

# Aktuelle Version

## Mathematik II für Elektrotechnik, TUD SS 2009

### Inhaltsverzeichnis

<b>22</b>	<b>Lineare Abbildungen, Matrizen, Transformationen, Determinanten</b>	<b>1</b>
22.1	Lineare Abbildungen und Matrizen . . . . .	1
22.1.1	Definition und Beispiele . . . . .	1
22.1.2	Koordinaten . . . . .	1
22.1.3	Beschreibung . . . . .	1
22.1.4	Komposition und Matrixmultiplikation . . . . .	2
22.1.5	Isomorphie und inverse Matrix . . . . .	3
22.1.6	Kern und Bild . . . . .	3
22.1.7	Charakterisierung und Berechnung der Inversen . . . . .	4
22.1.8	Summe und Vielfache . . . . .	5
22.2	Transponierte Matrix . . . . .	5
22.2.1	Definition und Eigenschaften . . . . .	5
22.2.2	Orthogonale Matrizen und Abbildungen . . . . .	5
22.2.3	Beispiele . . . . .	6
22.3	Basistransformation . . . . .	7
22.3.1	Koordinatentransformation für Vektoren . . . . .	7
22.3.2	Orthogonale Transformationsmatrizen . . . . .	8
22.3.3	Transformation der Abbildungsbeschreibung . . . . .	8
22.4	Determinanten . . . . .	9
22.4.1	Regeln . . . . .	9
22.4.2	Eindeutigkeit und Berechnung . . . . .	10
22.4.3	Produktsatz. . . . .	11
22.4.4	Transponieren und Zeilenumformungen . . . . .	11
22.4.5	Entwicklung . . . . .	12
22.4.6	Adjugierte Matrix . . . . .	13
22.4.7	Existenz . . . . .	13
<b>23</b>	<b>Transformation von Maßen und Integralen</b>	<b>14</b>
23.1	Unabhängigkeit des Jordan-Maßes . . . . .	14
23.2	Vorbemerkung . . . . .	14
23.2.1	Eulers Dreh . . . . .	14
23.2.2	Beweis . . . . .	15
23.3	Affine Abbildungen . . . . .	15
23.3.1	Affine Abbildungen und ihre Determinanten . . . . .	15
23.3.2	Bewegungen . . . . .	16
23.4	Maßtransformation bei affinen Abbildungen . . . . .	16
23.4.1	Bewegungsinvarianz des Jordan-Maßes . . . . .	16
23.4.2	Volumen eines Spats . . . . .	16
23.4.3	Affine Maßtransformation . . . . .	17

<b>27</b>	<b>Eigenwerte und Jordan-Normalform</b>	<b>18</b>
27.1	Eigenwerte und Eigenvektoren . . . . .	18
27.1.1	Motivation: Schwingungsgleichung . . . . .	18
27.1.2	Basistransformation . . . . .	19
27.1.3	Definition . . . . .	19
27.1.4	Beispiele . . . . .	20
27.1.5	Ansatz . . . . .	20
27.1.6	Beispiel . . . . .	20
27.1.7	Gekoppelte Schwingung . . . . .	20
27.2	Charakteristische Gleichung . . . . .	21
27.2.1	Algebraische Bestimmung der Eigenwerte . . . . .	21
27.2.2	Beispiel . . . . .	21
27.2.3	Mehr Beispiele . . . . .	22
27.2.4	Tricks . . . . .	23
27.3	Eigenräume und ihre Basen . . . . .	23
27.3.1	Vielfachheiten . . . . .	23
27.3.2	Unabhängigkeit von Eigenvektoren . . . . .	24
27.3.3	Diagonalisierbarkeit . . . . .	24
27.3.4	Potenzen und Inverse . . . . .	25
27.4	Komplexe Eigenwerte und Eigenvektoren . . . . .	25
27.4.1	Zerlegung des charakteristischen Polynoms . . . . .	25
27.4.2	Komplexe Eigenwerte reeller Matrizen . . . . .	26
27.5	Satz von Cayley-Hamilton . . . . .	27
27.6	Lineare Abbildungen in Ebene und Raum . . . . .	27
27.6.1	Klassifikation ebener linearer Abbildungen . . . . .	27
27.6.2	Schwingungsgleichung . . . . .	29
27.6.3	Lineare Abbildungen im Raum . . . . .	29
27.7	Jordan-Normalform . . . . .	31
27.7.1	Jordan-Ketten . . . . .	31
27.7.2	Jordan-Blöcke . . . . .	32
27.7.3	Jordan-Basen . . . . .	32
27.7.4	Jordan-Matrizen . . . . .	33
27.7.5	Hauptsatz . . . . .	33
27.7.6	Tricks . . . . .	34
27.7.7	Beispiel: Jordan-Basis bei nur einem Eigenwert . . . . .	34
27.7.8	Beispiel: Verallgemeinerte Eigenräume . . . . .	34
27.7.9	Hauptvektoren und verallgemeinerte Eigenräume . . . . .	36
27.7.10	Warnung . . . . .	37
27.7.11	Ausnahmefälle . . . . .	37
<b>28</b>	<b>Quadratische Formen und Hauptachsen</b>	<b>38</b>
28.1	Motivation . . . . .	38
28.1.1	Extrema von Funktionen zweier Veränderlicher . . . . .	38
28.1.2	Flächen zweiter Ordnung . . . . .	39
28.1.3	Metrischer Tensor . . . . .	39
28.1.4	Trägheitstensor + . . . . .	40

28.2	Bilineare und quadratische Formen . . . . .	40
28.2.1	Bilinearformen . . . . .	41
28.2.2	Quadratische Formen . . . . .	41
28.2.3	Transformation . . . . .	41
28.2.4	Definitheit . . . . .	42
28.3	Hauptachsentransformation . . . . .	42
28.3.1	Hauptsatz und Korollare . . . . .	42
28.3.2	Beweis . . . . .	43
28.3.3	Ausartung . . . . .	44
28.3.4	Klassifikation für $n = 2, 3$ . . . . .	44
28.3.5	Geometrisches Beispiel . . . . .	45
28.4	Definitheit . . . . .	45
28.4.1	Trägheitssatz . . . . .	45
28.4.2	Zerlegung . . . . .	46
28.4.3	Symmetrischer Gauss . . . . .	46
28.5	Hauptminorenkriterium . . . . .	47
28.6	Verwechslung . . . . .	48
28.7	Quadriken . . . . .	49
28.7.1	Quadratische Gleichungen . . . . .	49
28.7.2	Transformation . . . . .	49
28.7.3	Rezept . . . . .	50
28.7.4	Beispiel . . . . .	51
28.7.5	Normalform . . . . .	52
28.7.6	Mittelpunkte + . . . . .	52
28.8	Zur Numerik der symmetrischer Matrizen + . . . . .	52
28.8.1	Cholesky-Zerlegung . . . . .	52
28.8.2	Jacobiverfahren . . . . .	53
28.8.3	Matrixphilosophie . . . . .	54
28.9	Schneller Beweis der Trägheit . . . . .	55

## 22 Lineare Abbildungen, Matrizen, Transformationen, Determinanten

### 22.1 Lineare Abbildungen und Matrizen

#### 22.1.1 Definition und Beispiele

Sei  $K$  ein Körper und  $V, W$   $K$ -Vektorräume. Eine Abbildung  $\phi : V \rightarrow W$  ist *linear*, falls

$$\phi(\vec{x} + \vec{y}) = \phi(\vec{x}) + \phi(\vec{y}) \quad \text{und} \quad \phi(\lambda\vec{x}) = \lambda\phi(\vec{x}) \quad \text{für alle } \vec{x}, \vec{y} \in V, \lambda \in K$$

Beispiele:  $A \in K^{m \times n}$  eine Matrix

$$\phi : K^n \rightarrow K^m, \quad \phi(\mathbf{x}) = A\mathbf{x}$$

$\text{Pol}_n$  der Raum der reellen Polynome von Grad  $\leq n$

$$\frac{\partial}{\partial x} : \text{Pol}_n \rightarrow \text{Pol}_{n-1}, \quad p(x) \mapsto \frac{\partial}{\partial x}(p(x))$$

$$\int : \text{Pol}_n \rightarrow \text{Pol}_{n+1}, \quad p(x) \mapsto \int_0^x p(t) dt$$

Weitere Beispiele s. Kap.11.

#### 22.1.2 Koordinaten

**Korollar 22.1** Ist  $\alpha : \vec{v}_1, \dots, \vec{v}_m$  Basis des  $K$ -Vektorraumes  $V$ , so kann man jedem Vektor  $\vec{x} \in V$  umkehrbar eindeutig seine Koordinatenspalte  $\vec{x}^\alpha \in K^m$  zuordnen

$$\vec{x} = \sum x_i \vec{v}_i \mapsto \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = \vec{x}^\alpha = \begin{pmatrix} x_1^\alpha \\ \vdots \\ x_m^\alpha \end{pmatrix}$$

Es gilt

$$(\vec{x} + \vec{y})^\alpha = \vec{x}^\alpha + \vec{y}^\alpha \quad \text{und} \quad (r\vec{x})^\alpha = r\vec{x}^\alpha$$

$$\vec{v}_j^\alpha = \mathbf{e}_j$$

wobei  $\mathbf{e}_j$  die *Einheitsspalte* mit 1 in der  $j$ -ten Zeile, 0 sonst, bezeichnet. Insbesondere ist die Zuordnung  $\vec{x} \mapsto \vec{x}^\alpha$  eine lineare Abbildung  $V \rightarrow K^n$ . Beweis: vgl. Kap.5.5.

