{\perp\perp} = U.$$

#### 19.1.5 Beweis

Es ist klar, dass  $U \subseteq U^{\perp\perp}$ . Weiter ist aufgrund von Satz 19.1.2

$$\dim U^{\perp\perp} = n - \dim U^\perp = n - (n - \dim U) = \dim U.$$

Deswegen ist  $U^{\perp\perp} = U$ .

## 19.2 Der adjungierte Homomorphismus

### 19.2.1 Definition und Satz: Adjungierter Homomorphismus

Es seien  $f : V \rightarrow W$  und  $f^\dagger : W \rightarrow V$  zwei Endomorphismen zwischen unitären oder euklidischen Vektorräumen.

Die folgenden Aussagen sind äquivalent:

(A) (Def)  $f^\dagger$  heißt der zu  $f$  *adjungierte Homomorphismus*.

(B) Für alle  $v \in V, w \in W$  gilt

$$\langle f(v), w \rangle = \langle v, f^\dagger(w) \rangle.$$

(C) Sind  $B$  und  $C$  geordnete Orthonormalbasen für  $V$  bzw.  $W$ , so gilt für die Darstellungsmatrizen

$$f_{BC}^\dagger = \overline{f_{CB}^T}$$

Der adjungierte Homomorphismus ist eindeutig bestimmt.

### 19.2.2 Beweis

Die Eindeutigkeitsaussage ist aus (B) mit Hilfe der Beobachtung 19.1.1 klar.

(C)  $\Rightarrow$  (B) Es gilt für alle  $v, w \in V$

$$\langle f(v), w \rangle = (f_{CB} \cdot \eta_B(v))^T \cdot \overline{\eta_C(w)} = \eta_B(v)^T \cdot f_{CB}^T \cdot \overline{\eta_C(w)}$$

$$\langle v, f^\dagger(w) \rangle = \eta_B(v)^T \cdot \overline{f_{BC}^\dagger \cdot \eta_C(w)} = \eta_B(v)^T \cdot \overline{f_{BC}^\dagger} \cdot \overline{\eta_C(w)}$$

Sind also die beiden Matrixdarstellungen transponiert-konjugiert-komplex zueinander, so folgt, dass  $f$  und  $g$  adjungiert sind.

(B)  $\Rightarrow$  (C) Gilt umgekehrt Gleichheit zwischen den beiden Ausdrücken links, so folgt, wenn man die Vektoren der geordneten Orthonormalbasen  $v_j$  und  $w_k$  einsetzt, dass

$$(f_{BC}^\dagger)_{jk} = \overline{(f_{CB}^T)_{jk}}.$$

**19.2.3 Satz: Adjungierung ist Involution**

(i) Für einen Homomorphismus  $f : V \rightarrow W$  gilt

$$(f^\dagger)^\dagger = f.$$

(ii) Für zwei Homomorphismen  $f : V \rightarrow U$  und  $g : U \rightarrow W$  zwischen drei unitären oder euklidischen Vektorräumen gilt

$$(g \circ f)^\dagger = f^\dagger \circ g^\dagger$$

Die beiden Eigenschaften drücken aus, dass Adjungierung eine *Involution* ist.

**19.2.4 Beweis**

(i) Unter Bezug auf die Aussage (B) gilt

$$\langle f^{\dagger\dagger}(v), w \rangle = \langle w, f^{\dagger\dagger}(v) \rangle = \langle f^\dagger(w), v \rangle = \langle v, f^\dagger(w) \rangle = \langle f(v), w \rangle.$$

Die Beobachtung 19.1.1 liefert die Aussage (i).

(ii) Für alle  $v \in V$  und  $w \in W$  gilt

$$\langle v, (g \circ f)^\dagger(w) \rangle = \langle (g \circ f)(v), w \rangle = \langle f(v), g^\dagger(w) \rangle = \langle v, f^\dagger(g^\dagger(w)) \rangle.$$

Deshalb stimmen  $(g \circ f)^\dagger$  und  $f^\dagger \circ g^\dagger$  aufgrund der Beobachtung 19.1.1 überein.

**19.2.5 Satz: Kern und Bild beim adjungierten Homomorphismus**

$$\begin{aligned} \ker f &= (\operatorname{im} f^\dagger)^\perp & (\ker f)^\perp &= \operatorname{im} f^\dagger \\ \ker f^\dagger &= (\operatorname{im} f)^\perp & (\ker f^\dagger)^\perp &= \operatorname{im} f \end{aligned}$$

**19.2.6 Beweis**

Wir zeigen nur die erste Gleichung. Die anderen folgen dann durch Übergang zur Adjungierten bzw. zum orthogonalen Komplement.

Ist  $v \in \ker f$ , so folgt für beliebige  $w \in W$

$$\langle v, f^\dagger(w) \rangle = \langle f(v), w \rangle = 0$$

und daher  $v \in (\operatorname{im} f^\dagger)^\perp$ .

Sei  $v \in (\operatorname{im} f^\dagger)^\perp$ . Für alle  $w \in V$  ist dann

$$\langle f(v), w \rangle = \langle v, f^\dagger(w) \rangle = 0.$$

Daraus folgt  $f(v) = \vec{0}$  und damit  $v \in \ker f$ .

## 19.3 Der adjungierte Endomorphismus

### 19.3.1 Definition und Satz: Adjungierter Endomorphismus

Es seien  $f : V \rightarrow V$  und  $f^\dagger : V \rightarrow V$  zwei Endomorphismen in einem unitären oder euklidischen Vektorraum. Die folgenden Aussagen sind äquivalent:

- (A) (Def)  $f^\dagger$  heißt der zu  $f$  *adjungierte Endomorphismus*.  
 (B) Für alle  $u, v \in V$  gilt

$$\langle f(u), v \rangle = \langle u, f^\dagger(v) \rangle.$$

- (C) Ist  $B$  eine geordnete Orthonormalbasis von  $V$ , so gilt für die Darstellungsmatrizen

$$f_{BB}^\dagger = \overline{f_{BB}^T}$$

### 19.3.2 Beweis

Dies ist der Satz 19.2.1 aus dem letzten Unterkapitel mit  $W = V$ .

**19.3.3 Folgerung** Für einen Endomorphismus  $f : V \rightarrow V$  und seinen adjungierten Endomorphismus  $f^\dagger : V \rightarrow V$  gilt

$$\det f^\dagger = \overline{\det f}.$$

Dies folgt sofort aus der Matrixdarstellung:

$$\det f^\dagger = \det f_{BB}^\dagger = \det \overline{f_{BB}^T} = \overline{\det f_{BB}} = \overline{\det f}.$$

## 19.4 Normale Endomorphismen

### 19.4.1 Definition und Satz: Normale Endomorphismen

Für einen Endomorphismus  $f : V \rightarrow V$  in einem unitären Vektorraum  $V$  sind die folgenden Aussagen äquivalent:

(A)  $f$  heißt *normal*.

(B) Es ist  $f^\dagger \circ f = f \circ f^\dagger$ .

(C) Für alle  $u, v \in V$  gilt

$$\langle f(u), f(v) \rangle = \langle f^\dagger(u), f^\dagger(v) \rangle$$

(D) Ist  $B$  eine geordnete Orthonormalbasis von  $V$ , so gilt für die Darstellungsmatrizen

$$\overline{f_{BB}^T} \cdot f_{BB} = f_{BB} \cdot \overline{f_{BB}^T}$$

(E) Es gibt eine geordnete Orthonormalbasis  $B$  aus Eigenvektoren von  $f$ , so dass  $f_{BB}$  Diagonalmatrix ist.

Beachte, dass die Voraussetzung „ $V$  euklidisch“ für diesen Satz nicht ausreicht.

### 19.4.2 Beweis

(B)  $\Rightarrow$  (C) Für  $u, v \in V$  gilt

$$\langle f(u), f(v) \rangle = \langle u, (f^\dagger \circ f)(v) \rangle = \langle u, (f \circ f^\dagger)(v) \rangle = \langle f^\dagger(u), f^\dagger(v) \rangle.$$

(C)  $\Rightarrow$  (B) Gilt umgekehrt (C), so ist für alle  $u, v \in V$

$$\langle u, (f^\dagger \circ f)(v) \rangle = \langle f(u), f(v) \rangle = \langle f^\dagger(u), f^\dagger(v) \rangle = \langle u, (f \circ f^\dagger)(v) \rangle$$

und mit Beobachtung 19.1.1 folgt die Behauptung (B).

Die Äquivalenz (B)  $\Leftrightarrow$  (D) ist klar aus Satz 19.4.1.

Die Äquivalenz zu (E) wird weiter unten in Abschnitt 19.4.7 bewiesen.

### 19.4.3 Satz: Kern und Bild beim normalen Endomorphismus

Ist  $f : V \rightarrow V$  ein normaler Endomorphismus, so gilt

$$\ker f^\dagger = \ker f$$

$$\operatorname{im} f^\dagger = \operatorname{im} f$$

**19.4.4 Beweis** Für beliebiges  $v \in V$  haben wir aufgrund von

$$\langle f(v), f(v) \rangle = \langle f^\dagger(v), f^\dagger(v) \rangle$$

die Äquivalenz

$$f(v) = \vec{0} \iff f^\dagger(v) = \vec{0},$$

also  $\ker f^\dagger = \ker f$ . Weiter ist dann mit Satz 19.2.5

$$\operatorname{im} f^\dagger = (\ker f)^\perp = (\ker f^\dagger)^\perp = \operatorname{im} f.$$

**19.4.5 Satz: Eigenraum eines normalen Endomorphismus**

Es sei  $V$  ein unitärer oder euklidischer Vektorraum und  $f : V \rightarrow V$  ein normaler Endomorphismus,  $\lambda \in \mathbb{K}$ . Dann gilt

(i)  $f - \lambda$  ist normal.

(ii)  $\text{Eig}(f^\dagger, \bar{\lambda}) = \text{Eig}(f, \lambda)$ .

**19.4.6 Beweis** (i) wird einfach nachgerechnet:

$$\begin{aligned} (f - \lambda)^\dagger \circ (f - \lambda) &= (f^\dagger - \bar{\lambda}) \circ (f - \lambda) = f^\dagger \circ f - f^\dagger \circ \lambda - \bar{\lambda} \circ f + \bar{\lambda} \lambda \\ &= f \circ f^\dagger - \lambda \circ f^\dagger - f \circ \bar{\lambda} + \lambda \bar{\lambda} = (f - \lambda) \circ (f^\dagger - \bar{\lambda}) \\ &= (f - \lambda) \circ (f - \lambda)^\dagger \end{aligned}$$

(ii) folgt mit Satz 19.4.3

$$\text{Eig}(f, \lambda) = \ker(f - \lambda) = \ker(f - \lambda)^\dagger = \ker(f - \bar{\lambda}) = \text{Eig}(f^\dagger, \bar{\lambda})$$

**19.4.7 Fortsetzung des Beweises von Satz 19.4.1**

(B)  $\Rightarrow$  (E) Es seien  $\lambda_1, \dots, \lambda_n \in \mathbb{C}$  die nicht-notwendig verschiedenen Eigenwerte von  $f$ .

Es sei  $v_1 \in \text{Eig}(f, \lambda_1) \setminus \{\vec{0}\}$  und dann  $W_1 \subseteq V$ , so dass

$$V = \langle v_1 \rangle \oplus W_1, \quad \langle v_1 \rangle \perp W_1.$$

Wir zeigen, dass  $f(W_1) \subseteq W_1$  und  $f^\dagger(W_1) \subseteq W_1$ . Für  $w \in W_1$  gilt nämlich

$$\begin{aligned} \langle f(w), v_1 \rangle &= \langle w, f^\dagger(v_1) \rangle = \langle w, \bar{\lambda}_1 v_1 \rangle = \lambda_1 \langle w, v_1 \rangle = 0 \\ \langle f^\dagger(w), v_1 \rangle &= \langle w, f(v_1) \rangle = \langle w, \lambda_1 v_1 \rangle = \bar{\lambda}_1 \langle w, v_1 \rangle = 0. \end{aligned}$$

Es ist dann auch die Einschränkung von  $f$  auf  $W_1$  normal und wir haben insgesamt das Problem „eine Dimension“ tiefer gesetzt.

Diesen Prozess kann man — per Induktion — immer weiter fortsetzen, bis  $V$  ganz in orthogonale Eigenräume zerlegt ist.

(E)  $\Rightarrow$  (C) Sind  $v_i, v_j$  Eigenvektoren aus der geordneten ONB mit  $f(v_i) = \lambda_i v_i$  und  $f(v_j) = \lambda_j v_j$ , so gilt

$$\begin{aligned} \langle f(v_i), f(v_j) \rangle &= \langle \lambda_i v_i, \lambda_j v_j \rangle = \lambda_i \bar{\lambda}_j \langle v_i, v_j \rangle = \lambda_i \bar{\lambda}_j \delta_{ij} \\ \langle f^\dagger(v_i), f^\dagger(v_j) \rangle &= \langle \bar{\lambda}_i v_i, \bar{\lambda}_j v_j \rangle = \bar{\lambda}_i \lambda_j \langle v_i, v_j \rangle = \bar{\lambda}_i \lambda_j \delta_{ij} \end{aligned}$$

Eine getrennte Betrachtung für  $i \neq j$  und  $i = j$  zeigt auf, dass die beiden rechten Seiten — und damit die beiden linken Seiten — dieser beiden Gleichungsketten übereinstimmen.

Per Sesquilinearität gilt diese Gleichung dann für alle  $u, v \in V$

$$\langle f(u), f(v) \rangle = \langle f^\dagger(u), f^\dagger(v) \rangle$$

Das ist (C).

## 19.5 Unitäre Endomorphismen

### 19.5.1 Definition und Satz: Unitäre Endomorphismen

Für einen Endomorphismus  $f$  im unitären Vektorraum sind die folgenden Aussagen äquivalent:

(A)  $f$  heißt *unitär*.

(B)  $f$  ist bijektiv mit  $f^{-1} = f^\dagger$ .

(C) Für alle  $u, v \in V$  gilt

$$\langle f(u), f(v) \rangle = \langle u, v \rangle.$$

(D) Für alle  $v \in V$  gilt

$$\|f(v)\| = \|v\|.$$

(E)  $f$  bildet eine ONB auf eine ONB ab.

(F) Ist  $B$  eine geordnete Basis von  $V$ , so ist die Darstellungsmatrix  $f_{BB}$  unitär, d.h. sie ist regulär mit

$$f_{BB}^{-1} = \overline{f_{BB}^T}.$$

(G)  $f$  ist normal, alle Eigenwerte haben den Betrag 1.

### 19.5.2 Beweis

(B)  $\Rightarrow$  (C) Es ist

$$\langle f(u), f(v) \rangle = \langle u, f^\dagger(f(v)) \rangle = \langle u, v \rangle.$$

(C)  $\Rightarrow$  (B) Aufgrund von (C) gilt die Kette von Implikationen

$$f(v) = \vec{0} \implies \langle f(v), f(v) \rangle = 0 \implies \langle v, v \rangle = 0 \implies v = \vec{0}.$$

Daraus folgt  $\ker f = \{\vec{0}\}$  und dann  $f$  injektiv und dann  $f$  bijektiv.

Für die Vektoren einer Orthonormalbasis gilt dann

$$\begin{aligned} \langle f^\dagger(f(v_i)), v_j \rangle &= \langle f(v_i), f(v_j) \rangle = \langle v_i, v_j \rangle \\ \langle f(f^\dagger(v_i)), v_j \rangle &= \langle f(f^\dagger(v_i)), f(f^{-1}(v_j)) \rangle = \langle f^\dagger(v_i), f^{-1}(v_j) \rangle \\ &= \langle v_i, f(f^{-1}(v_j)) \rangle = \langle v_i, v_j \rangle \end{aligned}$$

und deshalb mit Beobachtung 19.1.1

$$f^\dagger \circ f = \text{id}_V \quad f \circ f^\dagger = \text{id}_V.$$

(C)  $\Rightarrow$  (D) ist ganz einfach:

$$\|f(v)\| = \sqrt{\langle f(v), f(v) \rangle} = \sqrt{\langle v, v \rangle} = \|v\|.$$

Die Umkehrung (D)  $\Rightarrow$  (C) scheint anschaulich klar zu sein, der Beweis erfordert aber den „Quadrat-Trick“. Es gilt

$$\begin{aligned}\|u + v\|^2 &= \langle u + v, u + v \rangle = \langle u, u \rangle + 2\langle u, v \rangle + \langle v, v \rangle \\ &= \|u\|^2 + 2\langle u, v \rangle + \|v\|^2 \\ \|f(u) + f(v)\|^2 &= \langle f(u) + f(v), f(u) + f(v) \rangle \\ &= \langle f(u), f(u) \rangle + 2\langle f(u), f(v) \rangle + \langle f(v), f(v) \rangle \\ &= \|f(u)\|^2 + 2\langle f(u), f(v) \rangle + \|f(v)\|^2.\end{aligned}$$

Aufgrund von (D) stimmen die Normquadrate in der oberen und unteren Gleichungskette jeweils überein. Also müssen auch die Skalarprodukte in der oberen und unteren Gleichungskette übereinstimmen.