#### 22.1.3 Beschreibung

**Satz 22.2** Sei  $\alpha : \vec{v}_1, \dots, \vec{v}_n$  eine Basis von  $V$  und  $\beta : \vec{w}_1, \dots, \vec{w}_m$  eine Basis von  $W$ . Dann gibt es eine 1-1-Entsprechung zwischen linearen Abbildungen  $\phi : V \rightarrow W$  und Matrizen  $A \in K^{m \times n}$  gegeben durch

$$(i) \quad \phi(\vec{v}_j)^\beta = A\mathbf{e}_j \quad \text{für } j = 1, \dots, n$$

bzw. äquivalent

$$(ii) \quad \phi(\vec{x})^\beta = A\vec{x}^\alpha \quad \text{für alle } \vec{x} \in V$$

Die  $j$ -te Spalte von  $A$  enthält die Koordinaten des Bildes des Basisvektors  $\vec{v}_j$

Beweis. Für alle  $\vec{x} \in V$  gilt

$$\sum_j x_j A \mathbf{e}_j = A \sum_j x_j \mathbf{e}_j = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A \vec{x}^\alpha$$

Sei  $\phi$  linear. Dann

$$(iii) \quad \phi\left(\sum_j x_j \vec{v}_j\right) = \sum_j x_j \phi(\vec{v}_j), \quad \left(\phi\left(\sum_j x_j \vec{v}_j\right)\right)^\beta = \sum_j x_j \phi(\vec{v}_j)^\beta$$

Also folgt (ii) aus (i) und  $\phi$  ist durch die  $\phi(\vec{v}_j)$  eindeutig bestimmt. Gilt (ii) und definiert man  $\phi$  durch (iii), so ist  $\phi$  linear:

$$\phi(\vec{x} + \vec{y})^\alpha = A(\vec{x} + \vec{y})^\alpha = A(\vec{x}^\alpha + \vec{y}^\alpha) = A\vec{x}^\alpha + A\vec{y}^\alpha = \phi(\vec{x})^\alpha + \phi(\vec{y})^\alpha$$

und daher  $\phi(\vec{x} + \vec{y}) = \phi(\vec{x}) + \phi(\vec{y})$ . Entsprechend für  $\lambda\vec{x}$ .

**Korollar 22.3** Zu  $\vec{x}_1, \dots, \vec{x}_n \in W$  gibt es genau eine lineare Abbildung  $\phi : V \rightarrow W$  so, dass  $\phi(\vec{v}_j) = \vec{x}_j$  für  $j = 1, \dots, n$  - nämlich mit der Matrix  $A$ , deren Spalten die Koordinaten der  $\vec{x}_j$  sind.

Ist  $\phi : V \rightarrow V$  linear und  $\alpha$  Basis von  $V$ , so ist  $A$  die Matrix von  $\phi$  (bzgl.  $\alpha$ , wenn  $\phi(\vec{x})^\alpha = A\vec{x}^\alpha$  für alle  $\vec{x} \in V$ ).

#### 22.1.4 Komposition und Matrixmultiplikation

Sind  $\phi : V \rightarrow W$  und  $\psi : W \rightarrow U$  Abbildungen, so ist die *Komposition* oder *Verkettung* definiert durch

$$\psi \circ \phi : V \rightarrow U, \quad (\psi \circ \phi)(\vec{x}) = \psi(\phi(\vec{x}))$$

Mit  $\chi : U \rightarrow X$  gilt das Assoziativgesetz

$$\chi \circ (\psi \circ \phi) = (\chi \circ \psi) \circ \phi$$

**Lemma 22.4** Seien  $\phi : V \rightarrow W$  und  $\psi : W \rightarrow U$  linear. Dann ist auch  $\psi \circ \phi : V \rightarrow U$  linear. Ist  $\alpha : \vec{v}_1, \dots, \vec{v}_n$  Basis von  $V$ ,  $\beta$  Basis von  $W$  und  $\gamma$  Basis von  $U$ , und  $A \in K^{m \times n}$  mit Spalten  $\mathbf{a}_1, \dots, \mathbf{a}_n$  die Matrix von  $\phi$  und  $B \in K^{m \times k}$  die Matrix von  $\psi$ , so hat  $\psi \circ \phi$  die Matrix  $BA$  mit Spalten  $B\mathbf{a}_1, \dots, B\mathbf{a}_n$

$$BA = (B\mathbf{a}_1 \quad \dots \quad B\mathbf{a}_n) = \left( \sum_{i=1}^m b_{hi} a_{ij} \right)_{h=1, \dots, k, j=1, \dots, n} \in K^{k \times n}$$

Beweis.  $\psi(\phi(\vec{x} + \vec{y})) = \psi(\phi(\vec{x}) + \phi(\vec{y})) = \psi(\phi(\vec{x})) + \psi(\phi(\vec{y}))$ . Ebenso für  $\lambda\vec{x}$ . Also  $\psi \circ \phi$  linear. Nun  $(\psi(\phi(\vec{v}_j)))^\gamma = B(\phi(\vec{v}_j))^\beta = B(A\mathbf{e}_j) = B\mathbf{a}_j$ .  $\square$

**Korollar 22.5** Die Matrizenmultiplikation ist, soweit ausführbar, assoziativ

$$C(BA) = (CB)A \quad \text{für } C \in K^{l \times k}, B \in K^{k \times m}, A \in K^{m \times n}$$

## 22.1.5 Isomorphie und inverse Matrix

Die Abbildungen  $\phi : V \rightarrow W$  und  $\psi : W \rightarrow V$  sind *Umkehrabbildungen* voneinander, wenn gilt

$$\psi(\phi(\vec{x})) = \vec{x} \text{ für alle } \vec{x} \in V, \quad \phi(\psi(\vec{y})) = \vec{y} \text{ für alle } \vec{y} \in W$$

oder anders ausgedrückt

$$\vec{y} = \phi(\vec{x}) \Leftrightarrow \vec{x} = \psi(\vec{y})$$

Dann ist  $\psi$  durch  $\phi$  eindeutig bestimmt und man schreibt  $\psi = \phi^{-1}$ .

**Lemma 22.6** Sei  $\phi : V \rightarrow W$  linear und besitze die Umkehrabbildung  $\phi^{-1} : W \rightarrow V$ . Dann ist  $\phi^{-1} : W \rightarrow V$  linear,  $\vec{v}_1, \dots, \vec{v}_n$  Basis von  $V$  genau dann, wenn  $\phi(\vec{v}_1), \dots, \phi(\vec{v}_n)$  Basis von  $W$  ist und es gilt  $\dim W = \dim V$ .

In diesem Falle heisst  $\phi : V \rightarrow W$  ein *Isomorphismus* und die Vektorräume  $V$  und  $W$  sind zueinander *isomorph*. Ein Beispiel ist der *Koordinatenisomorphismus*  $\vec{x} \mapsto \vec{x}^\alpha$ .

Beweis. Sind  $\vec{y}, \vec{y}' \in W$ , so  $\vec{y} = \phi(\vec{x})$  und  $\vec{y}' = \phi(\vec{x}')$  mit  $\vec{x} = \phi^{-1}(\vec{y}) \in V$  und  $\vec{x}' = \phi^{-1}(\vec{y}') \in V$ . Es folgt  $\psi(\vec{y} + \vec{y}') = \psi(\phi(\vec{x}) + \phi(\vec{x}')) = \psi(\phi(\vec{x} + \vec{x}')) = \vec{x} + \vec{x}' = \psi(\vec{y}) + \psi(\vec{y}')$ . Entsprechend für  $\lambda\vec{y}$ . Wegen  $\phi(\sum_j x_j \vec{v}_i) = \sum_j x_j \phi(\vec{v}_j)$  ist  $\vec{v}_1, \dots, \vec{v}_n$  in  $V$  erzeugend genau dann, wenn  $\phi(\vec{v}_1), \dots, \phi(\vec{v}_n)$  in  $W$  erzeugend ist und es folgt die entsprechende Aussage über Basen und damit über Dimension.  $\square$

Matrizen  $A, B \in K^{n \times n}$  sind *zueinander invers* wenn gilt (mit der Einheitsmatrix  $E \in K^{n \times n}$ )

$$AB = BA = E$$

**Korollar 22.7** Seien  $V$  und  $W$   $K$ -Vektorräume der Dimension  $n$  mit Basen  $\alpha$  bzw.  $\beta$  und  $\phi : V \rightarrow W$  und  $\psi : W \rightarrow V$  linear mit zugehörigen Matrizen  $A$  bzw.  $B$ . Dann sind  $\phi$  und  $\psi$  Umkehrabbildungen voneinander genau dann, wenn die Matrizen  $A$  und  $B$  zueinander invers sind.