Die Äquivalenz (B)  $\Leftrightarrow$  (F) folgt sofort mit Satz 19.3.1.

(C)  $\Rightarrow$  (E) ist trivial.

(E)  $\Rightarrow$  (C) Gilt für die Vektoren  $v_i, v_j$  einer ONB

$$\langle v_i, v_j \rangle = \delta_{ij} \implies \langle f(v_i), f(v_j) \rangle = \delta_{ij},$$

so heißt dies nichts anderes als

$$\langle f(v_i), f(v_j) \rangle \langle v_i, v_j \rangle = \langle v_i, v_j \rangle.$$

Mit der Bilinearität des Skalarprodukts folgt (C).

(B)  $\Rightarrow$  (G)  $f$  ist wegen  $f \circ f^\dagger = \text{id}_V = f^\dagger \circ f$  normal. Ist  $\lambda$  Eigenwert von  $f$ , so gilt für  $v \in \text{Eig}(f, \lambda) = \text{Eig}(f^\dagger, \bar{\lambda})$

$$v = f^\dagger(f(v)) = f^\dagger(\lambda v) = \lambda f^\dagger(v) = \lambda \bar{\lambda} v,$$

und damit  $|\lambda|^2 = \lambda \bar{\lambda} = 1$ .

(G)  $\Rightarrow$  (B) Da  $f$  normal ist, existiert eine geordnete ONB  $(v_1, \dots, v_n)$  aus Eigenvektoren mit

$$f(v_i) = \lambda_i v_i, \quad |\lambda_i| = 1.$$

Es gilt dann für  $i, j \in \{1, \dots, n\}$

$$\begin{aligned}\langle f(v_i), f(v_j) \rangle &= \langle \lambda_i v_i, \lambda_j v_j \rangle = \lambda_i \bar{\lambda}_j \langle v_i, v_j \rangle = \lambda_i \bar{\lambda}_j \delta_{ij} \\ &= \lambda_i \bar{\lambda}_i \delta_{ij} = \delta_{ij} = \langle v_i, v_j \rangle\end{aligned}$$

Mit der Bilinearität des Skalarprodukts folgt (C).

### 19.5.3 Zusatz

Bei den Aussagen (C) und (D) muss die Linearität gar nicht vorausgesetzt werden. Sie folgt automatisch.



### 19.5.4 Folgerung: Charakterisierung unitärer Matrizen

Für eine Matrix in  $A \in \mathbb{C}^{n \times n}$  sind die folgenden Aussagen äquivalent:

- (A)  $A$  ist unitär.
- (B)  $A$  ist regulär mit  $A^{-1} = \overline{A^T} = (\overline{A})^T$ .
- (C) Die Spaltenvektoren von  $A$  bilden eine ONB für  $(\mathbb{C}^n, \langle \cdot, \cdot \rangle)$ , wobei  $\langle \cdot, \cdot \rangle$  das Standard-Skalarprodukt ist.
- (D) Die Zeilenvektoren von  $A$  bilden eine ONB für  $(\mathbb{C}^{1 \times n}, \langle \cdot, \cdot \rangle)$ , wobei  $\langle \cdot, \cdot \rangle$  das Standard-Skalarprodukt ist.
- (E)  $A$  ist normal, alle Eigenwerte haben den Betrag 1.
- (F) Es existiert eine unitäre Matrix  $T$ , so dass

$$D = T \cdot A \cdot T^{-1}$$

eine Diagonalmatrix mit Diagonalelementen vom Betrag 1 ist.

### 19.5.5 Beweis

Im wesentlichen muss nur Satz 19.5.1 auf diese Situation angewandt werden. Die Aussage (C) dieses Satzes ist eine Umformulierung von Aussage (E) des letzten Satzes. Die in (F) konstatierte Eigenschaft „ $T$  unitär“ folgt aus (E) dadurch, dass  $T^{-1}$  die Eigenvektoren als Spalten enthält. Diese bilden aber eine Orthonormalbasis. Mit  $T^{-1}$  ist auch  $T$  unitär.

### 19.5.6 Satz: Die unitäre Gruppe $U(V)$

Die Menge der unitären Abbildungen in einem unitären Vektorraum  $V$  bildet unter Hintereinanderausführung eine Gruppe. Sie wird mit  $U(V)$  bezeichnet.

### 19.5.7 Beweis

Mit zwei unitären Abbildungen  $f, g$  ist auch die Hintereinanderausführung unitär, da

$$\langle f(g(u)), f(g(v)) \rangle = \langle g(u), g(v) \rangle = \langle u, v \rangle.$$

Die identische Abbildung ist ohne Zweifel unitär, sie ist das neutrale Element der Gruppe. Es sei  $f(u) = \vec{0}$ . Dann gilt:

$$\langle u, u \rangle = \langle f(u), f(u) \rangle = \langle \vec{0}, \vec{0} \rangle = 0,$$

wegen der Positiv-Definitheit des Skalarprodukts folgt  $u = \vec{0}$ . Das aber bedeutet, dass die Abbildung  $f$  injektiv ist. Eine injektive lineare Abbildung innerhalb eines endlich-dimensionalen Vektorraums  $V$  ist aber bijektiv. Die Umkehrabbildung  $f^{-1}$  ist das zu  $f$  inverse Element in  $U(V)$ . Diese ist ebenfalls unitär, da

$$\langle f^{-1}(u), f^{-1}(v) \rangle = \langle f(f^{-1}(u)), f(f^{-1}(v)) \rangle = \langle u, v \rangle.$$

**19.5.8 Satz: Die spezielle unitäre Gruppe  $\mathbf{SU}(V)$** 

- (i) Für einen unitären Endomorphismus  $f$  gilt:  $|\det f| = 1$ .
- (ii) Die Teilmenge der unitären Abbildungen mit Determinante  $+1$  bildet eine Untergruppe von  $\mathbf{U}(V)$ . Sie wird mit  $\mathbf{SU}(V)$  bezeichnet.

**19.5.9 Beweis**

- (i) Mit Folgerung 19.3.3 gilt

$$|\det f|^2 = \det f \cdot \overline{\det f} = \det f \cdot \overline{\det f} = \det f \cdot \det f^\dagger = \det(f \circ f^\dagger) = 1.$$

- (ii) Für zwei Abbildungen  $f, g \in \mathbf{SU}(V)$  gilt

$$\det(f \circ g) = \det f \cdot \det g = 1 \cdot 1 = 1,$$

also ist auch  $f \circ g$  in  $\mathbf{SU}(V)$ .

## 19.6 Hermitesche und schief-hermitesche Endomorphismen

### 19.6.1 Definition und Satz: Hermitesche Endomorphismen

Für einen Endomorphismus  $f$  im unitären Vektorraum sind die folgenden Aussagen äquivalent:

(A)  $f$  heißt *hermitesch*.

(B)  $f^\dagger = f$ .

(C) Für alle  $u, v \in V$  gilt

$$\langle f(u), v \rangle = \langle u, f(v) \rangle.$$

(D) Ist  $B$  eine geordnete Orthonormalbasis von  $V$ , so ist die Darstellungsmatrix hermitesch

$$f_{BB} = \overline{f_{BB}^T}.$$

(E)  $f$  ist normal, alle Eigenwerte sind reell.

(F) Es existiert eine unitäre Matrix  $T$ , so dass

$$D = T \cdot A \cdot T^{-1}$$

eine reelle Diagonalmatrix ist.

### 19.6.2 Beweis

Die Äquivalenzen (B)  $\Leftrightarrow$  (C)  $\Leftrightarrow$  (D) sind in Satz 19.3.1 enthalten.

(B)  $\Rightarrow$  (E) Wegen  $f^\dagger = f$  ist  $f$  normal. Ist  $\lambda$  Eigenwert von  $f$ , so gilt mit Satz 19.4.5 (ii)

$$\text{Eig}(f, \lambda) = \text{Eig}(f^\dagger, \bar{\lambda}) = \text{Eig}(f, \bar{\lambda}).$$

Es folgt, dass  $\lambda \in \mathbb{R}$ .

(E)  $\Rightarrow$  (B) Da  $f$  normal ist, existiert eine ONB aus Eigenvektoren von  $f$ . Für  $v \in \text{Eig}(f, \lambda_i) = \text{Eig}(f^\dagger, \bar{\lambda}_i) = \text{Eig}(f^\dagger, \lambda_i)$  gilt

$$f(v) = \lambda_i v = f^\dagger(v).$$

$f$  und  $f^\dagger$  stimmen also für die Eigenvektoren aus der ONB überein. Dann stimmen die Abbildungen  $f$  und  $f^\dagger$  selbst überein.

Die in (F) konstatierte Eigenschaft „ $T$  unitär“ folgt aus (E) so, dass  $T^{-1}$  die Eigenvektoren als Spalten enthält. Diese bilden aber eine Orthonormalbasis. Mit  $T^{-1}$  ist auch  $T$  unitär.

### 19.6.3 Folgerung: Charakterisierung hermitescher Matrizen

Für eine Matrix in  $A \in \mathbb{C}^{n \times n}$  sind die folgenden Aussagen äquivalent:

- (A)  $A$  ist hermitesch.
- (B) Es ist  $A = \overline{A}^T$ .
- (C)  $A$  ist normal, alle Eigenwerte sind reell.
- (D) Es existiert eine unitäre Matrix  $T$ , so dass

$$D = T \cdot A \cdot T^{-1}$$

eine reelle Diagonalmatrix ist.

### 19.6.4 Beweis

Im wesentlichen muss nur Satz 19.6.1 auf diese Situation angewandt werden. Die in (F) konstatierte Eigenschaft „ $T$  unitär“ folgt aus (E) dadurch, dass  $T^{-1}$  die Eigenvektoren als Spalten enthält. Diese bilden aber eine Orthonormalbasis. Mit  $T^{-1}$  ist auch  $T$  unitär.

### 19.6.5 Definition und Satz: Schief-hermitesche Endomorphismen

Für einen Endomorphismus  $f$  im unitären Vektorraum sind die folgenden Aussagen äquivalent:

- (A)  $f$  heißt *schief-hermitesch*
- (B)  $f = -f^\dagger$
- (C) Für alle  $u, v \in V$  gilt

$$\langle f(u), v \rangle = -\langle u, f(v) \rangle$$

- (D) Ist  $B$  eine geordnete Basis von  $V$ , so ist die Darstellungsmatrix schief-hermitesch

$$f_{BB} = -\overline{f_{BB}^T}$$

- (E)  $f$  ist normal, alle Eigenwerte sind rein imaginär.
- (F)  $i \cdot f$  ist hermitesch.

Der Beweis der Äquivalenzen (B)  $\Leftrightarrow \dots \Leftrightarrow$  (F) erfolgt vollkommen analog zu dem von Satz 19.6.1. Die Äquivalenz zu (F) ist ganz einfach nachzurechnen.

**19.6.6 Folgerung: Charakterisierung schief-hermitescher Matrizen**

Für eine Matrix in  $A \in \mathbb{C}^{n \times n}$  sind die folgenden Aussagen äquivalent:

- (A)  $A$  ist schief-hermitesch.
- (B) Es ist  $A = -\overline{A}^T$ .
- (C)  $A$  ist normal, alle Eigenwerte sind rein imaginär.
- (D) Es existiert eine unitäre Matrix  $T$ , so dass

$$D = T \cdot A \cdot T^{-1}$$

eine rein imaginäre Diagonalmatrix ist.

## 20 Endomorphismen im euklidischen Vektorraum

### 20.1 Orthogonale Endomorphismen

#### 20.1.1 Satz: Charakterisierung orthogonaler Endomorphismen

Die folgenden Aussagen über eine Abbildung  $f : V \rightarrow V$  in einem (reellen) euklidischen Vektorraum sind äquivalent:

- (A)  $f$  heißt orthogonal.
- (B)  $f$  ist bijektiv mit  $f^{-1} = f^\dagger$ .
- (C) Das Skalarprodukt zweier beliebiger Vektoren  $u, v \in V$  wird durch  $f$  nicht verändert:

$$\langle f(u), f(v) \rangle = \langle u, v \rangle.$$

- (D) Die Länge (Norm) eines beliebigen Vektors  $v \in V$  wird durch  $f$  nicht verändert:

$$\|f(v)\| = \|v\| \quad \text{für alle } v \in V.$$

- (E)  $f$  bildet eine ONB auf eine ONB ab.
- (F) Ist  $B$  eine geordnete Basis von  $V$ , so ist die Darstellungsmatrix  $f_{BB}$  orthogonal, d.h. sie ist regulär mit

$$f_{BB}^{-1} = f_{BB}^T.$$

#### 20.1.2 Beweis

Er erfolgt vollkommen analog zu dem von Satz 19.5.1.

#### 20.1.3 Folgerung

Ist eine Abbildung  $f : V \rightarrow V$  orthogonal, so wird der Cosinus des Zwischenwinkels von zwei beliebigen Vektoren durch  $f$  nicht verändert:

$$\cos \sphericalangle(f(u), f(v)) = \cos \sphericalangle(u, v) \quad \text{für alle } u, v \in V.$$

#### 20.1.4 Beweis

Das ist klar aufgrund der Beziehung zwischen Cosinus und Skalarprodukt

$$\cos \sphericalangle(f(u), f(v)) = \frac{\langle f(u), f(v) \rangle}{\sqrt{\langle f(u), f(u) \rangle} \sqrt{\langle f(v), f(v) \rangle}} = \frac{\langle u, v \rangle}{\sqrt{\langle u, u \rangle} \sqrt{\langle v, v \rangle}} = \cos \sphericalangle(u, v).$$

**20.1.5 Folgerung: Charakterisierung orthogonaler Matrizen**

Für eine Matrix in  $A \in \mathbb{R}^{n \times n}$  sind die folgenden Aussagen äquivalent:

- (A)  $A$  ist orthogonal.
- (B)  $A$  ist regulär mit  $A^{-1} = A^T$ .
- (C) Die Spaltenvektoren von  $A$  bilden eine ONB für  $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ , wobei  $\langle \cdot, \cdot \rangle$  das Standard-Skalarprodukt ist.
- (D) Die Zeilenvektoren von  $A$  bilden eine ONB für  $(\mathbb{R}^{1 \times n}, \langle \cdot, \cdot \rangle)$ , wobei  $\langle \cdot, \cdot \rangle$  das Standard-Skalarprodukt ist.

**20.1.6 Beweis**

Er erfolgt vollkommen analog zu dem von Folgerung 19.5.4.

**20.1.7 Satz: Die orthogonale Gruppe  $\mathbf{O}(V)$** 

Die Menge der orthogonalen Abbildungen in einem euklidischen Vektorraum  $V$  bildet unter Hintereinanderausführung eine Gruppe. Sie wird mit  $\mathbf{O}(V)$  bezeichnet.

**20.1.8 Beweis**

Analog zu dem Beweis des Satzes 19.5.6.

**20.1.9 Satz: Die spezielle orthogonale Gruppe  $\mathbf{SO}(V)$** 

- (i) Für einen orthogonalen Endomorphismus  $f$  gilt  $\det f \in \{+1, -1\}$ .
- (ii) Die Teilmenge von  $\mathbf{O}(V)$  der orthogonalen Abbildungen mit Determinante  $+1$  bildet eine Untergruppe von  $\mathbf{O}(V)$ . Sie wird mit  $\mathbf{SO}(V)$  bezeichnet.

**20.1.10 Beweis**

Analog zum Beweis des Satzes 19.5.8.

## 20.2 Orthogonale Endomorphismen im euklidischen $\mathbb{R}^2$

### 20.2.1 Einführung

Es sei

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

ein orthogonaler Endomorphismus in dem mit dem Standard-Skalarprodukt versehenen euklidischen Vektorraum  $V = \mathbb{R}^2$ .

Überlegen Sie mittels Anschauung: Wie schaut eine längen- und winkelerhaltende lineare Abbildung im  $\mathbb{R}^2$  aus?

### 20.2.2 Überlegung

(1) Die Orthogonalitätsbedingung an den Endomorphismus lässt sich in den beiden Formeln

$$I = A^T \cdot A \quad \text{und} \quad I = A \cdot A^T$$

darstellen. Ausgeschrieben lauten diese beiden Bedingungen

$$\begin{aligned} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} &= \begin{pmatrix} a & c \\ b & d \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{pmatrix} \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{pmatrix}, \end{aligned}$$

wir erhalten insgesamt sechs Gleichungen

$$\begin{aligned} a^2 + c^2 &= 1 \\ b^2 + d^2 &= 1 \\ a^2 + b^2 &= 1 \\ c^2 + d^2 &= 1 \\ ab + cd &= 0 \\ ac + bd &= 0 \end{aligned}$$

(2) Daraus folgen die Gleichungen

$$a^2 = d^2 \quad b^2 = c^2$$

und damit die vier Fälle

$$a = \pm d \quad b = \pm c.$$

(3) Die beiden Fälle

$$\begin{aligned} a = +d \neq 0 \quad \text{UND} \quad b = +c \neq 0 \\ a = -d \neq 0 \quad \text{UND} \quad b = -c \neq 0 \end{aligned}$$



widersprechen den Gleichungen mit den „gemischten“ Gliedern in (1). Es können also nur die beiden anderen Fälle

$$a = +d \quad \text{UND} \quad b = -c \quad \text{ODER}$$

$$a = -d \quad \text{UND} \quad b = +c$$

eintreten.

(4) Die Aussage in (3) bedeutet, dass die Matrix  $A$  die Form

$$A = \underbrace{\begin{pmatrix} a & -c \\ c & a \end{pmatrix}}_{=:D} \quad \text{oder} \quad A = \begin{pmatrix} a & c \\ c & -a \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}_{=:S} \underbrace{\begin{pmatrix} a & -c \\ c & a \end{pmatrix}}_{=:D}$$

mit  $a^2 + c^2 = 1$  hat.

(5) Man kann umgekehrt nachrechnen, dass Matrizen dieser Form orthogonal sind.

### 20.2.3 Satz und Definition: Charakterisierung von Drehmatrizen

Die folgenden Aussagen über eine Matrix  $D \in \mathbb{R}^{2 \times 2}$  sind äquivalent:

- (A) (Def)  $D$  heißt *Drehmatrix* im  $\mathbb{R}^2$ .
- (B)  $D \in \mathbf{SO}(\mathbb{R}^2)$ , d.h.  $D$  ist orthogonal mit Determinante  $+1$ .
- (C) Es ist

$$D = \begin{pmatrix} a & -c \\ c & a \end{pmatrix} \quad \text{mit } a, c \in \mathbb{R} \quad \text{und} \quad a^2 + c^2 = 1.$$

- (D) Es existiert ein  $\varphi \in \mathbb{R}$ , so dass

$$D = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$$

Diese Matrix dreht einen beliebigen Vektor um den Winkel  $\varphi$  (in Richtung  $e_2 \rightarrow e_1$ ), das heißt  $D$  ist eine Drehmatrix.

- (E) Wird  $D$  als Matrix in  $\mathbb{C}^{2 \times 2}$  aufgefasst, so hat  $D$  zwei Eigenwerte und Eigenvektoren (über  $\mathbb{C}$ ):

$$D \begin{pmatrix} 1 \\ i \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ i \end{pmatrix}, \quad D \begin{pmatrix} 1 \\ -i \end{pmatrix} = \lambda_2 \begin{pmatrix} 1 \\ -i \end{pmatrix},$$

wobei

$$\lambda_1 = \bar{\lambda}_2 \quad \text{und} \quad |\lambda_1| = |\lambda_2| = 1$$

gilt. Insbesondere ist  $D$  über  $\mathbb{C}$  diagonalisierbar.

(F) „Identifiziert“ man die Ebene  $\mathbb{R}^2$  mit der komplexen Zahlenebene  $\mathbb{C}$  mittels der bijektiven Abbildung

$$\begin{cases} \mathbb{R}^2 & \rightarrow \mathbb{C} \\ \begin{pmatrix} x \\ y \end{pmatrix} & \mapsto x + iy, \end{cases}$$

so korrespondiert zum Endomorphismus  $D : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  die Multiplikation in  $\mathbb{C}$  mit einer Zahl

$$d = \cos \varphi + i \sin \varphi, \quad \text{mit } \varphi \in \mathbb{R}.$$

auf dem Einheitskreis.

Zusatz: Die beiden Zahlen  $\varphi$  in (D) und (F) können übereinstimmend gewählt werden.

#### 20.2.4 Beweis

(B)  $\Rightarrow$  (C): Das wurde in den dem Satz vorhergehenden Überlegungen dargelegt.

(C)  $\Rightarrow$  (B) Die Gleichung  $D^{-1} = D^T$  ist leicht nachzurechnen.

(C)  $\Leftrightarrow$  (D): Das ist letztlich eine Aussage der Analysis.

(C)  $\Rightarrow$  (E): Setze  $\lambda_1 = a - ic$  und  $\lambda_2 = a + ic$ .

(D)  $\Leftrightarrow$  (F): Es ist

$$D \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \varphi x - \sin \varphi y \\ \sin \varphi x + \cos \varphi y \end{pmatrix}$$

$$(\cos \varphi + i \sin \varphi) \cdot (x + iy) = (\cos \varphi x - \sin \varphi y) + i(\sin \varphi x + \cos \varphi y).$$

(E)  $\Rightarrow$  (C) Wir setzen  $D$  als reelle  $2 \times 2$ -Matrix an:

$$D = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$$

Mit  $\lambda_1 = a - ic$  gilt dann:

$$\begin{pmatrix} x + iy \\ z + iw \end{pmatrix} = \begin{pmatrix} x & y \\ z & w \end{pmatrix} \begin{pmatrix} 1 \\ i \end{pmatrix} = (a - ic) \begin{pmatrix} 1 \\ i \end{pmatrix} = \begin{pmatrix} a - ic \\ c + ia \end{pmatrix}.$$

Daraus folgt, dass  $D$  die Darstellung in (B) hat. Es ist weiter  $a^2 + c^2 = |\lambda_1|^2 = 1$ .

#### 20.2.5 Diagonalisierbarkeit

Diagonalisierbarkeit über  $\mathbb{R}$  einer Matrix aus  $\mathbf{SO}(\mathbb{R}^2)$  liegt genau dann vor, wenn die in Aussage (E) genannten Eigenwerte reell sind:

$$\lambda_1 = \lambda_2 \in \{-1, +1\}.$$

In diesen Fällen ergibt sich:

$$D = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I \quad \text{bzw.} \quad D = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I.$$

Die zweite Abbildung ist die *Punktspiegelung* mit  $\vec{0}$  als Zentrum.

### 20.2.6 Definition Spiegelungsmatrix

Eine Matrix  $S \in \mathbb{R}^{2 \times 2}$  heißt *Spiegelungsmatrix*, wenn sie die beiden Eigenwerte  $+1$  und  $-1$  hat. Der Eigenraum  $\text{Eig}(S, +1)$  heißt die *Spiegelungsachse*.

### 20.2.7 Satz: Charakterisierung von orthogonalen Matrizen im $\mathbb{R}^2$

Die folgenden Aussagen über einen Endomorphismus  $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  sind äquivalent:

(A)  $A \in \mathbf{O}(\mathbb{R}^2)$ , d.h.  $A$  ist orthogonal. Siehe auch Folgerung 20.1.5.

(B) Es ist

$$A = D \quad \text{oder} \quad A = D \cdot S,$$

wobei  $D$  eine Drehmatrix und  $S$  eine Spiegelungsmatrix ist.

(C) Es ist

$$A = D \quad \text{oder} \quad A = S \cdot D,$$

wobei  $D$  eine Drehmatrix und  $S$  eine Spiegelungsmatrix ist.

(D) Es ist

$$A = \underbrace{\begin{pmatrix} a & -c \\ c & a \end{pmatrix}}_{=:D} \quad \text{oder} \quad A = \begin{pmatrix} a & c \\ c & -a \end{pmatrix}$$

mit  $a^2 + c^2 = 1$ .

### 20.2.8 Beweis

Die Äquivalenzen  $(A) \Leftrightarrow (B) \Leftrightarrow (D)$  sind in den Überlegungen in Abschnitt 20.2.2 begründet. Die Äquivalenz zu (C) ist leicht nachzurechnen.

## 20.3 Normalform eines orthogonalen Endomorphismus

### 20.3.1 Satz: Invarianz von orthogonalen Komplementen

Es sei  $V$  ein euklidischer Vektorraum. Ist  $f$  selbstadjungiert, schief-selbstadjungiert oder orthogonal, so gilt

$$U \text{ } f\text{-invariant} \quad \implies \quad U^\perp \text{ } f\text{-invariant.}$$

### 20.3.2 Beweis

Es sei  $U$   $f$ -invariant und  $v \in U^\perp$  beliebig. Die Aussage  $f(U^\perp) \subseteq U^\perp$  ist jeweils bewiesen, wenn wir

$$\langle f(v), u \rangle = 0 \quad \text{für alle } u \in U$$

zeigen können.

- (1) Ist  $f$  selbstadjungiert, so gilt für alle  $u \in U$ :

$$\langle f(v), u \rangle = \langle v, \underbrace{f(u)}_{\in U} \rangle = 0.$$

- (2) Wenn  $f$  schief-selbstadjungiert ist, so geht das (fast) ganz genauso:

$$\langle f(v), u \rangle = \langle v, \underbrace{-f(u)}_{\in U} \rangle = 0.$$

- (3) Es sei jetzt  $f$  orthogonal. Da  $f$  bijektiv ist, gilt allgemein  $\dim f(U) = \dim U$ . Deswegen folgt aus  $f(U) \subseteq U$  sogar  $f(U) = U$ , damit gilt aber auch  $f^{-1}(U) = U$ . Siehe Fußnote <sup>1</sup>

Jetzt können wir schließen:

$$\langle f(v), u \rangle = \langle v, f^\dagger(u) \rangle = \langle v, \underbrace{f^{-1}(u)}_{\in U} \rangle = 0.$$

### 20.3.3 Beispiel

Im  $\mathbb{R}^2$  mit dem Standard-Skalarprodukt sind die beiden Unterräume

$$U = \langle e_1 \rangle, \quad U^\perp = \langle e_2 \rangle$$

orthogonale Komplemente zueinander. Die Matrix

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

lässt  $U = \langle e_1 \rangle$  invariant, nicht aber  $U^\perp = \langle e_2 \rangle$ .

<sup>1</sup>Ist  $f : \mathbb{R} \rightarrow \mathbb{R}$  die bijektive Halbierungsabbildung, so gilt  $2\mathbb{Z} \subsetneq f(2\mathbb{Z})$ .



Es sei nun  $\widehat{A} = g_{\widehat{B}\widehat{B}}$  die Darstellungsmatrix für  $g$  bzgl. einer geordneten ONB  $\widehat{B}$  von  $W$ .

(3) Zu dem echt komplexen Eigenwert  $\lambda = \lambda_1$  gibt es einen komplexen Eigenvektor  $w$ , so dass

$$\widehat{A} w = \lambda w \quad \widehat{A} \bar{w} = \bar{\lambda} \bar{w}.$$

Es seien die reellen Vektoren  $u, v$  definiert durch

$$u := \operatorname{Re} w \quad v := \operatorname{Im} w.$$

(4) Wir zeigen, dass  $\langle u, v \rangle = 0$ . Zunächst gilt:

$$\lambda^2 \cdot \langle w, w \rangle = \langle \lambda w, \lambda w \rangle = \langle f(w), f(w) \rangle = \langle w, w \rangle.$$

Da  $\lambda^2 \neq 1$  ist, muss  $\langle w, w \rangle = 0$  sein. Dies ist kein Widerspruch zur Positiv-Definitheit des bilinearen euklidischen Skalarprodukts, da  $w$  ein komplexer Vektor ist.

Weiter ist dann

$$4i \langle u, v \rangle = \langle u + iv, u + iv \rangle - \langle u - iv, u - iv \rangle = \langle w, w \rangle - \overline{\langle w, w \rangle} = 0.$$

(5) Außerdem gilt:

$$\begin{aligned} \widehat{A} u &= \widehat{A}(\operatorname{Re} w) = \operatorname{Re}(\widehat{A} w) = \operatorname{Re}(\lambda w) = (\operatorname{Re} \lambda)u - (\operatorname{Im} \lambda)v \\ \widehat{A} v &= \widehat{A}(\operatorname{Im} w) = \operatorname{Im}(\widehat{A} w) = \operatorname{Im}(\lambda w) = (\operatorname{Im} \lambda)u + (\operatorname{Re} \lambda)v. \end{aligned}$$

Der durch  $u, v$  aufgespannte Unterraum von  $W$  ist also  $\widehat{A}$ -invariant. Die Einschränkung auf diesen Unterraum ist die im Satz beschriebene orthogonale  $2 \times 2$ -Matrix  $A_1$ .

(6) Führt man dieses Verfahren für alle Eigenwerte  $\lambda_2, \dots, \lambda_r$  fort, so erhält man die Aussage des Satzes.

## 20.4 Symmetrische und schiefsymmetrische Endomorphismen

### 20.4.1 Definition und Satz: Symmetrische Endomorphismen

Es sei  $V$  ein euklidischer Vektorraum. Für einen Endomorphismus  $f : V \rightarrow V$  sind die folgenden Aussagen äquivalent:

(A)  $f$  heißt *symmetrisch* oder *selbstadjungiert*.

(B)  $f = f^\dagger$

(C) Für alle  $u, v \in V$  gilt

$$\langle f(u), v \rangle = \langle u, f(v) \rangle$$

(D) Ist  $B$  eine geordnete Orthonormalbasis von  $V$ , so ist die Darstellungsmatrix symmetrisch

$$f_{BB} = f_{BB}^T$$

(E)  $f$  ist diagonalisierbar, alle Eigenwerte sind reell.

### 20.4.2 Beweis

Die Äquivalenzen (B)  $\Leftrightarrow$  (C)  $\Leftrightarrow$  (D) zeigt man wie im Beweis von Satz 19.6.1.

(D)  $\Rightarrow$  (E) Die Matrix  $f_{BB} \in \mathbb{R}^{n \times n}$  ist symmetrisch und daher als Matrix  $f_{BB} \in \mathbb{C}^{n \times n}$  Hermitesch. Nach Satz 19.6.1 ist die Hermitesche Matrix  $f_{BB}$  über  $\mathbb{C}$  diagonalisierbar mit Eigenwerten aus  $\mathbb{R}$ .

Ist  $v \in \text{Eig}(f, \lambda_i) \subseteq \mathbb{C}^n$ , so gilt wegen

$$f_{BB}v = \overline{f_{BB}v} = \overline{\lambda_i v} = \lambda_i \bar{v}$$

auch  $\bar{v} \in \text{Eig}(f, \lambda_i) \subseteq \mathbb{C}^n$ . Das bedeutet aber, dass man die komplexen Eigenvektoren in einer ONB durch ihre Realteile und/oder Imaginärteile ersetzen kann. Es existiert also auch eine ONB aus reellen Eigenvektoren.

Insgesamt ist also  $f_{BB}$  über  $\mathbb{R}$  diagonalisierbar.

(E)  $\Rightarrow$  (D) ist trivial.

### 20.4.3 Definition und Satz: Schiefsymmetrische Endomorphismen

Für einen Endomorphismus  $f$  im unitären Vektorraum sind die folgenden Aussagen äquivalent:

(A)  $f$  heißt *schiefsymmetrisch*

(B)  $f = -f^\dagger$

(C) Für alle  $u, v \in V$  gilt

$$\langle f(u), v \rangle = -\langle u, f(v) \rangle$$

(D) Ist  $B$  eine geordnete ONB von  $V$ , so ist die Darstellungsmatrix schiefsymmetrisch

$$f_{BB} = -f_{BB}^T$$

Der Beweis erfolgt vollkommen analog zu dem von Satz 20.4.1.



## 21 Dualität

### 21.1 Der duale Vektorraum

Es sei  $V$  ein  $\mathbb{K}$ -Vektorraum. Meist werden wir es mit dem Fall  $\dim V < \infty$  zu tun haben. Ein Skalarprodukt muss nicht vorhanden sein.

#### 21.1.1 Definition

Die Menge

$$V^* = \mathcal{L}(V, \mathbb{K}) = \{a \mid a \text{ ist lineare Abbildung von } V \text{ nach } \mathbb{K}\}$$

trägt die Struktur eines  $\mathbb{K}$ -Vektorraums. Er heißt der zu  $V$  *duale Vektorraum* oder *Dualraum*. Seine Elemente heißen *Linearformen* auf (über)  $V$ .