**Korollar 22.8** Hat die Matrix  $A \in K^{n \times n}$  eine inverse, so ist diese eindeutig bestimmt und wird mit  $A^{-1}$  bezeichnet.

## 22.1.6 Kern und Bild

Seien  $V, W$   $K$ -Vektorräume mit Basen  $\alpha$  bzw.  $\beta$ . Sei  $\phi : V \rightarrow W$  linear mit Matrix  $A$ . Wir definieren

$$\text{Kern}(\phi) = \{\vec{x} \in V \mid \phi(\vec{x}) = \vec{0}\} = \{\vec{x} \in V \mid A\vec{x}^\alpha = \mathbf{0}\} \subseteq V$$

$$\text{Bild}(\phi) = \{\phi(\vec{x}) \mid \vec{x} \in V\} = \{\vec{y} \in W \mid \vec{y}^\beta \in \text{Spann}(A)\} \subseteq W$$

Das sind offenbar Untervektorräume und es gilt

$$\dim \text{Kern}(\phi) = \dim\{\mathbf{x} \in K^n \mid A\mathbf{x} = \mathbf{0}\}$$

$$\dim \text{Bild}(\phi) = \dim \text{Spann}(A) = \text{Rang}(A)$$

und somit die *Dimensionsformel*

$$\dim \text{Kern}(\phi) + \dim \text{Bild}(\phi) = \dim V$$

## 22.1.7 Charakterisierung und Berechnung der Inversen

**Satz 22.9** Seien  $V$  und  $W$   $K$ -Vektorräume der endlichen Dimension  $n = \dim V = \dim W$  mit Basen  $\alpha$  bzw.  $\beta$  und  $\phi : V \rightarrow W$  linear mit zugehöriger Matrix  $A$ . Dann sind die folgenden Aussagen äquivalent

- (1)  $\phi : V \rightarrow W$  ist ein Isomorphismus
- (2)  $\text{Kern}(\phi) = \{\vec{0}\}$
- (3)  $\text{Bild}(\phi) = W$
- (4)  $\text{Rang}(A) = n$
- (5) Es gibt Matrix  $Y \in K^{n \times n}$  mit  $YA = E$
- (6) Es gibt Matrix  $X \in K^{n \times n}$  mit  $AX = E$

Die Matrizen  $X, Y$  in (5) bzw. (6) sind eindeutig bestimmt als die Matrix der Umkehrabbildung  $\phi^{-1}$  also

$$X = Y = A^{-1}$$

Beweis. Nach Kor.22.7 folgen (5) und (6) aus (1). Mit Lemma 22.4 folgt (3) aus (6) und (2) aus (5). Wegen  $\dim W = n$  sind (3) und (4) äquivalent und nach dem Dimensionsformel (2) und (3). Seien also (2) und (3) angenommen. Wir definieren

$$\psi(\vec{y}) = \vec{x} \Leftrightarrow \phi(\vec{x}) = \vec{y}$$

Wegen (3) ist  $\psi$  auf ganz  $W$  definiert. Wegen (2) ist es auch wohldefiniert: ist  $\phi(\vec{x}) = \vec{y} = \phi(\vec{x}')$  so  $\phi(\vec{x} - \vec{x}') = \phi(\vec{x}) - \phi(\vec{x}') = \vec{0}$  also  $\vec{x} = \vec{x}'$ .  $\square$ . Da  $B^{-1}A^{-1}AB = B^{-1}EB = B^{-1}B = E$  folgt

**Socke-Schuh-Regel:**  $(AB)^{-1} = B^{-1}A^{-1}$

**Korollar 22.10** Gauß-Jordan, Ist  $A$  invertierbar, so kann man die Spalten  $\mathbf{x}_j$  von  $X = A^{-1}$  als Lösungen der Gleichungssysteme  $A\mathbf{x}_j = \mathbf{0}$  bestimmen. Dazu überführt man  $A$  durch Zeilenumformungen in die Einheitsmatrix. Führt man beginnend mit der Matrix  $E$  statt  $A$  dieselben Umformungen aus, so erhält man  $A^{-1}$

$$\begin{aligned} & \left( \begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 2 & 2 & 4 & 0 & 1 & 0 \\ 2 & 1 & 1 & 0 & 0 & 1 \end{array} \right) \rightsquigarrow \left( \begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & -2 & 1 & 0 \\ 0 & -1 & -1 & -2 & 0 & 1 \end{array} \right) \rightsquigarrow \left( \begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 0 & 1 \\ 0 & 0 & 2 & -2 & 1 & 0 \end{array} \right) \\ & \rightsquigarrow \left( \begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 0 & 1 \\ 0 & 0 & 1 & -1 & \frac{1}{2} & 0 \end{array} \right) \rightsquigarrow \left( \begin{array}{ccc|ccc} 1 & 1 & 0 & 2 & -\frac{1}{2} & 0 \\ 0 & -1 & 0 & -3 & \frac{1}{2} & 1 \\ 0 & 0 & 1 & -1 & \frac{1}{2} & 0 \end{array} \right) \\ & \rightsquigarrow \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 0 & 1 \\ 0 & -1 & 0 & -3 & \frac{1}{2} & 1 \\ 0 & 0 & 1 & -1 & \frac{1}{2} & 0 \end{array} \right) \rightsquigarrow \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 0 & 1 \\ 0 & -1 & 0 & 3 & -\frac{1}{2} & -1 \\ 0 & 0 & -1 & -1 & \frac{1}{2} & 0 \end{array} \right) \end{aligned}$$

## 22.1.8 Summe und Vielfache

Für Matrizen  $A, B \in K^{m \times n}$  und  $\lambda$  definiert man (komponentenweise)

$$A + B = (a_{ij} + b_{ij}) \in K^{m \times n}, \quad \lambda A = (\lambda a_{ij}) \in K^{m \times n}$$

Mit dieser Addition und Multiplikation mit Skalaren erhält man einen zu  $K^{n^2}$  isomorphen  $K$ -Vektorraum.

Sind  $\phi, \psi$  lineare Abbildungen von  $V$  in  $W$  so hat man auch die linearen Abbildungen

$$\phi + \psi : V \rightarrow W \quad (\phi + \psi)(\vec{x}) = \phi(\vec{x}) + \psi(\vec{x}), \quad \lambda\phi : V \rightarrow W \quad (\lambda\phi)(\vec{x}) = \lambda(\phi(\vec{x}))$$

Bzgl. gegebener Basen von  $V$  und  $W$  entsprechen diesen die Matrizen  $A + B$  bzw.  $\lambda A$ .

## 22.2 Transponierte Matrix

## 22.2.1 Definition und Eigenschaften

Zu einer Matrix  $A = (a_{ij})_{i=1, \dots, m, j=1, \dots, n} \in K^{m \times n}$  definieren wir die *transponierte* als

$$A^t = A^T = (b_{kl})_{k=1, \dots, n, l=1, \dots, m} \in K^{n \times m} \quad \text{mit } b_{kl} = a_{lk}$$

Es gilt (wie man leicht nachrechnet)

$$(A + B)^t = A^t + B^t, \quad (\lambda A)^t = \lambda A^t, \quad A^{tt} = A$$

$$\boxed{(AB)^t = B^t A^t}$$

$$A \text{ invertierbar} \Leftrightarrow A^t \text{ invertierbar, und dann } A^{-t} := (A^t)^{-1} = (A^{-1})^t$$

## 22.2.2 Orthogonale Matrizen und Abbildungen

Eine reelle Matrix  $A$  heisst

- *orthogonal*, wenn  $A^t A = E$ , d.h.  $A^t = A^{-1}$ .

**Lemma 22.11** Sei  $V$  euklidischer Vektorraum mit Orthonormalbasis  $\alpha : \vec{v}_1, \dots, \vec{v}_n$  und  $\phi : V \rightarrow V$  linear mit Matrix  $A$  bzgl.  $\alpha$ . Dann sind die folgenden Aussagen äquivalent

- (i)  $\|\phi(\vec{x})\| = \|\vec{x}\|$  für alle  $\vec{x} \in V$
- (ii)  $\langle \phi(\vec{x}) | \phi(\vec{y}) \rangle = \langle \vec{x} | \vec{y} \rangle$  für alle  $\vec{x}, \vec{y} \in V$
- (iii)  $\phi(\vec{v}_1), \dots, \phi(\vec{v}_n)$  ist Orthonormalbasis von  $V$
- (iv)  $A$  ist orthogonal

Dann heisst  $\phi$  auch eine *orthogonale Abbildung*. Beweis. (ii) folgt aus (i), weil man das Skalarprodukt aus der Länge definieren kann. Gilt (ii), so ist das Bild der ON-Basis auch ON, also unabhängig und wegen  $\dim V = n$  auch Basis. Aus (iii) folgt (iv), da dies die Orthogonalität der Spalten von  $A$  bedeutet. Gilt (iv) so  $\|\vec{x}\| = \|\vec{x}^\alpha\| = \|\phi(\vec{x})^\alpha\| = \|\phi(\vec{x})\|$ .  $\square$

## 22.2.3 Beispiele

Wir geben im Folgenden Beispiele linearer Abbildungen  $\phi_0 : V \rightarrow V'$  an, wobei  $V, V'$  Vektor-Raum bzw. Ebene der anschaulichen Geometrie sind. Sind  $\mathcal{P}, \mathcal{P}'$  die zugehörigen Punkträume und wählt man Ursprünge  $O$  bzw.  $O'$ , so kann man der linearen Abbildung  $\phi_0$  eine (affine) Abbildung zuordnen

$$\phi : \mathcal{P} \rightarrow \mathcal{P}'. \quad \phi(\vec{x} + O) = \phi_0(\vec{x}) + O'$$

Diese ist der Anschauung meist besser zugänglich und wird dementsprechend illustrativ benannt. In Beispielen und Aufgaben wird die gemeinte lineare Abbildung  $\phi_0$  häufig auf diese Weise (d.h. durch Hinweise auf das zugehörige  $\phi$ ) mitgeteilt und der Leser ist dann aufgefordert, die notwendige Präzisierung selbst zu leisten. In der Regel führt Identifikation von Punkten und Vektoren nach der Zauberformel "Ortsvektor" zu dem vom Aufgabensteller gewünschten Ergebnis. Eine präzise Behandlung der Zusammenhänge unterbleibt. Ebenso die Rechtfertigung von Bezeichnungen wie Drehung, Spiegelung usw. sowie der Nachweis der Linearität der zugehörigen vektoriiellen Abbildungen. Die Begriffe "Eigenraum"  $E_\lambda$  und "Eigenwert" (EV) werden später behandelt.

1. Identische Abbildung  $\phi = \text{id}_{\mathcal{P}}$  mit  $\phi_0 = \text{id}_V$ , Matrix  $E$ ,  $E_1 = V$ . Isomorphismus
3. Punktspiegelung an  $O$  mit  $\phi_0(\vec{x}) = -\vec{x}$ . Matrix  $-E$ ,  $E_{-1} = V$ , Isomorphismus
4. Zentrische Streckung an  $O$  um  $r$  mit  $\phi_0(\vec{x}) = r\vec{x}$ . Matrix  $rE$ ,  $E_r = V$ , Isomorphismus falls  $r \neq 0$ , sonst Bild= 0.
5. Parallelprojektion mit Kern  $K = \text{Kern}\phi_0$  auf  $U + O$  mit  $U = \text{Bild}\phi_0$ .  $E_1 = U$ ,  $E_0 = K$ . Isomorphismus nur für  $U = V$ . Matrix bzgl. Basis  $\vec{v}_1, \dots, \vec{v}_n$  mit  $\vec{v}_1, \dots, \vec{v}_r$  Basis von  $U$  und  $\vec{v}_{r+1}, \dots, \vec{v}_n$  Basis von  $K$ .

$$\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}$$

Bei der Orthogonalprojektion ist  $K = \{\vec{x} \mid \vec{x} \perp U\}$  und bzgl. ON-Basis gilt

$$\phi_0(\vec{x}) = \sum_{i=1}^r \langle \vec{v}_i \mid \vec{x} \rangle \vec{v}_i$$

Ist dabei  $U$  Hyperebene ( $r = n - 1$ ) und  $\vec{n} = \vec{v}_n$  Normalenvektor, so

$$\phi_0(\vec{x}) = \vec{x} - 2\langle \vec{n} \mid \vec{x} \rangle \vec{n}$$

6. Spiegelung an Hyperebene  $U + O$  wie in 5. Achse=Normale. Bewegung.

$$\phi_0(\vec{x}) = \vec{x} - 2\langle \vec{n} \mid \vec{x} \rangle \vec{n}, \quad \text{Matrix} \begin{pmatrix} E_{n-1} & O \\ O & -1 \end{pmatrix}, \quad U = E_1, \mathbb{R}\vec{n} = E_{-1}$$

7. Drehung in der reellen Ebene mit Zentrum  $O$  um Winkel  $\omega$  - gegen die Uhr. Bewegung. Bzgl. ON-Basis Matrix

$$\begin{pmatrix} \cos \omega & -\sin \omega \\ \sin \omega & \cos \omega \end{pmatrix}$$

Keine reellen EV, komplexe EV von Betrag 1 und Argument  $\pm\omega$ .

8. Drehung im reellen Raum mit Achse  $\mathbb{R}\vec{v}_1 + O$  um Winkel  $\omega$  - positiv im Sinne der Rechten Hand. Bewegung. Bzgl. ON-Basis  $\vec{v}_1, \vec{v}_2, \vec{v}_3$  Matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \omega & -\sin \omega \\ 0 & \sin \omega & \cos \omega \end{pmatrix}, \quad E_1 = \mathbb{R}\vec{v}_1, \quad E_{-1} = E_1^\perp \Leftrightarrow \omega = \pi$$

9. Drehspiegelung im reellen Raum mit Achse  $\mathbb{R}\vec{v}_1 + O$  um Winkel  $\omega \neq \pi$  - positiv im Sinne der Rechten Hand. Bewegung. Bzgl. ON-Basis  $\vec{v}_1, \vec{v}_2, \vec{v}_3$  Matrix

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos \omega & -\sin \omega \\ 0 & \sin \omega & \cos \omega \end{pmatrix}, \quad E_{-1} = \mathbb{R}\vec{v}_1$$

10. Scherung in der Ebene längs der Achse  $\mathbb{R}\vec{v}_1 + O$ . Matrix bzgl. Basis  $\vec{v}_1, \vec{v}_2$

$$\begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix}, \quad E_1 = \mathbb{R}\vec{v}_1$$

11. Zentralprojektion  $\phi$  der Ebene  $U + O$  auf die parallele Ebene  $U + O'$  des Raumes mit Zentrum  $Z \notin U + O \cup U + O'$ . Dabei ist  $\phi(P)$  der Schnittpunkt der Geraden durch  $PZ$  mit der Ebene  $U + O'$ . Hier  $\phi_0 : U \rightarrow U$  und mit Normalenvektor  $\vec{n}$  von  $U$

$$\phi_0(\vec{x}) = \frac{s}{r}\vec{x} \quad \text{wobei } r\vec{n} + Z \in U + O, \quad s\vec{n} + Z \in U + O'$$

## 22.3 Basistransformation

### 22.3.1 Koordinatentransformation für Vektoren

Seien  $\alpha : \vec{v}_1, \dots, \vec{v}_m$  und  $\beta : \vec{w}_1, \dots, \vec{w}_m$  zwei Basen des  $K$ -Vektorraums  $V$ . Wie rechnet man die Koordinaten  $\vec{x}^\alpha$  und  $\vec{x}^\beta$  ineinander um? Wir machen den Ansatz

$$\vec{v}^\alpha = {}^\alpha T_\beta \vec{v}^\beta \quad \text{für alle } \vec{v}$$

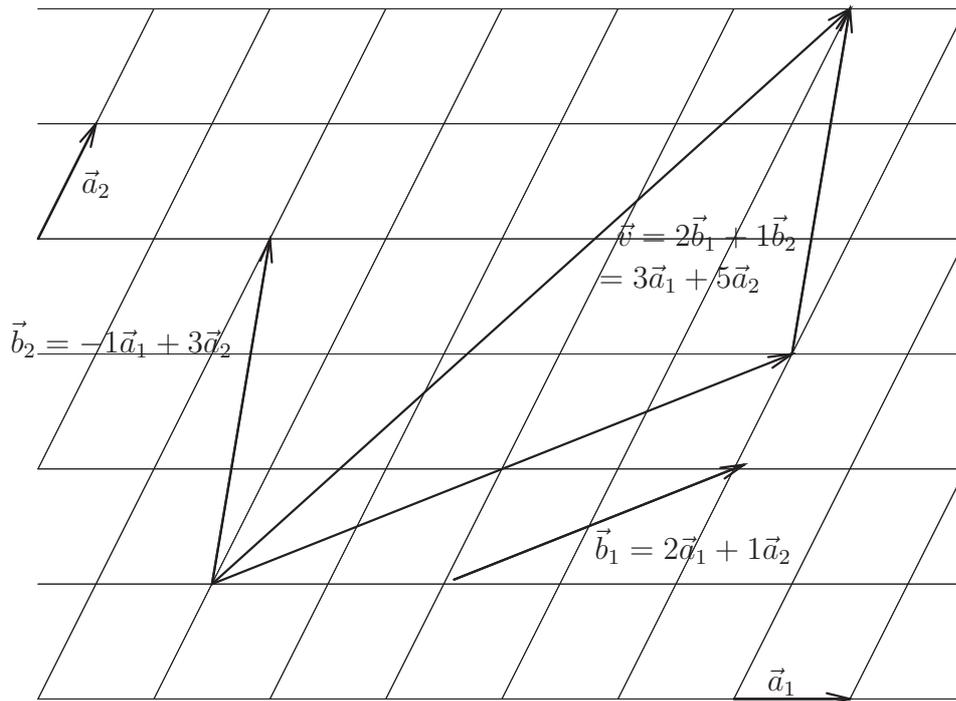
Die Transformationsmatrix ist dadurch eindeutig bestimmt: wir erhalten ihre Spalten als  ${}^\alpha T_\beta \mathbf{e}_j$  indem wir für  $\vec{v}$  die Vektoren aus  $\beta$  einsetzen.