#### 21.1.2 Definition: Duale Basis

Es sei in dem  $n$ -dimensionalen Vektorraum  $V$  eine geordnete Basis  $B = (v_1, \dots, v_n)$  gegeben. In Abhängigkeit davon definieren wir die Linearformen  $v_1^*, \dots, v_n^*$  durch die Vorgabe der Bilder auf der Basis  $B$ :

$$v_j^*(v_k) = \delta_{jk} = \begin{cases} 1 & \rightarrow j = k \\ 0 & \rightarrow j \neq k \end{cases}$$

Man beachte, dass die Definition von  $v_j^*$  von der ganzen geordneten Basis  $B$  — und nicht etwa nur von dem einzelnen Vektor  $v_j \in B$  — abhängt.

#### 21.1.3 Satz und Definition: Duale Basis

Das  $n$ -Tupel  $B^* = (v_1^*, \dots, v_n^*)$  ist eine geordnete Basis von  $V^*$ .

Es heißt die zu  $B$  *duale geordnete Basis*.

#### 21.1.4 Beweis

Für eine beliebige Linearform  $a \in V^*$  machen wir den Ansatz

$$a = \alpha_1 v_1^* + \dots + \alpha_n v_n^*.$$

Bei Anwendung dieser Linearform auf den Basisvektor  $v_k$  gilt

$$\begin{aligned} a(v_k) &= (\alpha_1 v_1^* + \dots + \alpha_n v_n^*)(v_k) \\ &= \alpha_1 v_1^*(v_k) + \dots + \alpha_n v_n^*(v_k) = \alpha_k \end{aligned}$$

Das bedeutet zunächst, dass die  $\alpha_k$  eindeutig bestimmt sind,  $B^*$  ist also linear unabhängig.

Setzt man umgekehrt im Ansatz  $\alpha_k := a(v_k)$ , so ist die Gleichung bei Anwendung auf alle  $v_k$  erfüllt, sie stimmt also als Gleichung zwischen Linearformen. Dies bedeutet, dass  $B^*$  ein Erzeugendensystem für  $V^*$  ist.

#### 21.1.5 Folgerung

- (i) Zu jedem Vektor  $v \in V \setminus \{\vec{0}\}$  gibt es ein  $a \in V^*$  mit  $a(v) \neq 0$ .
- (ii) Es ist  $\dim V^* = n$ .

### 21.1.6 Beweis

(i) Man setze  $v_1 = v$  und ergänze die Menge  $\{v_1\}$  zu einer Basis  $B$  von  $V$ . Für das Element  $v_1^*$  der dualen Basis gilt dann  $v_1^*(v_1) = 1$ .

### 21.1.7 Definition

Ist  $U \subseteq V$  eine Teilmenge von  $V$ , so heißt die Teilmenge

$$U^0 := \{a \in V^* \mid a(u) = 0 \text{ für alle } u \in U\} \subseteq V^*$$

der *Annihilator* von  $U$ .

### 21.1.8 Satz: Dimension des Annihilators

Die Teilmenge  $U^0$  ist ein Unterraum von  $V^*$ . Ist  $U$  ein Unterraum, so gilt

$$\dim U^0 = \dim V - \dim U$$

### 21.1.9 Beweis

(1) Sind  $a, \tilde{a} \in U^0$ , so gilt für alle  $u \in U$

$$(\alpha a + \tilde{\alpha} \tilde{a})(u) = \alpha a(u) + \tilde{\alpha} \tilde{a}(u) = 0,$$

also  $\alpha a + \tilde{\alpha} \tilde{a} \in U^0$ .

(2) Man betrachte eine Basis

$$B_U = \{v_1, \dots, v_r\} \quad \text{von } U$$

und ergänze sie zu einer Basis

$$B_V = \{v_1, \dots, v_r, v_{r+1}, \dots, v_n\} \quad \text{von } V$$

(3) Wir überzeugen uns nun davon, dass die Teilmenge

$$B_{U^0} := \{v_{r+1}^*, \dots, v_n^*\} \subseteq B_V^*$$

eine Basis von  $U^0$  ist.

(3a) Die lineare Unabhängigkeit ist klar.

(3b) Wegen  $v_j^*(v_k) = 0$  für  $j = r+1, \dots, n$  und  $k = 1, \dots, r$  ist  $\langle B_{U^0} \rangle \subseteq U^0$ .

(3c) Ist  $a \in U^0$ , so gibt es eine Basisentwicklung bzgl.  $B_V^*$

$$a = \alpha_1 v_1^* + \dots + \alpha_r v_r^* + \alpha_{r+1} v_{r+1}^* + \dots + \alpha_n v_n^*$$

Setzt man hier die Vektoren  $v_1, \dots, v_r$  ein, so folgt  $\alpha_1 = \dots = \alpha_r = 0$  und deshalb

$$a = \alpha_{r+1} v_{r+1}^* + \dots + \alpha_n v_n^*.$$

Wir haben also  $U^0 \subseteq \langle B_{U^0} \rangle$  gezeigt.

## 21.2 Der duale Homomorphismus

### 21.2.1 Definition: Der duale Homomorphismus

Es sei  $f : V \rightarrow W$  ein Homomorphismus zwischen zwei  $\mathbb{K}$ -Vektorräumen.

Ist  $a \in W^*$  gegeben, so ist die lineare Abbildung

$$a \circ f : \begin{cases} V & \rightarrow \mathbb{K} \\ v & \mapsto a(f(v)) \end{cases}$$

ein Element von  $V^*$ . Wir definieren nun den *dualen Homomorphismus*

$$f^* : \begin{cases} W^* & \rightarrow V^* \\ a & \mapsto a \circ f. \end{cases}$$

Er bildet also Linearformen über  $W$  auf Linearformen über  $V$  ab.

### 21.2.2 Satz: Dualisierung entspricht Transponierung

Es seien  $V, W$  zwei  $\mathbb{K}$ -Vektorräume und  $V^*, W^*$  die dazu dualen  $\mathbb{K}$ -Vektorräume mit  $\dim V = \dim V^* = n$  sowie  $\dim W = \dim W^* = m$ .

Sie seien mit geordneten Basen  $B$  für  $V$ ,  $C$  für  $W$ ,  $B^*$  für  $V^*$  und  $C^*$  für  $W^*$  ausgestattet.

Die Darstellungsmatrix  $f_{CB} \in \mathbb{K}^{m \times n}$  für einen Homomorphismus  $f : V \rightarrow W$  und die Darstellungsmatrix  $f_{B^*C^*}^* \in \mathbb{K}^{n \times m}$  für den dazu dualen Homomorphismus  $f^* : W^* \rightarrow V^*$  sind transponiert zueinander:

$$f_{B^*C^*}^* = f_{CB}^T.$$

### 21.2.3 Beweis

Es seien

$$\begin{aligned} B &= (v_1, \dots, v_n), & C &= (w_1, \dots, w_m) \\ B^* &= (v_1^*, \dots, v_n^*), & C^* &= (w_1^*, \dots, w_m^*) \end{aligned}$$

die im Satz genannten geordneten Basen. Die  $j$ -te Spalte der Matrix  $A = f_{CB}$  enthält die Entwicklungskoeffizienten von  $f(v_j)$  bzgl. der Basis  $C$ :

$$f(v_j) = a_{1j}w_1 + \dots + a_{mj}w_m.$$

Die  $k$ -te Spalte der Matrix  $A^* = f_{B^*C^*}^*$  enthält die Entwicklungskoeffizienten von  $f^*(w_k^*)$  bzgl. der Basis  $B^*$ :

$$f^*(w_k^*) = a_{1k}^*v_1^* + \dots + a_{nk}^*v_n^*.$$

Wir setzen die obere Gleichung ein in  $w_k^*$ , die untere wenden wir an auf  $v_j$ . Dann erhalten wir

$$\begin{aligned} w_k^*(f(v_j)) &= a_{kj} \\ f^*(w_k^*)(v_j) &= a_{jk}^* \end{aligned}$$

Gemäß Definition von  $f^*$  stimmen die linken Seiten überein, deshalb dann  $a_{kj} = a_{jk}^*$ .

### 21.2.4 Satz: Dualisierung ist Involution

(i) Für einen Homomorphismus  $f : V \rightarrow W$  gilt

$$(f^*)^* = f$$

(ii) Für zwei Homomorphismen  $f : V \rightarrow U$  und  $g : U \rightarrow W$  zwischen drei Vektorräumen gilt

$$(g \circ f)^* = f^* \circ g^*$$

Die beiden Eigenschaften drücken aus, dass Dualisierung eine *Involution* ist.

### 21.2.5 Beweis

Beide Aussagen folgen sofort aus Satz 21.2.2.

### 21.2.6 Satz: Kern und Bild des dualen Homomorphismus

Es seien  $V, W$  zwei endlich-dimensionale  $\mathbb{K}$ -Vektorräume und  $V^*, W^*$  die dazu dualen  $\mathbb{K}$ -Vektorräume.

Dann gilt

$$\begin{aligned} (\operatorname{im} f)^* &= (\ker f)^0 \\ (\ker f)^* &= (\operatorname{im} f)^0 \end{aligned}$$

### 21.2.7 Beweis

(1) Es sei  $a \in \operatorname{im} f^*$ . Dann gibt es  $b \in W^*$  mit  $a = f^*(b)$ .

Ist  $v \in \ker f$ , so gilt dann

$$a(v) = (f^*b)(v) = b(f(v)) = 0.$$

Das bedeutet aber  $a|_{\ker f} = 0$ , weiter  $a \in (\ker f)^0$ .

(2) Es sind Basen wie folgt vorhanden

$$\begin{aligned} B &= \{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}, & \ker f &= \langle v_{r+1}, \dots, v_n \rangle \\ C &= \{w_1, \dots, w_r, w_{r+1}, \dots, w_m\}, & \operatorname{im} f &= \langle w_1, \dots, w_r \rangle. \end{aligned}$$

Ist nun  $a \in (\ker f)^0$ , also  $a|_{\ker f} = 0$ , so definieren wird  $b \in W^*$  durch

$$b(w_i) = \begin{cases} a(v_i), & \text{falls } i = 1, \dots, r \\ 0, & \text{falls } i = r + 1, \dots, m. \end{cases}$$

Es ist dann

$$a(v_i) = \begin{cases} b(w_i) = b(f(v_i)), & \text{falls } i = 1, \dots, r \\ 0 = b(f(v_i)), & \text{falls } i = r + 1, \dots, n. \end{cases}$$

Das bedeutet  $a = b \circ f = f^*(b)$  und demzufolge  $a \in \text{im } f^*$ .

(3) Sei  $b \in \text{im } f^*$ , das heißt  $f^*(b) = 0$ . Zu beliebigem  $w \in \text{Im } f$  existiert  $v \in V$  mit  $w = f(v)$ . Dann

$$b(w) = b(f(v)) = f^*(b)(v) = 0.$$

also insgesamt  $b|_{\text{im } f} = 0$ , d.h.  $b \in (\text{im } f)^0$ .

(4) Sei  $b \in (\text{im } f)^0$ , d.h.  $b|_{\text{im } f} = 0$ . Dann ist für alle  $v \in V$

$$f^*(b)(v) = b(f(v)) = 0.$$

Insgesamt also  $f^*(b) = 0$ , demzufolge  $b \in \ker f^*$ .

## 21.3 Der biduale Vektorraum

### 21.3.1 Definition: Die Evaluation

Es sei  $V$  ein  $\mathbb{K}$ -Vektorraum. Ein fester Vektor  $v \in V$  kann in eine beliebige Linearform  $a$  über  $V$  eingesetzt werden. Auf diese Weise ist eine Abbildung

$$\varepsilon_v : \begin{cases} V^* & \rightarrow \mathbb{K} \\ a & \mapsto a(v) \end{cases}$$

definiert. Wir nennen sie *Evaluation (von  $a$ ) an der Stelle  $v$* .

Diese Situation ist etwas ungewöhnlich. Wir betrachten nicht mehr die Vektoren  $v \in V$  als Objekte und die Linearformen  $a \in V^*$  als Abbildungen, sondern umgekehrt: Die Linearformen  $a \in V^*$  sind die Objekte, die Vektoren  $v \in V$  werden zu Abbildungen  $\varepsilon_v$  auf  $V^*$ .

### 21.3.2 Der Bidualraum

Der Vektorraum der Linearformen auf  $V^*$  heißt der *biduale Vektorraum* oder *Bidualraum*. Er wird — naheliegend — bezeichnet mit

$$V^{**} := (V^*)^*$$

Die Abbildung  $\varepsilon_v$  ist linear und damit ein Element von  $V^{**}$ . Sieht man jetzt auch das  $v$  als variabel an, so haben wir es insgesamt mit einer Abbildung

$$\varepsilon : \begin{cases} V & \rightarrow V^{**} \\ v & \mapsto \varepsilon_v \end{cases}$$

zu tun. Sie ist ohne Zweifel linear.

### 21.3.3 Satz:

Ist  $V$  endlich-dimensional, so ist die Abbildung

$$\varepsilon : V \rightarrow V^{**}$$

ein (natürlicher) Isomorphismus.

### 21.3.4 Beweis

Wir wissen bereits (21.1.5 (ii)), dass  $\dim V^{**} = \dim V^* = \dim V$ .

Ist  $\varepsilon_v$  für ein  $v \in V$  die Null-Linearform auf  $V^*$ , so ist für alle  $b \in V^*$

$$b(v) = \varepsilon_v(b) = 0.$$

Gemäß Folgerung 21.1.5 (i) gibt es aber zu jedem  $v \neq \vec{0}$  eine Linearform in  $b \in V^*$  mit  $b(v) \neq 0$ . Es bleibt also nur, dass  $v = \vec{0}$  ist.

Demzufolge ist  $\varepsilon$  injektiv, wegen der Gleichheit der Dimensionen von  $V$  und  $V^{**}$  dann auch bijektiv.

### 21.3.5 Identifizierung

Es ist allgemein üblich, bei Vorliegen eines natürlichen Isomorphismus  $\varepsilon : M \rightarrow N$  zwischen zwei mathematischen Strukturen (Mengen), die einander durch  $\varepsilon$  zugeordneten Elemente von  $M$  und  $N$  zu identifizieren, d.h. als gleich anzusehen.

Im hier vorliegenden Fall heißt das, dass wir zwischen  $v \in V$  und  $\varepsilon_v \in V^{**}$  nicht unterscheiden. Eine Linearform auf  $V^*$  wird einfach durch den Vektor aus  $V$  beschrieben, der diese Linearform durch „eingesetzt werden“ hervorbringt:

$$\varepsilon_v : \begin{cases} V^* & \rightarrow \mathbb{K} \\ a & \mapsto a(v) \end{cases} = v : \begin{cases} V^* & \rightarrow \mathbb{K} \\ a & \mapsto a(v) \end{cases}$$

Auf diese Weise werden auch die Mengen  $V$  und  $V^{**}$  gleich:

$$V^{**} = V.$$

Der natürliche Isomorphismus zwischen diesen beiden Mengen wird „überflüssig“.

Durch diese Sichtweise werden die Rollen von  $V$  und  $V^*$  „gleichberechtigt“ bzw. „symmetrisch zueinander“. Dies kommt durch die bilineare Abbildung bzw. Schreibweise

$$(\cdot | \cdot) \begin{cases} V \times V^* & \rightarrow \mathbb{K} \\ (v, a) & \mapsto (v|a) = a(v) = \varepsilon_v(a) \end{cases}$$

zum Ausdruck. Im englischen spricht man hier von einem „pairing“ zwischen  $V$  und  $V^*$ .

Beachte aber den Unterschied zum Skalarprodukt auf einem Vektorraum: Das ist eine bilineare Abbildung  $V \times V \rightarrow \mathbb{K}$ .

Der nächste Satz erhält so eine besondere und kompakte Formulierung:

### 21.3.6 Satz: Der Doppel-Annihilator

Für einen Unterraum  $U \subseteq V$  und seinen „Doppel-Annihilator“  $U^{00} \subseteq V^{**}$  gilt

$$U^{00} = U.$$

### 21.3.7 Beweis

Wir schließen mit Satz 21.1.8, dass

$$\dim U^{00} = \dim V - \dim U^0 = \dim V - (\dim V - \dim U) = \dim U.$$

Es sei jetzt  $u \in U$  und  $a \in U^0$  beliebig. Dann ist

$$(u|a) = a(u) = 0,$$

da  $a$  im Annihilator von  $U$  enthalten ist. Die umgekehrte Deutung ist aber, dass  $u$  im Annihilator von  $U^0$  enthalten ist. Insgesamt ist so  $U \subseteq U^{00}$  und dann aufgrund der Dimensionsgleichheit  $U = U^{00}$ .

## 22 Affine Räume

### 22.1 Einführung

Was sind freie Vektoren, Verschiebungsvektoren, Ortsvektoren, Punkte, Klassen paralleler Pfeile, ...?