Es sei also  $S = {}^\alpha T_\beta$  definiert als die *Transformationsmatrix* mit den Koordinatenspalten  $\vec{w}_j^\alpha$  der  $\vec{w}_j$  bzgl.  $\alpha$  als Spalten, also

$$S = {}^\alpha T_\beta = (\vec{w}_1^\alpha \dots \vec{w}_m^\alpha), \quad \vec{w}_j = s_{1j}\vec{v}_1 + \dots + s_{mj}\vec{v}_m$$

$$\boxed{\vec{x}^\alpha = {}^\alpha T_\beta \vec{x}^\beta, \quad \vec{x}^\beta = {}^\beta T_\alpha \vec{x}^\alpha \quad \text{mit} \quad {}^\beta T_\alpha = {}^\alpha T_\beta^{-1}}$$

Die Transformationsmatrix  $S = {}^\alpha T_\beta$  leistet die Koordinatenumrechnung von der 'neuen' Basis  $\beta$  in die 'alte' Basis  $\alpha$ . In ihren Spalten stehen die Koordinaten der neuen Basisvektoren bzgl. der alten Basis. Für die Umrechnung von 'alt' auf 'neu' benutzt man die inverse Matrix  $S^{-1} = {}^\beta T_\alpha$



Beweis.

$$\vec{x} = \sum_j y_j \vec{w}_j = \sum_j y_j \sum_i s_{ij} \vec{v}_i = \sum_{i,j} s_{ij} y_j \vec{v}_i = \sum_i \left( \sum_j s_{ij} y_j \right) \vec{v}_i$$

Also  $\vec{x}^\alpha = {}^\alpha T_\beta \vec{x}^\beta$ . Durch Vertauschen der Rollen  $\vec{x}^\beta = {}^\beta T_\alpha \vec{x}^\alpha$ , also  $\vec{x}^\alpha = {}^\alpha T_\beta {}^\beta T_\alpha \vec{x}^\alpha$ . Das gilt insbesondere für die  $\vec{x}^\alpha = \mathbf{e}_j$ , also  $E = {}^\alpha T_\beta {}^\beta T_\alpha E$ .  $\square$

Ist eine Basis  $\alpha$  des  $m$ -dimensionalen  $K$ -Vektorraums  $V$  gegeben, so kann man jede invertierbare  $m \times m$ -Matrix  $S$  auf genau eine Weise als Transformationsmatrix  $S = {}^\alpha T_\beta$  auffassen, nämlich mit der Basis  $\beta$  deren Koordinaten bzgl.  $\alpha$  durch die Spalten von  $S$  gegeben sind.

### 22.3.2 Orthogonale Transformationsmatrizen

**Lemma 22.12** *Ist  $\alpha$  Orthonormalbasis eines euklidischen Vektorraumes, so ist die Transformationsmatrix  ${}^\alpha T_\beta$  genau dann orthogonal, wenn  $\beta$  ebenfalls Orthonormalbasis ist*

Beweis. Sind die  $\vec{f}_j$  die Vektoren von  $\beta$  und berechnet man das Skalarprodukt in den Koordinaten bzgl.  $\alpha$ , so

$${}^\alpha T_\beta^t {}^\alpha T_\beta = (\langle \vec{f}_i | \vec{f}_j \rangle)_{1 \leq i, j \leq n}$$

Und  $\beta$  ist Orthonormalbasis genau dann, wenn letztere Matrix  $E$  ist.  $\square$

### 22.3.3 Transformation der Abbildungsbeschreibung

Sei  $\phi : V \rightarrow V$  linear. Sei  $A$  die Matrix von  $\phi$  bzgl. der (alten) Basis  $\alpha$  von  $V$  und  $B$  die Matrix von  $\phi$  bzgl. der neuen Basis von  $V$  also

$$\phi(\vec{x})^\alpha = A\vec{x}^\alpha, \quad \phi(\vec{x})^\beta = B\vec{x}^\beta$$

Dann gilt

$$B = {}^\alpha T_\beta^{-1} A {}^\alpha T_\beta$$

Beweis.  $B\vec{x}^\beta = (\phi(\vec{x}))^\beta = {}^\beta T_\alpha(\phi(\vec{x}))^\alpha = {}^\beta T_\alpha(A\vec{x}^\alpha) = ({}^\beta T_\alpha A)\vec{x}^\alpha = {}^\beta T_\alpha A {}^\alpha T_\beta \vec{x}^\beta$

- Anwendung 1. Sei  $B$  bezüglich einer ‘günstigen’ Basis  $\beta$  bekannt. Man bestimme Matrix  $A$  von  $\phi$  bzgl. der Ausgangs-Basis  $\alpha$ .
- Anwendung 2. Die Matrix  $A$  von  $\phi$  bezüglich der Ausgangs-Basis  $\alpha$  sei bekannt, sagt aber wenig über die Struktur. Man bestimme eine ‘günstige’ Basis  $\beta$ , sodass man der Matrix  $B$  von  $\phi$  bzgl.  $\beta$  geometrische Eigenschaften von  $\phi$  ansehen kann. Bzw. so, dass man das durch  $B$  gegebene, zu  $A$  gleichwertige, (Differential)Gleichungssystem lösen kann. Das ist das Thema der Eigenwerttheorie.
- Warnung: Der Gauss-Algorithmus ist in diesem Zusammenhang wenig hilfreich.

Beispiel: Euklidische Ebene mit ON-Basis  $\alpha : \vec{e}_1, \vec{e}_2$ , Spiegelung  $\phi$  Gerade  $\mathbb{R}(2\vec{e}_1 + \vec{e}_2)$ .  
Neue ON-Basis

$$\beta : \frac{1}{\sqrt{5}}(2\vec{e}_1 + \vec{e}_2), \frac{1}{\sqrt{5}}(-\vec{e}_1 + 2\vec{e}_2)$$

$$B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad {}^\alpha T_\beta = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}, \quad {}^\beta T_\alpha = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}$$

$$A = {}^\alpha T_\beta B {}^\beta T_\alpha = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 3 & 4 \\ 4 & -3 \end{pmatrix}$$

## 22.4 Determinanten

### 22.4.1 Regeln

Seien der Körper  $K$  und eine natürliche Zahl  $n$  gegeben. Eine (normierte) *Determinantenform* ist eine Abbildung  $\det$ , die jeder  $n \times n$ -Matrix  $A$

$$A = (\mathbf{a}_1 \dots \mathbf{a}_n)$$

über  $K$  einen Skalar

$$|A| = \det A = \det(\mathbf{a}_1, \dots, \mathbf{a}_n)$$

aus  $K$  zuordnet so, dass die folgenden Regeln gelten:

$$(D1) \det(\mathbf{a}_1, \dots, \mathbf{a}_j + \mathbf{b}_j, \dots, \mathbf{a}_n) = \det(\mathbf{a}_1, \dots, \mathbf{a}_j, \dots, \mathbf{a}_n) + \det(\mathbf{a}_1, \dots, \mathbf{b}_j, \dots, \mathbf{a}_n)$$

$$(D2) \det(\mathbf{a}_1, \dots, r\mathbf{a}_j, \dots, \mathbf{a}_n) = r \det(\mathbf{a}_1, \dots, \mathbf{a}_j, \dots, \mathbf{a}_n)$$

d.h.  $\det$  ist linear in jeder Spalte. Eine Determinante mit zwei benachbarten gleichen Spalten ist Null:

$$(D3) \det(\dots, \mathbf{a}, \mathbf{a}, \dots) = 0.$$

$$(D4) \det E_n = 1.$$

Ob's sowas gibt, wissen wir vorerst nicht, ziehen aber munter unsere Folgerungen über “die Determinante”.

Eine Determinante mit zwei gleichen Spalten ist Null:

$$(D3^+) \det(\dots, \mathbf{a}, \dots, \mathbf{a}, \dots) = 0.$$

(D5) Die Determinante ändert sich nicht, wenn man zu einer Spalte ein Vielfaches einer anderen addiert;

(D6) Die Determinante ändert das Vorzeichen, d.h. man muss mit  $-1$  multiplizieren, wenn man zwei Spalten vertauscht.