#### 22.1.1 Definitionen: Affiner Raum

Eine Menge  $\mathcal{P}$  heißt (reeller) *affiner Raum der Dimension  $n$* , wenn es einen reellen Vektorraum  $V$  der Dimension  $n$  gibt und eine Abbildung

$$\tau : \begin{cases} V \times \mathcal{P} & \rightarrow \mathcal{P} \\ (v, P) & \mapsto \tau_v(P) =: P + v, \end{cases}$$

so dass die folgenden drei Eigenschaften erfüllt sind:

- (1) Für jedes feste  $v \in V$  ist  $\tau_v : \mathcal{P} \rightarrow \mathcal{P}$  eine bijektive Abbildung.
- (2) Zu  $P, Q \in \mathcal{P}$  gibt es genau ein  $v \in V$ , so dass

$$\tau_v(P) = Q.$$

Dieser *Verbindungsvektor*  $v$  wird dann mit  $\overrightarrow{PQ}$  bezeichnet.

- (3) Für  $v, w \in V$  gilt:

$$\tau_v \circ \tau_w = \tau_{v+w}.$$

#### 22.1.2 Bemerkungen

- (1) Die Elemente von  $\mathcal{P}$  heißen die *Punkte* des affinen Raumes.
- (2) Die Abbildung  $\tau_v$  heißt *Translation* oder *Verschiebung* (in  $\mathcal{P}$ ) um den Vektor  $v$ .
- (3) Deshalb kann der Vektorraum  $V$  auch als der *Vektorraum der Translationen* von  $\mathcal{P}$  bezeichnet werden.
- (4) Man kann  $\tau$  auch auffassen als Abbildung

$$\begin{cases} V & \rightarrow \mathcal{F}(\mathcal{P}, \mathcal{P}) \quad (= \text{Menge der Abbildungen } \mathcal{P} \rightarrow \mathcal{P}) \\ v & \mapsto \tau_v \end{cases}$$

Man sagt dann auch, dass der Vektorraum  $V$  auf  $\mathcal{P}$  *operiert*.

- (5) Die Schreibweise in der Definition von  $\tau$  macht bereits deutlich, dass man sich die Operation von  $V$  auf  $\mathcal{P}$  vorstellen kann als eine Art Anheften des Vektorpfeils  $v$  an einen Punkt  $P \in \mathcal{P}$ . Das Bild  $Q = P + v$  ist dann durch die Spitze des Pfeils festgelegt.
- (6) Mit dieser Definition wird eine Unterscheidung von Punkten (den statischen Objekten) und den Verschiebungen (den dynamischen Objekten) in der Geometrie möglich.



### 22.1.3 Satz

Wird in  $\mathcal{P}$  irgendein Punkt  $O$  („Origin“) als Ursprung ausgezeichnet, so kommt dadurch eine bijektive Abbildung

$$\begin{cases} V & \rightarrow \mathcal{P} \\ v & \mapsto \tau_v(O) = O + v \end{cases} \quad \text{mit Umkehrabbildung} \quad \begin{cases} \mathcal{P} & \rightarrow V \\ P & \mapsto \overrightarrow{OP} \end{cases}$$

zustande. Die Räume sind „als Mengen gleich“. In diesem Zusammenhang heißt der Vektor  $\overrightarrow{OP}$  auch der *Ortsvektor* (bezüglich  $O$ ) für den Punkt  $P$ .

Jeder affine Raum erbt also in gewisser Weise Strukturen von seinem Vektorraum  $V$ . Der umgekehrte Sachverhalt kommt in dem folgenden Beispiel zum Ausdruck.

### 22.1.4 Beispiel

Jeder (reelle) Vektorraum  $V$  kann als affiner Raum  $\mathcal{P} = V$  (mit sich selbst als Translationenvektorraum aufgefasst werden. Die zugehörige Abbildung ist einfach durch

$$\tau \begin{cases} V \times V & \rightarrow V \\ (v, w) & \mapsto \tau_v(w) =: w + v. \end{cases}$$

gegeben. Der kanonische Vektorraum  $\mathbb{R}^n$  wird so zum kanonischen affinen Raum  $\mathcal{A}^n$ . Als Ursprung wird hier oft der Punkt  $\vec{0}$  gewählt.

Der wesentliche Unterschied von affinen Räumen gegenüber Vektorräumen besteht darin, dass in ihnen kein Nullpunkt von vornherein (auf natürliche Weise) ausgezeichnet ist.

Unsere „Anschauungsräume“ der Dimension 2 (Zeichenebene) oder der Dimension 3 (der Raum, in dem wir leben) werden geeigneter als Affine Räume denn als Vektorräume modelliert.

Erste Beobachtungen und Rechenregeln werden in dem folgenden Satz zusammengefasst:

### 22.1.5 Satz

(i) Die Translationsoperation hat die folgenden Eigenschaften:

$$\begin{aligned} \tau_{\vec{0}} &= \text{id}_{\mathcal{P}}. \\ \tau_v^{-1} &= \tau_{-v} \end{aligned}$$

(ii) Es seien  $P, Q, R \in \mathcal{P}$ . Für Verbindungsvektoren gilt:

$$\begin{aligned} \overrightarrow{PP} &= \vec{0} \\ \overrightarrow{PQ} + \overrightarrow{QR} &= \overrightarrow{PR} \\ \overrightarrow{QP} &= -\overrightarrow{PQ} \end{aligned}$$

(iii) Es seien  $P, Q \in \mathcal{P}$ . Für Ortsvektoren gilt:

$$\overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP}.$$

**22.1.6 Beweis** (i) Es sei  $P$  ein beliebiger Punkt von  $\mathcal{P}$ . Dann gilt:

$$\tau_{\vec{0}} \circ \tau_{\vec{0}}(P) \stackrel{(3)}{=} \tau_{\vec{0}+\vec{0}}(P) = \tau_{\vec{0}}(P).$$

Da  $\tau_{\vec{0}}$  gemäß Eigenschaft (1) bijektiv ist, können wir die Umkehrabbildung  $\tau_{\vec{0}}^{-1}$  auf diese Gleichung anwenden. Es folgt:

$$\tau_{\vec{0}}(P) = P.$$

Also ist  $\tau_{\vec{0}}$  die identische Abbildung von  $\mathcal{P}$ .

Die zweite Aussage ist jetzt ganz einfach zu zeigen: Es gilt nämlich

$$\tau_{-v} \circ \tau_v = \tau_{\vec{0}} = \text{id}_{\mathcal{P}} \quad \text{und} \quad \tau_v \circ \tau_{-v} = \tau_{\vec{0}} = \text{id}_{\mathcal{P}}.$$

und das heißt, dass  $\tau_{\vec{0}}$  die Umkehrabbildung von  $\tau_v$  ist.

(ii) Der Vektor  $\overrightarrow{PQ}$  ist gemäß (2) durch die Eigenschaft

$$\tau_{\overrightarrow{PQ}}(P) = Q$$

eindeutig charakterisiert.

Dann folgt die erste Behauptung mit

$$\tau_{\overrightarrow{P\vec{P}}}(P) = P = \tau_{\vec{0}}(P).$$

Die zweite Behauptung sieht man so:

$$\tau_{\overrightarrow{P\vec{Q}}+\overrightarrow{Q\vec{R}}}(P) = \tau_{\overrightarrow{Q\vec{R}}}(\tau_{\overrightarrow{P\vec{Q}}}(P)) = \tau_{\overrightarrow{Q\vec{R}}}(Q) = R = \tau_{\overrightarrow{P\vec{R}}}(R).$$

Die dritte Behauptung ergibt sich mit der ersten und zweiten so:

$$\overrightarrow{P\vec{Q}} + \overrightarrow{Q\vec{P}} = \overrightarrow{P\vec{P}} = \vec{0}.$$

(iii) ist klar.

## 22.2 Affine Unterräume

### 22.2.1 Definition

Es sei  $\mathcal{P}$  ein affiner Raum mit Translationsvektorraum  $V$ . Eine Teilmenge  $\mathcal{Q} \subseteq \mathcal{P}$  heißt affiner Unterraum von  $\mathcal{P}$  der Dimension  $m$ , wenn es einen Punkt  $Q \in \mathcal{Q}$  und einen Unterraum  $U \subseteq V$  mit  $\dim U = m$  gibt, so dass

$$\mathcal{Q} = Q + U := \{Q + u \mid u \in U\}$$

Der Unterraum  $U$  ist durch  $\mathcal{Q}$  eindeutig bestimmt, da dann

$$U = \{\overrightarrow{PQ} \mid P, Q \in \mathcal{Q}\}.$$

Ein affiner Raum heißt

- *Gerade*, wenn seine Dimension gleich 1 ist,
- *Ebene*, wenn seine Dimension gleich 2 ist,
- *Hyperebene*, wenn seine Dimension gleich  $n - 1$  ist.

Eine Menge  $\{P\}$ , die genau einen Punkt von  $\mathcal{P}$  enthält, ist ein 0-dimensionaler affiner Unterraum.

### 22.2.2 Definition: Affine Basis

Ist  $\mathcal{Q}$  ein affiner Unterraum eines affinen Raums  $\mathcal{P}$ , so wähle dazu

- einen Punkt  $P \in \mathcal{Q}$  und
- eine geordnete Basis  $B = (w_1, \dots, w_m)$  des zugeordneten Unterraums  $U$  von  $V$ .

Dann gibt es zu jedem Punkt  $Q \in \mathcal{Q}$  ein-eindeutig ein  $m$ -Tupel  $(\mu_1, \dots, \mu_m) \in \mathbb{K}^m$ , so dass

$$Q = P + \mu_1 w_1 + \dots + \mu_m w_m.$$

In diesem Zusammenhang wird das  $(m + 1)$ -Tupel  $(P, w_1, \dots, w_m)$  auch eine *affine Basis* (oder ein *affines Koordinatensystem*) für  $\mathcal{Q}$  genannt.

### 22.2.3 Definition: Parallelität

Zwei affine Unterräume  $\mathcal{Q}_1$  und  $\mathcal{Q}_2$  heißen *parallel (zueinander)*, wenn für die zugehörigen Unterräume  $U_1$  und  $U_2$  gilt:

$$U_1 \subseteq U_2 \quad \text{oder} \quad U_2 \subseteq U_1 \quad (*)$$

Symbolisch schreibt man für Parallelität bzw. Nicht-Parallelität:

$$\mathcal{Q}_1 \parallel \mathcal{Q}_2, \quad \mathcal{Q}_1 \not\parallel \mathcal{Q}_2.$$

Gilt  $\dim \mathcal{Q}_1 = \dim \mathcal{Q}_2$ , so ist die Bedingung  $(*)$  äquivalent zu

$$U_1 = U_2.$$

### 22.2.4 Definition: Windschief

Zwei affine Unterräume  $\mathcal{Q}_1$  und  $\mathcal{Q}_2$  heißen *windschief (zueinander)*, wenn

$$\mathcal{Q}_1 \cap \mathcal{Q}_2 = \emptyset \quad \text{und} \quad \mathcal{Q}_1 \not\parallel \mathcal{Q}_2.$$

### 22.2.5 Bemerkung

Für zwei affine Unterräume  $\mathcal{Q}_1$  und  $\mathcal{Q}_2$  mit  $\dim \mathcal{Q}_1 = \dim \mathcal{Q}_2$  kann man dann die folgenden Fälle unterscheiden:

- $\mathcal{Q}_1 = \mathcal{Q}_2$ ,
- $\mathcal{Q}_1 \neq \mathcal{Q}_2$  mit  $\mathcal{Q}_1 \cap \mathcal{Q}_2 \neq \emptyset$ .
- $\mathcal{Q}_1 \neq \mathcal{Q}_2$  mit  $\mathcal{Q}_1 \parallel \mathcal{Q}_2$ ,
- $\mathcal{Q}_1$  windschief zu  $\mathcal{Q}_2$ .

Betrachte als Beispiele Geraden und Ebenen im zwei- oder drei-dimensionalen Anschauungsraum. Welche Fälle können dann gar nicht auftreten?

### 22.2.6 Satz: LGS-Lösungsmengen als affine Unterräume

Wir können die Lösungsmengen von Linearen Gleichungssystemen geometrisch beschreiben. Es sei

$$Ax = b$$

ein lineares Gleichungssystem mit einer reellen  $m \times n$ -Matrix  $A$  und einem Vektor  $b \in \mathbb{R}^n$ .

Dann ist die Lösungsmenge  $\mathcal{L}(A|b)$  genau dann nicht-leer, wenn

$$\text{Rang}(A|b) = \text{Rang } A.$$

Ist die Lösungsmenge nicht-leer, so bildet sie einen affinen Unterraum der Dimension

$$\dim \mathcal{L}(A|b) = n - \text{Rang } A.$$

Die Gleichung  $Ax = b$  heißt in diesem Zusammenhang auch einfach die *Gleichung des affinen Unterraums*.

## 23 Euklidische affine Räume

### 23.1 Einführung

#### 23.1.1 Definition: Euklidischer affiner Raum

Es sei  $\mathcal{P}$  ein affiner Raum über  $\mathbb{R}$ .

Ist auf dem zugehörigen  $\mathbb{R}$ -Vektorraum  $V$  der Translationen ein Skalarprodukt  $\langle \cdot, \cdot \rangle$  gegeben, so heißt  $\mathcal{P}$  ein *euklidischer affiner Raum*.

Für zwei Punkte  $P, Q \in \mathcal{P}$  heißt die nicht-negative reelle Zahl

$$\overline{PQ} = d(P, Q) := \|\overrightarrow{PQ}\| = \sqrt{\langle \overrightarrow{PQ}, \overrightarrow{PQ} \rangle}$$

der *Abstand* zwischen  $P$  und  $Q$ .

#### 23.1.2 Beobachtung

Für alle  $P, Q, R \in \mathcal{P}$  gilt:

$$\begin{aligned} d(P, Q) &= 0 \iff P = Q \\ d(P, Q) &= d(Q, P) \\ d(P, R) &\leq d(P, Q) + d(Q, R). \end{aligned}$$

Die Funktion  $d : M \times M \rightarrow \mathbb{R}_0^+$  ist also eine Metrik auf  $\mathcal{P}$ , vgl. Definition 11.8.3.

## 23.2 Abstand zu einem Unterraum

### 23.2.1 Satz und Definition: Der Lotfußpunkt

Es sei  $\mathcal{Q}$  ein  $m$ -dimensionaler affiner Unterraum von  $\mathcal{P}$  und  $P \in \mathcal{P}$ . Die folgenden Aussagen über einen Punkt  $L \in \mathcal{Q}$  sind äquivalent.

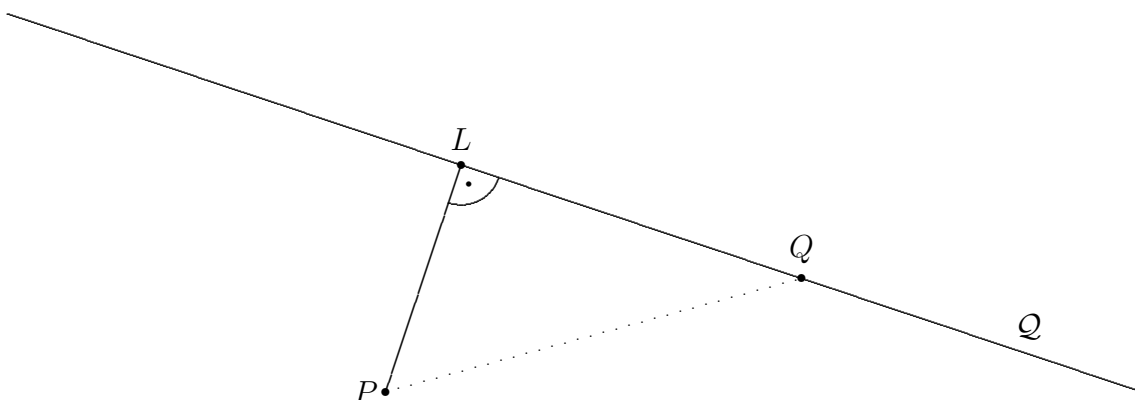
- (A) (Def) Der Punkt  $L$  heißt *Lotfußpunkt* (bzgl. des Lots von  $P$  auf  $\mathcal{Q}$ ).
- (B) Ist  $\{w_1, \dots, w_m\}$  eine Orthonormalbasis des Translationsvektorraums  $U$  von  $\mathcal{Q}$ , so gilt für alle  $Q \in \mathcal{Q}$

$$\overrightarrow{QL} = \langle \overrightarrow{QP}, w_1 \rangle w_1 + \dots + \langle \overrightarrow{QP}, w_m \rangle w_m,$$

d.h.  $\overrightarrow{QL}$  ist die „senkrechte Projektion des Vektors  $\overrightarrow{QP}$  auf  $\mathcal{Q}$ “.

- (C) Es gilt  $\langle \overrightarrow{PL}, \overrightarrow{LQ} \rangle = 0$  für alle  $Q \in \mathcal{Q}$ , d.h.  $\overrightarrow{PL}$  steht senkrecht auf  $\mathcal{Q}$ .
- (D) Es ist  $\|\overrightarrow{PL}\| < \|\overrightarrow{PQ}\|$  für alle  $Q \in \mathcal{Q} \setminus \{L\}$ , d.h.  $L \in \mathcal{Q}$  ist näher an  $P$  als alle anderen Punkte von  $\mathcal{Q}$ .
- (E) Es gilt:  $\|\overrightarrow{PL}\| \leq \|\overrightarrow{PQ}\|$  für alle  $Q \in \mathcal{Q}$ , d.h. bei  $L$  wird der Abstand der Punkte  $Q \in \mathcal{Q}$  zu  $P$  minimal.

Durch diese Aussagen ist der Punkt  $L$  eindeutig bestimmt.



### 23.2.2 Beweis

(B)  $\Rightarrow$  (C): Der Vektor  $\overrightarrow{LQ}$  hat die folgende Entwicklung bzgl. der ONB

$$\overrightarrow{LQ} = \langle \overrightarrow{LQ}, w_1 \rangle w_1 + \dots + \langle \overrightarrow{LQ}, w_m \rangle w_m,$$

was man leicht durch Skalarmultiplikation dieser Gleichung mit den Vektoren  $w_i$  der Orthonormalbasis nachweisen kann. Es gilt dann weiter:

$$\begin{aligned} \langle \overrightarrow{PL}, \overrightarrow{LQ} \rangle &= \langle \overrightarrow{PQ} + \overrightarrow{QL}, \overrightarrow{LQ} \rangle = \langle \overrightarrow{PQ}, \overrightarrow{LQ} \rangle + \langle \overrightarrow{QL}, \overrightarrow{LQ} \rangle \\ &\stackrel{(B)}{=} \langle \overrightarrow{PQ}, \overrightarrow{LQ} \rangle + \left\langle \langle \overrightarrow{QP}, w_1 \rangle w_1 + \dots + \langle \overrightarrow{QP}, w_m \rangle w_m, \overrightarrow{LQ} \right\rangle \\ &= \langle \overrightarrow{PQ}, \overrightarrow{LQ} \rangle + \langle \overrightarrow{QP}, w_1 \rangle \langle w_1, \overrightarrow{LQ} \rangle + \dots + \langle \overrightarrow{QP}, w_m \rangle \langle w_m, \overrightarrow{LQ} \rangle \\ &= \langle \overrightarrow{PQ}, \overrightarrow{LQ} \rangle - \langle \overrightarrow{PQ}, w_1 \rangle \langle w_1, \overrightarrow{LQ} \rangle - \dots - \langle \overrightarrow{PQ}, w_m \rangle \langle w_m, \overrightarrow{LQ} \rangle \\ &= \langle \overrightarrow{PQ}, \overrightarrow{LQ} - \langle w_1, \overrightarrow{LQ} \rangle w_1 - \dots - \langle w_m, \overrightarrow{LQ} \rangle w_m \rangle = \langle \overrightarrow{PQ}, \vec{0} \rangle = 0. \end{aligned}$$

(C)  $\Rightarrow$  (B): Der Vektor  $\overrightarrow{QL}$  ist in  $U$  enthalten, besitzt also eine Entwicklung bzgl. der in (B) vorgegebenen Orthonormalbasis:

$$\overrightarrow{QL} = \lambda_1 w_1 + \dots + \lambda_m w_m.$$

Die Koeffizienten ergeben sich zu

$$\lambda_i = \langle \overrightarrow{QL}, w_i \rangle = \langle \overrightarrow{QP}, w_i \rangle + \underbrace{\langle \overrightarrow{PL}, w_i \rangle}_{=0 \text{ wegen (C)}} = \langle \overrightarrow{QP}, w_i \rangle.$$

(C)  $\Rightarrow$  (D): Für  $Q \in \mathcal{Q} \setminus \{L\}$  gilt:

$$\begin{aligned} \|\overrightarrow{PQ}\|^2 - \|\overrightarrow{PL}\|^2 &= \|\overrightarrow{PL} + \overrightarrow{LQ}\|^2 - \|\overrightarrow{PL}\|^2 \\ &= \langle \overrightarrow{PL} + \overrightarrow{LQ}, \overrightarrow{PL} + \overrightarrow{LQ} \rangle - \langle \overrightarrow{PL}, \overrightarrow{PL} \rangle \\ &= 2\langle \overrightarrow{PL}, \overrightarrow{LQ} \rangle + \langle \overrightarrow{LQ}, \overrightarrow{LQ} \rangle = \|\overrightarrow{LQ}\|^2 > 0. \end{aligned}$$

(D)  $\Rightarrow$  (E): Die Aussage (D) impliziert, dass  $\|\overrightarrow{PL}\| \leq \|\overrightarrow{PQ}\|$  für alle  $Q \in \mathcal{Q}$ . Das ist aber gerade die Aussage (E).

(E)  $\Rightarrow$  (C): Für  $Q = L$  ist die Aussage (C) trivial. Sei also O.B.d.A.  $Q \in \mathcal{Q} \setminus \{L\}$  beliebig. Der Punkt  $R$ , definiert durch

$$\overrightarrow{PR} = \overrightarrow{PL} - \frac{\langle \overrightarrow{PL}, \overrightarrow{LQ} \rangle}{\langle \overrightarrow{LQ}, \overrightarrow{LQ} \rangle} \cdot \overrightarrow{LQ}$$

ist wegen  $\overrightarrow{LR} = \overrightarrow{PR} - \overrightarrow{PL} \in \mathcal{Q}$  ebenfalls in  $\mathcal{Q}$ . Daher gilt mit (E)

$$\begin{aligned} 0 &\leq \|\overrightarrow{PR}\|^2 - \|\overrightarrow{PL}\|^2 = \|\overrightarrow{PL} - \frac{\langle \overrightarrow{PL}, \overrightarrow{LQ} \rangle}{\langle \overrightarrow{LQ}, \overrightarrow{LQ} \rangle} \cdot \overrightarrow{LQ}\|^2 - \|\overrightarrow{PL}\|^2 \\ &= -2 \frac{\langle \overrightarrow{PL}, \overrightarrow{LQ} \rangle^2}{\langle \overrightarrow{LQ}, \overrightarrow{LQ} \rangle} + \frac{\langle \overrightarrow{PL}, \overrightarrow{LQ} \rangle^2}{\langle \overrightarrow{LQ}, \overrightarrow{LQ} \rangle^2} \cdot \langle \overrightarrow{LQ}, \overrightarrow{LQ} \rangle = -\frac{\langle \overrightarrow{PL}, \overrightarrow{LQ} \rangle^2}{\langle \overrightarrow{LQ}, \overrightarrow{LQ} \rangle} \leq 0. \end{aligned}$$

Daraus folgt aber  $\langle \overrightarrow{PL}, \overrightarrow{LQ} \rangle = 0$ .

Die Eindeutigkeit folgt direkt aus Aussage (B).

### 23.2.3 Definition: Abstand Punkt — Affiner Unterraum

Ist  $\mathcal{Q}$  ein affiner Unterraum von  $\mathcal{P}$  und  $P \in \mathcal{P}$  ein Punkt, so heißt der Abstand zwischen Punkt und zugehörigem Lotfußpunkt

$$d(P, \mathcal{Q}) := \inf\{d(P, Q) \mid Q \in \mathcal{Q}\} = d(P, L)$$

der *Abstand* zwischen dem Punkt  $P$  und dem affinen Unterraum  $\mathcal{Q}$ .

## 23.3 Hyperebenen und Normalenvektoren

### 23.3.1 Definition: Normaleneinheitsvektor

Es sei  $\mathcal{Q}$  eine Hyperebene in  $\mathcal{P}$  und  $U$  der zugehörige  $(n - 1)$ -dimensionale Unterraum von  $V$ . Eine Orthonormalbasis  $B = \{w_1, \dots, w_{n-1}\}$  von  $U$  lässt sich durch einen Vektor  $\vec{n}$  zu einer Orthonormalbasis von  $V$  ergänzen. Er ist bis auf einen Faktor  $\pm 1$  eindeutig bestimmt. Ein solcher Vektor  $\vec{n}$  hat also die Länge 1 und steht senkrecht auf den Verbindungsvektoren in  $\mathcal{Q}$ , er heißt deshalb *Normaleneinheitsvektor* für  $\mathcal{Q}$ .

### 23.3.2 Satz

Für den Abstand eines Punktes  $P \in \mathcal{P}$  von einer Hyperebene  $\mathcal{Q}$  gilt

$$d(P, \mathcal{Q}) = |\langle \vec{n}, \overrightarrow{PQ} \rangle|,$$

wobei  $Q \in \mathcal{Q}$  beliebig ist.

### 23.3.3 Beweis

Ist nämlich  $L \in \mathcal{Q}$  der zugehörige Lotfußpunkt, so stehen sowohl  $\vec{n}$  als auch  $\overrightarrow{PL}$  (gemäß obigem Satz) auf (allen Verbindungsvektoren innerhalb) der Hyperebene  $\mathcal{Q}$  senkrecht. Damit sind sie aber linear abhängig.

Mit der Cauchy-Schwarz-Ungleichung bei linearer Abhängigkeit, vgl. Satz 11.4.2 (iv), gilt dann

$$\begin{aligned} d(P, \mathcal{Q}) &= d(P, L) = \|\overrightarrow{PL}\| = \|\vec{n}\| \cdot \|\overrightarrow{PL}\| = |\langle \vec{n}, \overrightarrow{PL} \rangle| \\ &= |\langle \vec{n}, \overrightarrow{PL} \rangle + \langle \vec{n}, \overrightarrow{LQ} \rangle| = |\langle \vec{n}, \overrightarrow{PQ} \rangle|. \end{aligned}$$

### 23.3.4 Hessesche Normalform

Ist  $P \in \mathcal{Q}$  ein fest gewählter Punkt in einer Hyperebene, so gilt für alle  $Q \in \mathcal{Q}$

$$\langle \vec{n}, \overrightarrow{OQ} - \overrightarrow{OP} \rangle = 0.$$

Führt man eine ONB in dem zugehörigen Translationsvektorraum  $U$  ein und ergänzt sie durch  $\vec{n}$  zu einer ONB von  $V$ , so kann diese Gleichung auch in Koordinatenform aufgeschrieben werden.

Diese Gleichung heißt dann die *Hessesche Normalform* der Hyperebene  $\mathcal{Q}$ .



## 23.4 Isometrien

### 23.4.1 Definition

Eine Abbildung  $f : \mathcal{P} \rightarrow \mathcal{P}$  in einem euklidischen affinen Raum heißt eine *Isometrie* oder *Kongruenzabbildung*, wenn

$$d(f(P), f(Q)) = d(P, Q) \quad \text{für alle } P, Q \in \mathcal{P}.$$

### 23.4.2 Satz: Isometrien

- (i) Translationen sind Isometrien.
- (ii) Es sei  $D \in \mathcal{P}$ , die beiden Abbildungen  $f : \mathcal{P} \rightarrow \mathcal{P}$  und  $f^\# : V \rightarrow V$  seien durch die Beziehung

$$f^\#(\overrightarrow{DP}) = \overrightarrow{Df(P)} \quad \text{für alle } P \in \mathcal{P}$$

wechselseitig definiert.

Genau dann ist  $f$  eine Isometrie mit Fixpunkt  $D$ , wenn  $f^\#$  orthogonal ist.

- (iii) Ist  $f$  eine Isometrie, so gibt es einen Vektor  $v$  und einen Punkt  $D \in \mathcal{P}$ , so dass

$$f = g \circ \tau_v,$$

wobei  $g$  eine Isometrie mit  $D$  als Fixpunkt ist.

### 23.4.3 Beweis

- (i) Es ist

$$\begin{aligned} d(\tau_v(P), \tau_v(Q)) &= \|\overrightarrow{\tau_v(P)\tau_v(Q)}\| = \|\overrightarrow{\tau_v(P)P} + \overrightarrow{PQ} + \overrightarrow{\tau_v(Q)Q}\| \\ &= \|-v + \overrightarrow{PQ} + v\| = \|\overrightarrow{PQ}\| = d(P, Q). \end{aligned}$$

- (ii) Es sei  $f$  eine Isometrie mit Fixpunkt  $D$ . Dann gilt für beliebiges  $v \in V$ ,  $P := \tau_v(D)$ :

$$\|f^\#(v)\| = \|f^\#(\overrightarrow{DP})\| = \|\overrightarrow{Df(P)}\| = \|\overrightarrow{f(D)f(P)}\| = \|\overrightarrow{DP}\| = \|v\|,$$

daraus folgt mit dem Satz über die Eigenschaften orthogonaler Abbildungen (Satz 20.1.1), dass  $f^\#$  orthogonal ist.

Ist umgekehrt eine Abbildung  $f^\#$  mit der im Satz angegebenen Eigenschaft gegeben, so gilt:

$$\overrightarrow{Df(D)} = f^\#(\overrightarrow{DD}) = f^\#(\vec{0}) = \vec{0} \implies f(D) = D.$$

Außerdem ist

$$\begin{aligned} d(f(P), f(Q)) &= \|\overrightarrow{f(P)f(Q)}\| = \|\overrightarrow{f(P)D} + \overrightarrow{Df(Q)}\| = \|\overrightarrow{-f^\#(\overrightarrow{DP}) + f^\#(\overrightarrow{DQ})}\| \\ &= \|f^\#(\overrightarrow{DQ} - \overrightarrow{DP})\| = \|\overrightarrow{DQ} - \overrightarrow{DP}\| = \|\overrightarrow{PQ}\| = d(P, Q). \end{aligned}$$

- (iii) Wähle als Vektor

$$v := \overrightarrow{f^{-1}(D)D}.$$

Dann gilt:

$$f \circ \tau_{-v}(D) = f \circ f^{-1}(D) = D.$$

Daraus folgt, dass  $g := f \circ \tau_{-v}$  eine Isometrie mit Fixpunkt  $D$  ist.

## 23.5 Anwendungen in der Elementargeometrie

Es sei  $\mathcal{P}$  ein euklidischer affiner Raum, beispielsweise die klassische „Zeichenebene“ der Schulgeometrie.

Im folgenden seien  $A, B, C \in \mathcal{P}$ . Man stelle sich ein Dreieck mit den Eckpunkten  $A, B, C$  vor.

### 23.5.1 Satz des Pythagoras

Es gilt die Äquivalenz

$$\langle \overrightarrow{AC}, \overrightarrow{BC} \rangle = 0 \iff \overline{AC}^2 + \overline{BC}^2 = \overline{AB}^2$$

**23.5.2 Beweis** Aufgrund der Bilinearität des Skalarprodukts ist

$$\begin{aligned} \overline{AC}^2 + \overline{BC}^2 - \overline{AB}^2 &= \langle \overrightarrow{AC}, \overrightarrow{AC} \rangle + \langle \overrightarrow{BC}, \overrightarrow{BC} \rangle - \langle \overrightarrow{AB}, \overrightarrow{AB} \rangle \\ &= \langle \overrightarrow{AC}, \overrightarrow{AC} \rangle + \langle \overrightarrow{BC}, \overrightarrow{BC} \rangle - \langle \overrightarrow{AC} - \overrightarrow{BC}, \overrightarrow{AC} - \overrightarrow{BC} \rangle \\ &= 2\langle \overrightarrow{AC}, \overrightarrow{BC} \rangle \end{aligned}$$

Damit zeigt sich, dass der Satz von Pythagoras schon in der Definition des euklidischen Skalarprodukts enthalten ist.

Die eigentliche Bedeutung des klassischen Satzes von Pythagoras liegt stärker in der Feststellung, dass die uns umgebende Welt durch das mathematische Modell des affinen euklidischen Raums gut beschrieben ist.