Beweis: Wir zeigen zunächst (D5) und (D6) für den Fall, dass die  $j$ -te und  $k$ -te Spalte benachbart sind.

$$\begin{aligned} \det(\dots, \mathbf{a}_j, \dots, r\mathbf{a}_j + \mathbf{a}_k, \dots) &=_{(D1)} \det(\dots, \mathbf{a}_j, \dots, r\mathbf{a}_j, \dots) + \det(\dots, \mathbf{a}_j, \dots, \mathbf{a}_k, \dots) \\ &=_{(D2)} r \det(\dots, \mathbf{a}_j, \dots, \mathbf{a}_j, \dots) + \det(\dots, \mathbf{a}_j, \dots, \mathbf{a}_k, \dots) =_{(D3)} 0 + \det(\dots, \mathbf{a}_j, \dots, \mathbf{a}_k, \dots). \\ \det(\dots, \mathbf{a}_j, \dots, \mathbf{a}_k, \dots) + \det(\dots, \mathbf{a}_k, \dots, \mathbf{a}_j, \dots) \\ &=_{(D5)} \det(\dots, \mathbf{a}_j, \dots, \mathbf{a}_j + \mathbf{a}_k, \dots) + \det(\dots, \mathbf{a}_k, \dots, \mathbf{a}_k + \mathbf{a}_j, \dots) \\ &=_{(D1)} \det(\dots, \mathbf{a}_j + \mathbf{a}_k, \dots, \mathbf{a}_k + \mathbf{a}_j, \dots) =_{(D3)} 0. \end{aligned}$$

Hat man (D6) für benachbarte Spalten, so folgt  $(D3^+)$  aus (D3) durch Induktion über den Positions-Abstand der zwei gleichen Spalten

$$\det(\dots \mathbf{a} \dots \mathbf{b} \mathbf{a} \dots) = -\det(\dots \mathbf{a} \dots \mathbf{a} \mathbf{b} \dots) = 0$$

Und dann folgen (D5) und (D6) mit obigem Beweis allgemein.  $\square$

#### 22.4.2 Eindeutigkeit und Berechnung

**Satz 22.13** Zu gegebenem  $n$  und  $K$  gibt es höchstens eine normierte Determinantenform. Man berechnet  $\det A$  durch Umformen nach (D5) und (D6) auf Dreiecksform (mit Berücksichtigung der Vorzeichenwechsel) und dann Produkt über die Diagonale. Es gilt  $\det A \neq 0$  genau dann, wenn  $A$  invertierbar ist.

Beweis. Durch (elementare) Spaltenumformungen nach (D5) und (D6) kann man  $A$  in eine Matrix  $A'$  in unterer Stufenform mit  $\det A' = \det A$  oder  $\det A' = -\det A$  überführen - und man weiß, welcher Fall vorliegt. Hat  $A'$  eine Spalte  $\mathbf{0}$ , so  $\det A' = 0$  nach (D2). Andernfalls handelt es sich wegen der Stufenform um eine untere Dreiecksmatrix mit Diagonaleinträgen  $a_{ii} \neq 0$  und man hat nach (D2)  $\det A' = a'_{11} \cdot \dots \cdot a'_{nn} \det A''$ , wobei  $A''$  untere Dreiecksmatrix mit Diagonaleinträgen 1 ist. Weitere Umformung nach (D5) überführt  $A''$  in  $E_n$ , also  $\det A'' = \det E_n = c$ .

Nach Gauß-Jordan ist  $A$  genau dann invertierbar, wenn  $A'$  eine (strikte) untere Dreiecksmatrix (auf der Diagonalen keine Null, darunter nur Nullen) ist, d.h.  $\det A = \det A' \neq 0$ .

**Korollar 22.14** Eine Abbildung von  $K^{n \times n}$  in  $K$ , die (D2), (D5) und (D6) erfüllt, ist durch ihren Wert an der Stelle  $E_n$  schon eindeutig bestimmt.

Beispiel 1 Hier mit Zeilen statt Spalten.

$$\det \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 4 & 2 \\ 0 & 3 & 0 & -3 \\ 1 & 1 & 2 & 6 \end{pmatrix} = 2 \cdot 3 \det \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 1 \\ 0 & 1 & 0 & -1 \\ 1 & 1 & 2 & 6 \end{pmatrix} = 6 \cdot \det \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & -1 & -1 & 2 \end{pmatrix}$$

$$= 6 \cdot \det \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & -2 & -2 \\ 0 & 0 & 1 & 3 \end{pmatrix} = 6 \cdot (-2) \det \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 3 \end{pmatrix} = -12 \det \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

und nach Satz ist die gesuchte Determinante gleich  $-12 \cdot 2 = -24$ .  $\blacksquare$

### 22.4.3 Produktsatz.

$$\det(AB) = \det A \det B$$

Beweis. Sei  $A$  fest,  $B$  variabel. Die Abbildung  $\mathbf{b} \mapsto A\mathbf{b}$  erfüllt (M1) und (M2), also erfüllen beide Abbildungen

$$B \mapsto \det AB \quad \text{und} \quad B \mapsto \det A \cdot \det B$$

die Bedingungen (D1-3) und  $E_n \mapsto \det A$ . Daher stimmen nach Kor.22.14 sie überein.  $\square$

Wir vermerken die wichtigen Spezialfälle

#### Korollar 22.15

$$\det(rA) = r^n \det A, \quad \det A^{-1} = \frac{1}{\det A}, \quad \det(S^{-1}AS) = \det(A).$$

### 22.4.4 Transponieren und Zeilenumformungen

**Satz 22.16**  $\det A = \det A^t$  und die Regeln (D1)- (D6) gelten entsprechend für Zeilenumformungen

**Korollar 22.17**  $\det(S^tAS) = (\det S)^2 \det A$ .  $|\det S| = 1$  falls  $S$  orthogonal.

Beweis. Wir betrachten jeweils eine (D2), (D5), bzw. (D6) entsprechende Zeilenumformung  $A \mapsto z(A)$ , d.h.  $z(A)$  entsteht aus  $A$  indem

- (a) die  $i$ -te Zeile mit  $s$  multipliziert wird ( $i$  und  $s$  fest)
- (b) das  $s$ -fache der  $i$ -ten Zeile zur  $l$ -ten addiert wird ( $i \neq l$  und  $s$  fest)
- (c) die  $i$ -te und die  $l$ -te Zeile vertauscht werden ( $i \neq l$  fest)

Wir behaupten, dass die Abbildung  $A \mapsto \det(z(A))$  jeweils (D2), (D5) und (D6) erfüllt. Dazu genügt es zu beobachten, dass in allen drei Fällen

$$\begin{aligned} z(\mathbf{a}_1, \dots, r\mathbf{a}_j, \dots, \mathbf{a}_n) &= (z(\mathbf{a}_1), \dots, rz(\mathbf{a}_j), \dots, z(\mathbf{a}_n)) \\ z(\dots, \mathbf{a}_j, \dots, r\mathbf{a}_j + \mathbf{a}_k, \dots) &= (\dots, z(\mathbf{a}_j), \dots, rz(\mathbf{a}_j) + z(\mathbf{a}_k), \dots) \\ z(\dots, \mathbf{a}_j, \dots, \mathbf{a}_k, \dots) &= (\dots, z(\mathbf{a}_j), \dots, z(\mathbf{a}_k), \dots) \end{aligned}$$

Wir behaupten, dass für jede Umformung  $z$  nach (a), (b) bzw. (c) die Abbildung  $A \mapsto \det(z(A))$  mit einer (passenden) der Abbildungen

$$A \mapsto s \det(A), \quad A \mapsto \det(A), \quad A \mapsto -\det(A)$$

übereinstimmt. Letztere erfüllen offensichtlich (D2), (D5) und (D6). Wegen Kor.22.14 haben wir daher die Übereinstimmung nur an der Stelle  $E_n$  nachprüfen. Das ist aber nicht schwer.

Somit haben wir für  $A \mapsto \det(A)$  die (D2),(D5) und (D6) entsprechenden Regeln für Zeilenumformungen bewiesen. Es folgen (D2), (D5) und (D6) für  $A \mapsto \det(A^t)$ . Für  $E_n$  erhalten wir 1 und haben daher  $\det(A^t) = \det(A)$  nach Kor.22.14.  $\square$

### 22.4.5 Entwicklung

Als zweckmässige Notation führt man ein: Der *Minor*  $A^{k \wedge l}$  ist die  $(n-1) \times (n-1)$ -Matrix, die man aus  $A$  durch Weglassen der  $k$ -ten Zeile und  $l$ -ten Spalte erhält. Dann hat man die Entwicklung nach der  $j$ -ten Spalte bzw. der  $i$ -ten Zeile

$$\det A = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det A^{i \wedge j}, \quad \det A = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det A^{i \wedge j}$$

Die Faktoren  $(-1)^{i+j}$  merkt man sich am besten nach der 'Schachbrettregel'

$$\begin{pmatrix} + & - & + & \dots \\ - & + & - & \dots \\ + & - & + & \dots \\ \vdots & & & \end{pmatrix}$$