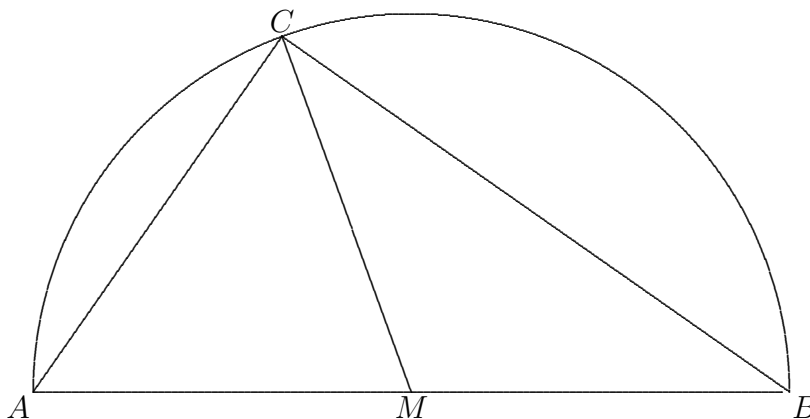
### 23.5.3 Satz von Thales

Es sei  $M := A + \frac{1}{2}\overrightarrow{AB}$  der Mittelpunkt der Strecke  $[AB]$ .

Es gilt die Äquivalenz

$$\overline{MC} = \overline{MA} \iff \langle \overrightarrow{AC}, \overrightarrow{BC} \rangle = 0$$

Geometrisch: Der Innenwinkel bei  $C$  ist genau dann ein rechter Winkel, wenn  $C$  auf dem Kreis liegt, der die Strecke  $[AB]$  als Durchmesser hat.



### 23.5.4 Beweis

(1) Aus der Definition von  $M$  folgt die (aus der Anschauung völlig klare) Tatsache, dass

$$B = A + \overrightarrow{AB} = M + \frac{1}{2}\overrightarrow{AB}$$

und deswegen

$$\overrightarrow{MB} = \frac{1}{2}\overrightarrow{AB} = \overrightarrow{AM} = -\overrightarrow{MA}.$$

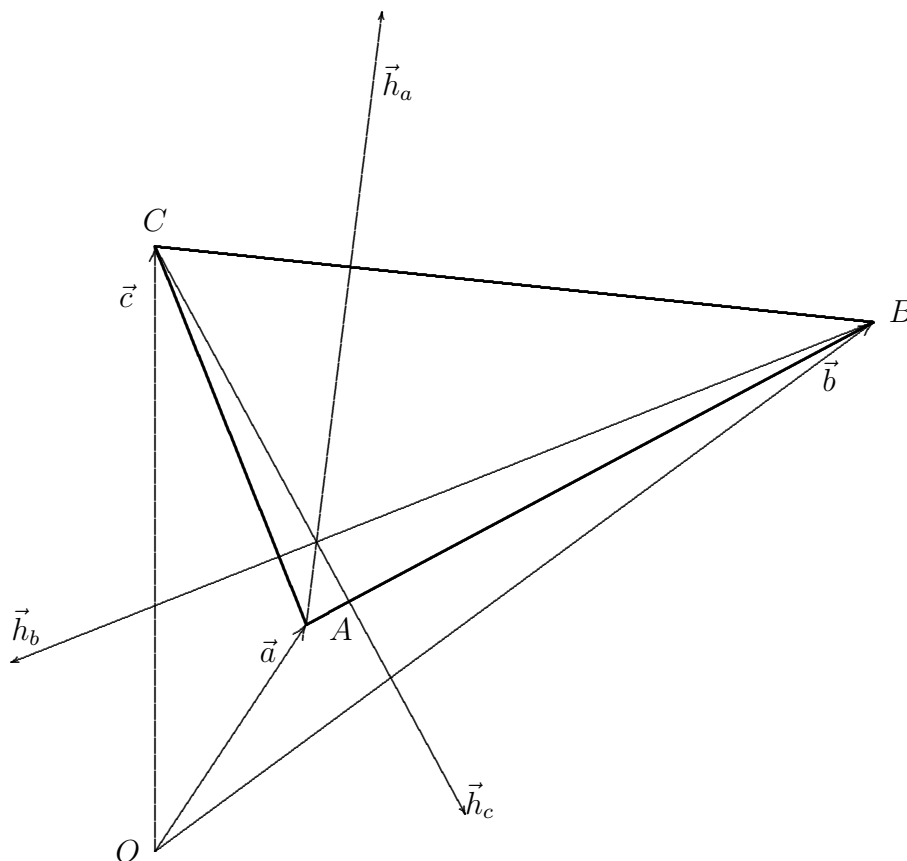
(2) Die Äquivalenz geht nun aus der folgenden Rechnung hervor:

$$\begin{aligned} |MC|^2 - |MA|^2 &= \langle \overrightarrow{MC}, \overrightarrow{MC} \rangle - \langle \overrightarrow{MA}, \overrightarrow{MA} \rangle \\ &= \langle \overrightarrow{MC} - \overrightarrow{MA}, \overrightarrow{MC} + \overrightarrow{MA} \rangle \\ &= \langle \overrightarrow{MC} - \overrightarrow{MA}, \overrightarrow{MC} - \overrightarrow{MB} \rangle \\ &= \langle \overrightarrow{AC}, \overrightarrow{BC} \rangle. \end{aligned}$$

### 23.5.5 Satz: Schnittpunkt der Höhen in einem Dreieck

Die Höhen eines Dreiecks  $ABC$  schneiden sich in einem Punkt.

### 23.5.6 Beweis



(1) Wir wählen irgendeinen Punkt  $O$  als Ursprung und bezeichnen die Ortsvektoren von  $A, B, C$  mit  $\vec{a}, \vec{b}, \vec{c}$ .

Weiter bezeichnen wir mit  $\vec{\ell}_a, \vec{\ell}_b, \vec{\ell}_c \in V \setminus \{\vec{0}\}$  Vektoren, die jeweils senkrecht (lotrecht) auf den Seiten des Dreiecks stehen, d.h. die Eigenschaft

$$\langle \vec{\ell}_a, \overrightarrow{BC} \rangle = \langle \vec{\ell}_b, \overrightarrow{CA} \rangle = \langle \vec{\ell}_c, \overrightarrow{AB} \rangle = 0 \quad (*)$$

haben.

(2) Wir berechnen den Ortsvektor  $\vec{h}_{ab}$  des Schnittpunkts der Höhen auf  $a$  und  $b$  durch den Ansatz

$$\vec{h}_{ab} = \vec{a} + \mu_a \vec{\ell}_a = \vec{b} + \mu_b \vec{\ell}_b.$$

Um  $\mu_b$  zu bestimmen, bilden wir das Skalarprodukt mit  $\overrightarrow{BC}$  und beachten dabei (\*):

$$\langle \vec{a}, \overrightarrow{BC} \rangle = \langle \vec{b}, \overrightarrow{BC} \rangle + \mu_b \langle \vec{\ell}_b, \overrightarrow{BC} \rangle$$

Daraus folgt

$$\mu_b = \frac{\langle \vec{a} - \vec{b}, \overrightarrow{BC} \rangle}{\langle \vec{\ell}_b, \overrightarrow{BC} \rangle} = \frac{\langle \overrightarrow{BA}, \overrightarrow{BC} \rangle}{\langle \vec{\ell}_b, \overrightarrow{BC} \rangle}$$

und weiter

$$\vec{h}_{ab} = \vec{b} + \frac{\langle \overrightarrow{BA}, \overrightarrow{BC} \rangle}{\langle \vec{\ell}_b, \overrightarrow{BC} \rangle} \vec{\ell}_b.$$

(3) Die genau gleiche Prozedur — mit den Rollen von  $C$  und  $A$  vertauscht — liefert den Ortsvektor  $\vec{h}_{cb}$  des Höhenschnittpunkts:

$$\vec{h}_{cb} = \vec{b} + \frac{\langle \overrightarrow{BC}, \overrightarrow{BA} \rangle}{\langle \vec{\ell}_b, \overrightarrow{BA} \rangle} \vec{\ell}_b.$$

(4) Weiter gilt mit (\*):

$$\langle \vec{\ell}_b, \overrightarrow{BC} \rangle = \langle \vec{\ell}_b, \overrightarrow{BA} \rangle + \langle \vec{\ell}_b, \overrightarrow{AC} \rangle = \langle \vec{\ell}_b, \overrightarrow{BA} \rangle$$

und damit  $\vec{h}_{ab} = \vec{h}_{cb}$ .

### 23.5.7 Satz: Schnittpunkt der Mittelsenkrechten in einem Dreieck

Die Mittelsenkrechten eines Dreiecks  $ABC$  schneiden sich in einem Punkt.

#### 23.5.8 Beweis

(1) Wir wählen irgendeinen Punkt  $O$  als Ursprung und bezeichnen die Ortsvektoren von  $A, B, C$  mit  $\vec{a}, \vec{b}, \vec{c}$ .

Weiter bezeichnen wir mit  $\vec{\ell}_a, \vec{\ell}_b, \vec{\ell}_c \in V \setminus \{\vec{0}\}$  Vektoren, die jeweils senkrecht (lotrecht) auf den Seiten des Dreiecks stehen, d.h. die Eigenschaft

$$\langle \vec{\ell}_a, \overrightarrow{BC} \rangle = \langle \vec{\ell}_b, \overrightarrow{CA} \rangle = \langle \vec{\ell}_c, \overrightarrow{AB} \rangle = 0 \quad (*)$$

haben.

(2) Wir berechnen den Ortsvektor  $\vec{m}_{ab}$  des Schnittpunkts der Mittelsenkrechten auf  $a$  und  $b$  durch den folgenden Ansatz:

$$\vec{m}_{ab} = \frac{\vec{b} + \vec{c}}{2} + \mu_a \vec{\ell}_a = \frac{\vec{a} + \vec{c}}{2} + \mu_b \vec{\ell}_b.$$

Die Lote werden jetzt also von dem Mittelpunkt der Seiten aus gefällt. Wir bilden wieder das Skalarprodukt mit  $\overrightarrow{BC}$  und beachten dabei wieder (\*):

$$\langle \frac{\vec{b} + \vec{c}}{2}, \overrightarrow{BC} \rangle = \langle \frac{\vec{a} + \vec{c}}{2}, \overrightarrow{BC} \rangle + \mu_b \langle \vec{\ell}_b, \overrightarrow{BC} \rangle$$

Daraus folgt:

$$\mu_b = \frac{1}{2} \cdot \frac{\langle \overrightarrow{AB}, \overrightarrow{BC} \rangle}{\langle \vec{\ell}_b, \overrightarrow{BC} \rangle}$$

und weiter

$$\vec{m}_{ab} = \frac{\vec{a} + \vec{c}}{2} + \frac{1}{2} \cdot \frac{\langle \overrightarrow{AB}, \overrightarrow{BC} \rangle}{\langle \vec{\ell}_b, \overrightarrow{BC} \rangle} \vec{\ell}_b.$$

(3) Die genau gleiche Prozedur — mit den Rollen von  $C$  und  $A$  vertauscht — liefert den Ortsvektor  $\vec{m}_{cb}$  des Schnittpunkts der Mittelsenkrechten

$$\vec{m}_{cb} = \frac{\vec{a} + \vec{c}}{2} + \frac{1}{2} \cdot \frac{\langle \overrightarrow{CB}, \overrightarrow{BA} \rangle}{\langle \vec{\ell}_b, \overrightarrow{BA} \rangle} \vec{\ell}_b.$$

(4) Weiter gilt mit (\*):

$$\langle \vec{\ell}_b, \overrightarrow{BC} \rangle = \langle \vec{\ell}_b, \overrightarrow{BA} \rangle + \langle \vec{\ell}_b, \overrightarrow{AC} \rangle = \langle \vec{\ell}_b, \overrightarrow{BA} \rangle$$

und damit  $\vec{m}_{ab} = \vec{m}_{cb}$ .

## 24 Nachtrag: Orientierung

In diesem Abschnitt ist der zugrundeliegende Körper der der reellen Zahlen:  $\mathbb{K} = \mathbb{R}$ .

## 24.1 Orientierung eines $\mathbb{R}$ -Vektorraums

### 24.1.1 Einstieg

Wir betrachten zu einem gegebenen  $\mathbb{R}$ -Vektorraum  $V$  der Dimension  $n \geq 1$  die Menge aller geordneten Basen

$$G = \{(v_1, \dots, v_n) \in V^n \mid \{v_1, \dots, v_n\} \text{ ist Basis}\}.$$

Die folgenden Aussagen über zwei geordnete Basen  $B = (v_1, \dots, v_n)$  und  $\tilde{B} = (\tilde{v}_1, \dots, \tilde{v}_n)$  sind äquivalent.

(A) (Def)  $B$  und  $\tilde{B}$  heißen *gleich-orientiert*.

(B) Der durch die Festlegung

$$f : \begin{cases} V & \rightarrow V \\ v_i & \mapsto \tilde{v}_i \quad \text{für alle } i \in \{1, \dots, n\} \end{cases}$$

definierte Automorphismus hat positive Determinante.

(C) Für die zugehörige Übergangsmatrix gilt

$$\det \Theta_{\tilde{B}B} > 0.$$

Es ist leicht zu sehen, dass diese Relation zwischen zwei geordneten Basen eine Äquivalenzrelation auf  $G$  ist. Es zerfällt dann  $G$  in zwei Äquivalenzklassen

$$G = \mathcal{B}' \dot{\cup} \mathcal{B}''.$$

Zwei geordnete Basen heißen *gleich-orientiert*, wenn sie zur gleichen Klasse gehören, anderenfalls *verschieden-orientiert*.

### 24.1.2 Definition: Orientierter Vektorraum

Wird in einem  $\mathbb{R}$ -Vektorraum eine der beiden Äquivalenzklassen ausgezeichnet, so spricht man von einem *orientierten  $\mathbb{R}$ -Vektorraum*, symbolisch wird er als Paar  $(V, \mathcal{B})$  notiert.

Eine geordnete Basis  $B$  aus  $\mathcal{B}$  heißt dann *positiv orientiert*.

### 24.1.3 Bemerkungen

1. Zu einem gegebenen  $\mathbb{R}$ -Vektorraum  $V$  gibt es also zwei orientierte  $\mathbb{R}$ -Vektorräume.
2. Für zwei geordnete Basen, die die gleichen Vektoren enthalten, gilt

$$\begin{aligned} (v_1, \dots, v_n) &\sim (\tilde{v}_{\pi(1)}, \dots, \tilde{v}_{\pi(n)}) \\ \iff \pi &\text{ ist eine gerade Permutation.} \end{aligned}$$

3. Die Vertauschung zweier Vektoren in einer geordneten Basis führt zu einem Wechsel der Orientierung.
4. Die Ersetzung eines Vektors in einer geordneten Basis durch seinen (additiv) inversen führt zu einem Wechsel der Orientierung.
5. Ein reeller affiner Raum  $\mathcal{P}$  wird zu einem orientierten affinen Raum, wenn der zugeordnete Vektorraum  $V$  der Translationen orientiert wird.

#### 24.1.4 Direkte Summe

Man kann die direkte Summe zweier orientierter Vektorräume  $(V, \mathcal{B})$  und  $(\tilde{V}, \tilde{\mathcal{B}})$  dadurch orientieren, dass man die Äquivalenzklasse der geordneten Basis von  $V \oplus \tilde{V}$  auswählt, die durch das Nebeneinanderschreiben von zwei geordneten Basen aus  $\mathcal{B}$  bzw.  $\tilde{\mathcal{B}}$  entsteht.

$$\mathcal{B}_{\oplus} := [(B, \tilde{B})], \quad \text{wobei } B, \tilde{B} \text{ positiv orientiert.}$$

Man schreibt dann

$$(V, \mathcal{B}) \oplus (\tilde{V}, \tilde{\mathcal{B}}) = (V \oplus \tilde{V}, \mathcal{B}_{\oplus}).$$

Aus dieser Gleichung ist auch ersichtlich, dass die Orientierung zweier dieser Räume die Orientierung des dritten nach sich zieht.

Genau dann, wenn beide Vektorräume  $V$  und  $\tilde{V}$  ungerade Dimension haben, sind die beiden orientierten Vektorräume  $(V, \mathcal{B}) \oplus (\tilde{V}, \tilde{\mathcal{B}})$  und  $(\tilde{V}, \tilde{\mathcal{B}}) \oplus (V, \mathcal{B})$  verschieden orientiert.

#### 24.1.5 Orientierungserhaltender Isomorphismus

Ein Isomorphismus  $f : (V, \mathcal{B}) \rightarrow (\tilde{V}, \tilde{\mathcal{B}})$  zwischen zwei orientierten Vektorräumen heißt *orientierungserhaltend*, wenn das Bild einer positiv orientierten Basis von  $V$  wieder eine positiv orientierte Basis von  $\tilde{V}$  ist.

Im anderen Fall heißt  $f$  *orientierungsumkehrend*.

#### 24.1.6 Orientierungserhaltender Automorphismus

Für einen Automorphismus  $f : (V, \mathcal{B}) \rightarrow (V, \mathcal{B})$  in einem orientierten Vektorraum sind die drei folgenden Aussagen äquivalent:

- (A)  $f$  ist orientierungserhaltend
- (B)  $f$  ist bei Auswahl der anderen Orientierung orientierungserhaltend
- (C) Es ist  $\det f > 0$ .

Die Eigenschaft „orientierungserhaltend“ eines Automorphismus kann also gefasst werden, ohne dass der Vektorraum  $V$  orientiert sein muss.

## 24.2 Orientierung des $\mathbb{R}^n$

### 24.2.1 Definition

Wählt man für den  $\mathbb{R}^n$  als Äquivalenzklasse von geordneten Basen diejenige aus, die die kanonische Basis enthält,

$$\mathcal{B} = [(e_1, \dots, e_n)],$$

so heißt  $(\mathbb{R}^n, \mathcal{B})$  *kanonisch orientiert*.

**24.2.2 Beobachtung** Eine geordnete Basis  $(v_1, \dots, v_n)$  von  $\mathbb{R}^n$  ist also genau dann positiv orientiert, wenn die zugehörige Matrix der Spaltenvektoren positive Determinante hat:

$$\det(v_1, \dots, v_n) > 0.$$

Diese Zahl heißt auch das „*Volumen* des von  $(v_1, \dots, v_n)$  aufgespannten Parallelepipeds“.

### 24.2.3 Orientierung des $\mathbb{R}^3$

Sind zwei Vektoren  $v_1, v_2$  im kanonisch orientierten  $\mathbb{R}^3$  linear unabhängig, so ist die Basis

$$(v_1, v_2, v_1 \times v_2)$$

positiv orientiert. Die Definition des Vektorprodukts ist gerade so gemacht.

### 24.2.4 Orientierung von $\mathbb{C}$

Fasst man  $\mathbb{C}$  als 2-dimensionalen  $\mathbb{R}$ -Vektorraum auf, so kann er — kanonisch — orientiert werden durch Auswahl der geordneten Basis  $(1, i)$ .



## 24.3 Orientierung eines euklidischen Raums

### 24.3.1 Existenz einer positiv orientierten ONB

Es sei  $V$  ein euklidischer  $n$ -dimensionaler Raum, der zusätzlich mit einer Orientierung versehen sei.

Ist  $B' = (v'_1, v'_2, \dots, v'_n)$  eine Orthonormalbasis, die nicht positiv orientiert ist, so kann sie durch Vertauschung der ersten beiden Vektoren in eine solche umgewandelt werden.

In jedem orientierten euklidischen Raum existiert also eine positiv orientierte ONB.

### 24.3.2 Die Gruppe $\mathbb{S}$

Wir definieren die kommutative (und damit additiv geschriebene) Gruppe  $(\mathbb{S}, +)$  als Quotientengruppe der Gruppe  $(\mathbb{R}, +)$  nach der diskreten Untergruppe

$$2\pi\mathbb{Z} = \{x \in \mathbb{R} \mid \exists k \in \mathbb{Z} : x = k \cdot 2\pi\}.$$

Also

$$\mathbb{S} := \mathbb{R}/(2\pi\mathbb{Z}).$$

Die Elemente von  $\mathbb{S}$  lassen sich eineindeutig durch reelle Zahlen in  $[0, 2\pi[$  repräsentieren. Wenn die Addition oder Subtraktion zweier Elemente einen Wert außerhalb von  $[0, 2\pi[$  produziert, so muss er durch Addition von geeignet vielen Vielfachen von  $2\pi$  wieder nach  $[0, 2\pi[$  „zurückgeholt“ werden. Man rechnet „modulo  $2\pi$ “.

Man mache sich klar, dass ein Winkel  $\varphi \in \mathbb{S}$  durch den Vektor  $\begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} \in \mathbb{R}^2$  eindeutig bestimmt ist.

### 24.3.3 Satz und Definition: Drehung

Es sei  $V$  ein orientierter euklidischer Raum der Dimension 2.

Die folgenden Aussagen über einen Endomorphismus  $f : V \rightarrow V$  stimmen überein.

- (A)  $f$  heißt *Drehung*.
- (B)  $f$  ist orientierungserhaltend und orthogonal. Vgl. auch Abschnitt 24.1.6.
- (C) Es gibt einen „Drehwinkel“  $\varphi \in \mathbb{S}$ , so dass bzgl. jeder positiv orientierten ONB  $B = (w_1, w_2)$  gilt:

$$\begin{aligned} f(w_1) &= \cos \varphi w_1 + \sin \varphi w_2 \\ f(w_2) &= -\sin \varphi w_1 + \cos \varphi w_2, \end{aligned}$$

oder — äquivalent dazu —

$$f_{BB} = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}.$$

Beachte, dass diese Matrix — und damit der Drehwinkel — bereits durch einen einzigen Vektor ungleich Nullvektor und Bildvektor festgelegt ist.

### 24.3.4 Beweis

Er ist im wesentlichen mit Satz 20.2.3 erbracht.

### 24.3.5 Bemerkungen

(1) Man könnte in der Aussage (C) das „Minuszeichen beim Sinus“ jeweils diagonal in die andere Ecke verschieben. Dann ist der Drehwinkel ebenfalls eindeutig bestimmt, aber gleich dem negativen des durch die ursprüngliche Aussage bestimmten.

(2) Bemerkenswert an der Aussage (C) ist, dass sie für beliebige positiv orientierte ONBen gilt.

(3) Würde man in der Aussage (C) des Satzes alle geordneten Basen — und nicht nur die positiv orientierten — zulassen, so würde der Satz falsch sein. Man könnte dann gleichzeitig den Drehwinkel  $\varphi$  durch  $-\varphi$  und eine geordnete Basis  $(w_1, w_2)$  durch  $(w_2, w_1)$  ersetzen. Um die Eindeutigkeit des Drehwinkels zu erzwingen, muss der Vektorraum  $V$  orientiert sein.

### 24.3.6 Satz und Definition: Winkelmaß

Es sei  $V$  ein orientierter euklidischer Raum der Dimension 2. Dann gibt es genau eine Abbildung — das *Winkelmaß* —

$$\mu : \begin{cases} (V \setminus \{\vec{0}\})^2 & \rightarrow \mathbb{S} \\ (v, \tilde{v}) & \mapsto \mu(v, \tilde{v}), \end{cases}$$

so dass die folgenden Aussagen erfüllt sind:

- Längeninvarianz: Sind  $\lambda, \tilde{\lambda} > 0$ , so gilt

$$\mu(\lambda v, \tilde{\lambda} \tilde{v}) = \mu(v, \tilde{v}).$$

- Additivität: Für drei Vektoren  $v_1, v_2, v_3 \in V \setminus \{\vec{0}\}$  gilt:

$$\mu(v_1, v_2) + \mu(v_2, v_3) = \mu(v_1, v_3).$$

- Invarianz unter Drehungen: Ist  $D$  eine Drehung, so gilt für zwei Vektoren  $v, \tilde{v} \in V \setminus \{\vec{0}\}$

$$\mu(D(v), D(\tilde{v})) = \mu(v, \tilde{v}).$$

- Normierung: Für jede positive orientierte ONB  $(w_1, w_2)$  gilt  $\mu(w_1, w_2) = +\frac{\pi}{2}$ .

- Stetigkeit: Die Abbildung  $\mu$  ist stetig. Das heißt, wenn

$$\|v - v'\| \rightarrow 0 \quad \text{und} \quad \|\tilde{v} - \tilde{v}'\| \rightarrow 0, \quad \text{dann} \quad \mu(v, \tilde{v}) - \mu(v', \tilde{v}') \rightarrow 0.$$

### 24.3.7 Beweis Existenz

Wir konstruieren die Abbildung für normierte Vektoren  $v, \tilde{v}$ . Mit Hilfe der Längeninvarianz kann sie dann für alle Vektoren ungleich Nullvektor fortgesetzt werden. Seien also  $v, \tilde{v} \in V$  mit  $\|v\| = \|\tilde{v}\| = 1$ .

(1) Bzgl. einer positiv orientierten ONB  $B$  seien

$$\eta_B(v) = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad \text{und} \quad \eta_B(\tilde{v}) = \begin{pmatrix} \gamma \\ \delta \end{pmatrix}.$$

Dann ist die Gleichung

$$\begin{pmatrix} \gamma \\ \delta \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

äquivalent zu

$$\begin{pmatrix} \gamma \\ \delta \end{pmatrix} = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}$$

und zu

$$\begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} \begin{pmatrix} \gamma \\ \delta \end{pmatrix}.$$

Es gibt also ein eindeutig bestimmtes  $\varphi \in \mathbb{S}$ , so dass  $\tilde{v} = D_\varphi(v)$ , wobei  $D_\varphi$  die in Satz ?? beschriebene Drehung mit Drehwinkel  $\varphi$  ist.

(2) Durch

$$\mu(v, \tilde{v}) := \varphi, \quad \text{wobei } \tilde{v} = D_\varphi(v).$$

ist also das Winkelmaß wohldefiniert.

(3) Sind drei normierte Vektoren  $v_1, v_2, v_3$  mit

$$v_2 = D_\varphi(v_1) \quad \text{und} \quad v_3 = D_\psi(v_2)$$

und damit

$$v_3 = D_\psi(v_2) = D_\psi(D_\varphi(v_1)) = D_{\varphi+\psi}(v_1)$$

gegeben, so gilt

$$\mu(v_1, v_2) + \mu(v_2, v_3) = \varphi + \psi = \mu(v_1, v_3).$$

(4) Die Drehinvarianz ist offensichtlich.

(5) Ist  $(w_1, w_2)$  eine positiv orientierte ONB, so gilt mit Satz ??

$$D_{\frac{\pi}{2}}(w_1) = \cos \frac{\pi}{2} w_1 + \sin \frac{\pi}{2} w_2 = w_2$$

und damit  $\mu(w_1, w_2) = \frac{\pi}{2}$ .

(6) Den Überlegungen in Schritt (1) kann die Stetigkeit der Abbildung

$$(v, \tilde{v}) \mapsto (\alpha, \beta, \gamma, \delta) \mapsto (\cos \varphi, \sin \varphi) \mapsto \varphi$$

entnommen werden.

### 24.3.8 Beweis Eindeutigkeit

Es sei  $\mu$  die eben konstruierte Abbildung und  $\tilde{\mu}$  eine weitere Abbildung mit den im Satz genannten Eigenschaften. Wir müssen zeigen, dass  $\mu(v, \tilde{v}) = \mu'(v, \tilde{v})$  für alle normierten  $v, \tilde{v} \in V$ .

(1) Ist  $\mu(v, \tilde{v}) = \frac{\pi}{2n}$  für ein  $n \in \mathbb{N}$ , so gilt  $Dv = \tilde{v}$ , wenn  $D = D_{\frac{\pi}{2n}}$  die Drehung um diesen Winkel ist. Weiter ist dann

$$\begin{aligned} n \cdot \mu'(v, \tilde{v}) &= \mu'(v, \tilde{v}) + \mu'(Dv, D\tilde{v}) + \dots + \mu'(D^{n-1}v, D^{n-1}\tilde{v}) \\ &= \mu'(v, Dv) + \mu'(Dv, D^2v) + \dots + \mu'(D^{n-1}v, D^n v) \\ &= \mu'(v, D^n v) = \mu'(v, D_{\frac{\pi}{2}} v) = \mu'(v, w_2) = \frac{\pi}{2} \\ &= n \cdot \mu(v, \tilde{v}), \end{aligned}$$

also

$$\mu'(v, \tilde{v}) = \mu(v, \tilde{v}).$$

(2) Es sei nun  $\mu(v, \tilde{v}) = \frac{m\pi}{2n}$  für  $m, n \in \mathbb{N}$ . Wir definieren den Vektor  $w \in V$  durch  $w = Dv$ , wobei  $D = D_{\frac{\pi}{2n}}$  ist. Aufgrund von (1) ist dann

$$\mu(v, w) = \mu'(v, w)$$

und weiter

$$\begin{aligned} \frac{m\pi}{2n} &= m \cdot \mu(v, w) = m \cdot \mu'(v, w) \\ &= \mu'(v, w) + \mu'(Dv, Dw) + \dots + \mu'(D^{m-1}v, D^{m-1}w) \\ &= \mu'(v, Dv) + \mu'(Dv, D^2v) + \dots + \mu'(D^{m-1}v, D^m v) \\ &= \mu'(v, D^m v) = \mu'(v, \tilde{v}) \end{aligned}$$

also

$$\mu'(v, \tilde{v}) = \mu(v, \tilde{v}).$$

(3) Es sei schließlich  $\mu(v, \tilde{v}) = a \cdot \frac{\pi}{2}$  mit  $a \in \mathbb{R} \setminus \mathbb{Q}$ .

Da  $\mu(v, \cdot)$  surjektiv ist, gibt es zu jedem  $\varepsilon > 0$  einen normierten Vektor  $w$  mit

$$\begin{aligned} |\mu(\tilde{v}, w)| &\leq \frac{\varepsilon}{2} \\ |\mu'(\tilde{v}, w)| &\leq \frac{\varepsilon}{2} \end{aligned}$$

und

$$|\mu(v, w)| = \frac{m}{n} \frac{\pi}{2}, \quad \text{wobei } m, n \in \mathbb{N}.$$

Es ist dann  $\mu'(v, w) = \mu(v, w)$  und deshalb

$$\begin{aligned} |\mu(v, \tilde{v}) - \mu'(v, \tilde{v})| &= |[\mu(v, \tilde{v}) - \mu(v, w)] - [\mu'(v, \tilde{v}) - \mu'(v, w)]| \\ &= |\mu(w, \tilde{v}) - \mu'(w, \tilde{v})| \\ &\leq |\mu(w, \tilde{v})| + |\mu'(w, \tilde{v})| \leq \varepsilon. \end{aligned}$$

Da diese Ungleichung für alle  $\varepsilon > 0$  gilt, muss

$$\mu'(v, \tilde{v}) = \mu(v, \tilde{v}).$$

gelten.

## 24.4 Orientierung der zweidimensionalen „Zeichenebene“

Als Modell für die aus der Schule bekannte Zeichenebene wählen wir einen (bis auf Isomorphie: den) orientierten euklidischen Raum der Dimension 2.

Die Zeichenebene wird durch Festlegung einer geordneten Basis

$$(v_1, v_2)$$

orientiert, die aus zwei senkrecht zueinander stehenden Vektoren besteht, so dass der zweite Vektor aus dem ersten durch Drehung um  $90^\circ$  **gegen** den Uhrzeigersinn hervorgeht.

Die in der Schule und im Alltag häufig anzutreffende Sprechweise, dass es sich bei „gegen den Uhrzeigersinn“ um den „mathematisch positiven Drehsinn“ handelt, ist höchst fragwürdig, da diese „Situation aus der Wahrnehmung der Wirklichkeit“ eben nicht innermathematisch erfasst werden kann.

**24.4.1 Achsenspiegelungen** in der Zeichenebene sind orientierungsumkehrend.

## 24.5 Orientierung des dreidimensionalen „Anschauungsraums“

### 24.5.1 Einstieg

Rene Descartes hat entdeckt, dass der uns umgebende Raum (bzgl. vieler Aspekte unserer Wahrnehmung) durch den dreidimensionalen (affinen) euklidischen Raum modelliert werden kann.

Eine Orientierung des Anschauungsraums sei durch eine der folgenden gleich-orientierten geordneten Basen vorgenommen:

- (Daumen, Zeigefinger, Mittelfinger) der vom Herzen abgewandten Hand
- (Blick nach vorne, Blick nach oben, Blick seitlich in Richtung „herzabgewandt“)
- (Minus Bewegung eines Elektrons, Richtung des Magnetfelds, Lorentzkraft)
- (Erdachse Nordpol  $\rightarrow$  Südpol, Radius Mittelpunkt  $\rightarrow$  Eichstädt, Bewegung von Eichstädt aufgrund Erddrehung)

Man nennt diese Orientierung die *Rechts-Orientierung*.

### 24.5.2 Beispiel: Schrauben

Weltweit werden Schrauben aller Art (technisch, Wasserhähne, Korkenzieher, ...) so hergestellt, dass die durch sie definierte geordnete Basis

(Kopf  $\rightarrow$  Spitze, radial nach außen, Bewegung eines Punktes auf dem Umfang beim Hinein- bzw. Zudrehen)

rechts-orientiert ist.

### 24.5.3 Beispiel: Spielwürfel

Untersuchen Sie einen Spielwürfel daraufhin, ob er bei Zuordnung einer geordneten Basis mittels

(Mitte  $\rightarrow$  1-Seitenfläche, Mitte  $\rightarrow$  2-Seitenfläche, Mitte  $\rightarrow$  3-Seitenfläche)

rechtsorientiert ist.

**24.5.4 Ebenenspiegelungen** im dreidimensionalen Raum sind orientierungsumkehrend.