Beweis für Spaltenentwicklung - Zeilen durch Transponieren. Seien  $\mathbf{a}_1, \dots, \mathbf{a}_n$  die Spalten von  $A$  und  $j$  fest. Dann  $\mathbf{a}_j = a_{1j}\mathbf{e}_1 + \dots + a_{nj}\mathbf{e}_j$ , also wegen (D1-2)

$$\det A = a_{1j} \det(\mathbf{a}_1, \dots, \mathbf{e}_1, \dots, \mathbf{a}_n) + \dots + a_{nj} \det(\mathbf{a}_1, \dots, \mathbf{e}_n, \dots, \mathbf{a}_n)$$

wobei jeweils die  $j$ -te Spalte ersetzt wurde. Durch Subtraktion des  $a_{ik}$ -fachen der  $j$ -ten Spalte von jeweils der  $k$ -ten werden alle Einträge der  $i$ -ten Zeile 0, ausser der 1 in der  $j$ -ten Spalte. Durch  $i+j$  Zeilen- bzw. Spaltenvertauschungen bringen wir diese 1 in die linke obere Ecke und haben

$$\det(\mathbf{a}_1, \dots, \mathbf{e}_i, \dots, \mathbf{a}_n) = (-1)^{i+j} \det \begin{pmatrix} 1 & O \\ O & A^{i \wedge j} \end{pmatrix} = (-1)^{i+j} \det A^{i \wedge j}.$$

Für den letzten Schluss benutzen wir wieder die Eindeutigkeit, hier für  $n-1$  und die Abbildung

$$B \mapsto \det \begin{pmatrix} 1 & \mathbf{0}^t \\ \mathbf{0} & B \end{pmatrix}, \quad B \in K^{(n-1) \times (n-1)}$$

Beispiel 2 Entwicklung nach der letzten Spalte liefert

$$\begin{aligned} \det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} &= 3 \det \begin{pmatrix} 4 & 5 \\ 7 & 8 \end{pmatrix} - 6 \det \begin{pmatrix} 1 & 2 \\ 7 & 8 \end{pmatrix} + 9 \det \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix} \\ &= 3 \cdot (-3) - 6(-6) + 9(-3) = 0. \end{aligned}$$

Beispiel 3 Entwicklung jeweils nach der ersten Spalte liefert

$$\begin{aligned} \det \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{pmatrix} &= a_{11} \det \begin{pmatrix} a_{22} & a_{23} & a_{24} \\ 0 & a_{33} & a_{34} \\ 0 & 0 & a_{44} \end{pmatrix} \\ &= a_{11} a_{22} \det \begin{pmatrix} a_{33} & a_{34} \\ 0 & a_{44} \end{pmatrix} = a_{11} a_{22} a_{33} \det(a_{44}) = a_{11} a_{22} a_{33} a_{44}. \end{aligned}$$

#### 22.4.6 Adjugierte Matrix

$$(adA)A = (\det A)E_n \quad \text{mit } adA = ((-1)^{i+j} \det A^{i\wedge j})^t$$

Beweis. Die Berechnung des  $j$ -ten Diagonalelements von  $(adA)A$  entspricht gerade dem Entwickeln nach der  $j$ -ten Spalte. Um zu sehen, dass in Position  $j, k$  mit  $k \neq j$  eine Null steht, betrachte man die Matrix  $B$ , die aus  $A$  entsteht, wenn man die  $k$ -te Spalte durch die  $j$ -te ersetzt - und selbige beibehält. Dann ist  $A^{i\wedge k}$  gleich  $B^{i\wedge j}$  bis auf eine (für alle  $i$  gleiche) Vertauschung und man erhält den Eintrag bis auf das Vorzeichen als die Entwicklung von  $B$  nach der  $j$ -ten Spalte, was aber wegen zweier gleicher Spalten Null ergibt.  $\square$

#### 22.4.7 Existenz

Im Falle  $n = 1$  sei  $\det A = a_{11}$ . Sei nun  $\det$  für  $n-1$  schon definiert. Definiere für  $n \times n$  Matrizen  $\det A$  durch die Entwicklung nach der ersten Zeile. Beim Nachweis von (D1-2) betrachten wir festes  $j$  und Matrizen  $A, B, C$  mit Spalten  $\mathbf{a}_k = \mathbf{b}_k = \mathbf{c}_k$  für alle  $k \neq j$ , also  $A^{1\wedge j} = B^{1\wedge j} = C^{1\wedge j}$ . Gilt  $\mathbf{c}_j = \mathbf{a}_j + \mathbf{b}_j$ , so folgt mit Induktion  $\det C^{1\wedge k} = \det A^{1\wedge k} + \det B^{1\wedge k}$  für  $k \neq j$  und man berechnet

$$\begin{aligned} \det C &= (-1)^{1+j}(a_{1j} + b_{1j}) \det C^{1\wedge j} + \sum_{k \neq j} (-1)^{1+k} a_{1k} \det C^{1\wedge k} \\ &= (-1)^{1+j}(a_{1j} \det A^{1\wedge j} + b_{1j} \det B^{1\wedge j}) + \sum_{k \neq j} (-1)^{1+k} a_{1k} \det(A^{1\wedge k} + B^{1\wedge k}) \\ &= (-1)^{1+j} a_{1j} \det A^{1\wedge j} + \sum_{k \neq j} (-1)^{1+k} a_{1k} \det A^{1\wedge k} \\ &\quad + (-1)^{1+j} b_{1j} \det B^{1\wedge j} + \sum_{k \neq j} (-1)^{1+k} b_{1k} \det B^{1\wedge k} \\ &= \det A + \det B \end{aligned}$$

Gilt  $\mathbf{c}_j = r\mathbf{a}_j$ , so gilt nach Induktion  $\det C^{1\wedge k} = r \det A^{1\wedge k}$  für  $k \neq j$  und somit

$$\begin{aligned} \det C &= (-1)^{1+j} r a_{1j} \det C^{1\wedge j} + \sum_{k \neq j} (-1)^{1+k} a_{1k} \det C^{1\wedge k} \\ &= (-1)^{1+j} r a_{1j} \det A^{1\wedge j} + \sum_{k \neq j} (-1)^{1+k} a_{1k} r \det A^{1\wedge k} = r \det A \end{aligned}$$

Hat  $A$  zwei gleiche Spalten  $\mathbf{a}_k = \mathbf{a}_l$ ,  $l = k + 1$ , so gilt das auch für alle  $A^{1\wedge j}$  mit  $j \neq k, l$ , also  $\det A^{1\wedge j} = 0$ . Die Minoren  $A^{1\wedge k}$  und  $A^{1\wedge l}$  stimmen überein und treten mit demselben Vorfaktor  $a_{1k} = a_{1l}$  aber entgegengesetztem Vorzeichen in der Entwicklung auf. Also  $\det A = 0$ .

Schliesslich  $\det E_n = (-1)^2 \det(E_n)^{1\wedge i} + \sum_{j \neq i} 0 \det(E_n)^{1\wedge j} = \det E_{n-1} = 1$ .

## 23 Transformation von Maßen und Integralen

### 23.1 Unabhängigkeit des Jordan-Maßes

#### 23.2 Vorbemerkung

Die Grundlage zur Behandlung des Raumes und physikalischer Größen ist der Begriff des euklidischen Raumes, eines affinen Raums mit zugehörigem Vektorraum, ausgestattet mit einer Längenmessung bzw. Skalarprodukt. Für diese kann man das Konzept einer Orthonormalbasis bzw. kartesischen Koordinatensystems einführen. Ist ein solches gegeben, kann man Vektoren bzw. Punkte durch Koordinatenspalten in  $\mathbb{R}^n$  bezeichnen - also den Raum auf diese Weise mit  $\mathbb{R}^n$  jeweils identifizieren. Dabei sollte man aber nicht vergessen, dass diese Koordinatisierung sekundär ist, dass primär die von Koordinaten unabhängige vektorielle und metrische Struktur gegeben ist.

$\mathbb{R}^3$  kommt heraus, wenn man den Raum mithilfe der Koordinaten bzgl. eines kartesischen Koordinatensystems beschreibt. Es gibt aber unendlich viele solche Koordinatensysteme, also unendlich viele gleichberechtigte Möglichkeiten, den Raum als  $\mathbb{R}^3$  aufzufassen. Berechnet man Längen bzw. Skalarprodukte über solche Koordinaten, so hängt das Ergebnis nicht vom Koordinatensystem ab, weil der Längenbegriff den Koordinaten vorausgeht.

Um Flächen bzw. Volumen zu messen, haben wir Zerlegungen benutzt, die zwei wesentlichen Anforderungen genügen mussten: Das Mass einer Zelle ergibt sich unmittelbar aus zugehörigen Längenmaßen; je zwei Zerlegungen haben eine gemeinsame Verfeinerung. Die zweite Bedingung war der Grund dafür, ein kartesisches Koordinatensystem auszuwählen und das Jordan-Maß nur über achsenparallele Rechteck-Zerlegungen bzgl. dieses einen Koordinatensystems zu definieren. Darauf aufbauend, haben wir allgemeinere Zerlegungen definiert, deren Zellen bzgl. des eingeführten Maßes messbar sind, und gezeigt, dass diese dann zum gleichen Maß führen. Um das anwenden zu können, müssen wir jedoch die Messbarkeit und das Maß geeigneter Zellen herleiten, insbesondere für Rechtecke bzw. Quader in beliebiger Lage. Für diese sollte sich jedoch das Maß als Produkt der Kantenlängen ergeben - d.h. das Maß unabhängig vom gewählten kartesischen Koordinatensystem sein. Wir haben also Folgendes zu zeigen.

**Satz 23.1** *Sind  $\alpha$  und  $\beta$  kartesische Koordinatensysteme der Ebene oder des Raumes mit den zugehörigen Jordan-Maßen  $\mu_\alpha$  bzw.  $\mu_\beta$ , so ist eine Punktmenge  $B$  genau dann bzgl.  $\mu_\alpha$  messbar, wenn sie bzgl.  $\mu_\beta$  messbar ist und zwar  $\mu_\alpha(B) = \mu_\beta(B)$ .*

Eine äquivalente Sichtweise ist, dass Messbarkeit und Maß einer Menge unter längenerhaltender Abbildung erhalten bleiben. Allgemeiner studieren wir Abbildungen  $\sigma$ , für die  $\mu(\sigma(B))$  aus  $\mu(B)$  bestimmt werden kann und wenden das auf die Berechnung von Integralen an.

#### 23.2.1 Eulers Dreh

**Lemma 23.2** *Euler. Zu je zwei ON-Basen  $\alpha, \beta$  des Raumes gibt es ON-Basen  $\beta_0, \dots, \beta_4$  mit  $\alpha = \beta_0$ ,  $\beta = \beta_4$  und so, dass  $\beta_{i+1}$  aus  $\beta_i$  durch Drehung um eine Achse oder Spiegelung an einer Koordinatenebene von  $\beta_i$  hervorgeht.*

Beweis. Nach Spiegelung an einer Koordinatenebene kann man annehmen, dass  $\alpha$  und  $\beta$  gleichorientiert sind. Haben  $\alpha$  und  $\beta$  einen Vektor gemeinsam, so drehe man das ‘Dreibein’  $\alpha$  um diesen, bis die beiden anderen Beine auf die von  $\beta$  zu liegen kommen. Um auf ein gemeinsames Bein zu kommen, wähle man ein Bein von  $\alpha$  und drehe  $\alpha$  um eine seiner anderen Achsen, bis das Bein in eine der Koordinatenebenen von  $\beta$  zu liegen kommt. Danach drehe man das Bein in dieser Ebene, bis es auf einem Bein von  $\beta$  liegt.  $\square$

### 23.2.2 Beweis

Wir zeigen, dass sich für jedes kartesische Koordinatensystem  $\alpha$  das Volumen  $\mu_\alpha(Q)$  eines  $n$ -dimensionalen Quaders mit Seitenlängen  $a_1, \dots, a_n$  unabhängig von der Lage ergibt als

$$(+)\quad \mu_\alpha(Q) = a_1 \cdot \dots \cdot a_n$$

Dabei spielt die Wahl des Ursprungs offenbar keine Rolle, es geht also nur um die Orthonormalbasen. Ist nun  $\beta$  eine solche, so ist  $\mu_\beta$  über zu  $\beta$  achsenparallele Quader-Zerlegungen definiert  $Z$  definiert, diese sind wegen der  $\mu_\alpha$ -Messbarkeit dieser Quader Zerlegungen bzgl.  $\alpha$  nach Def. 21.3.1. Also ist jedes  $\mu_\beta$ -messbare  $B$  nach Satz 21.9 auch  $\mu_\alpha$  messbar und  $\mu_\alpha(B) = \mu_\beta(B)$ .

Um (+) zu zeigen, betrachten wir zunächst  $n = 2$ , eine gegebene ON-Basis  $\alpha$  und ein Dreieck mit zwei zueinander senkrechten Seiten der Längen  $a$  und  $b$ . Wie in der Übung am Beispiel ausgeführt, kann man  $D$  in dreieckige Normalbereiche zerlegen und so  $\mu(D) = \frac{1}{2}ab$  herleiten. Um  $\mu(R) = ab$  für ein Rechteck der Seitenlängen  $a, b$  zu zeigen, zerlege man dieses durch die Diagonale in zwei rechtwinklige Dreiecke.

Sei nun  $n = 3$ . Ist eine Kante des Quaders  $Q$  achsenparallel, o.B.d.A.  $a_3$  in  $z$ -Richtung, so kann  $Q$  als Ordinatenmenge

$$\{(x, y, z) \mid (x, y) \in B, 0 \leq z \leq a_3\}$$

aufgefasst werden und nach dem Fall  $n = 2$  hat man  $\mu(B) = a_1 a_2$  also

$$\mu(Q) = \int_B a_3 d(x, y) = a_3 \mu(B) = a_1 a_2 a_3$$

(+) folgt nun zunächst für den Fall, dass  $\beta$  aus  $\alpha$  durch Drehung um eine Koordinatenachse oder Spiegelung an einer Koordinatenebene von  $\alpha$  hervorgeht, nach Euler dann auch allgemein. Für  $n > 3$  geht’s ähnlich weiter.  $\square$

## 23.3 Affine Abbildungen

### 23.3.1 Affine Abbildungen und ihre Determinanten

Eine *affine Abbildung*  $\phi : V \rightarrow V$  ist von der Form  $\phi(\vec{x}) = \phi_0(\vec{x}) + \vec{c}$  mit einer linearen Abbildung  $\phi_0$  und konstanten  $\vec{c} \in V$ , also bezgl. einer Basis  $\alpha$

$$\phi(\vec{x})^\alpha = A\vec{x}^\alpha + \vec{c}^\alpha$$

(wir fassen hier  $V$  als affinen Raum mit Ursprung  $\vec{0}$  auf). Wir definieren

$$\det(\phi) = \det A$$

was nach dem Produktsatz von der Basis  $\alpha$  unabhängig ist. In der Tat  $\det({}^\beta T_\alpha A {}^\alpha T_\beta) = \det A$  da  ${}^\beta T_\alpha = {}^\alpha T_\beta^{-1}$ .

**Lemma 23.3** *Affine Abbildungen sind stetig. Eine affine Abbildung  $\phi$  ist umkehrbar genau dann, wenn  $\det \phi \neq 0$ . Die Umkehrung  $\phi^{-1}$  ist dann auch affin und es gilt  $\det(\phi^{-1}) = (\det \phi)^{-1}$ .*

Beweis. Die Stetigkeit folgt aus der Matrixbeschreibung. Die Umkehrbarkeit von  $\phi$  ist gleichbedeutend zur Invertierbarkeit von  $A$  - und die Inverse hat Matrix  $A^{-1}$ .  $\square$

### 23.3.2 Bewegungen

Eine affine Abbildung mit  $\phi_0$  d.h.  $A$  orthogonal, so handelt ist eine *Bewegung*, d.h. es gilt

$$\|\phi(x) - \phi(y)\| = \|x - y\| \quad \text{für all } x, y \in \mathbb{R}^n$$

Nämlich  $\|Ax + c - (Ay + c)\| = \|A(x - y)\| = \|x - y\| = \|x + c - (y + c)\|$ . Umgekehrt folgt aus dieser Bedingung, dass  $\phi$  in dieser Form beschrieben werden kann.  $\phi$  hat eine Umkehrabbildung  $\phi^{-1}$  und diese ist ebenfalls eine Bewegung. Hier  $|\det \phi| = \pm 1$ .

Ist  $\phi_0 = \text{id}$  die identische Abbildung, also  $\phi(\vec{x}) = \vec{x} + \vec{c}$  so ist  $\phi$  eine *Verschiebung* oder *Translation*.

## 23.4 Maßtransformation bei affinen Abbildungen

### 23.4.1 Bewegungsinvarianz des Jordan-Maßes

Ist  $Z$  Zerlegung einer Menge  $B \subseteq \mathbb{R}^n$  und  $\sigma : \mathbb{R}^n \rightarrow \mathbb{R}^n$  eine Abbildung, so sei

$$\sigma(Z) = \{\sigma(C) \mid C \in Z\}$$

**Korollar 23.4** *Ist  $\sigma$  eine Bewegung von  $\mathbb{R}^n$  und  $B \subseteq \mathbb{R}^n$  messbar, so ist auch  $\sigma(B)$  messbar und  $\mu(\sigma(B)) = \mu(B)$ .*

Beweis. Ist  $\mu$  bzgl. der ON-Basis  $\alpha$  definiert, so sei  $\beta$  das Bild von  $\alpha$  unter  $\sigma$ . Ist nun  $Z_n$  eine Folge von Gitterzerlegungen  $B$  bzgl.  $\alpha$  mit Weite  $\rightarrow 0$ , so ist  $\sigma(Z_n)$  eine Folge von Gitterzerlegungen von  $\sigma(B)$  mit Weite  $\rightarrow 0$  und  $\mu(\sigma(C)) = \mu(C)$  für  $C \in Z_n$ , da  $C$  und  $\sigma(C)$  Quader mit gleichen Kantenlängen sind. Also  $\mu_\beta(\sigma(B)) = \mu_\alpha(B)$  und die Behauptung folgt aus der Basisunabhängigkeit des Jordan-Maßes.  $\square$

### 23.4.2 Volumen eines Spats

**Lemma 23.5** *Ist  $S$  das von den unabhängigen Vektoren  $\vec{a}_1, \dots, \vec{a}_n$  aufgespannte Parallelepiped (Parallelogramm bzw. Spat) und  $A$  die Matrix, deren Spalten die Koordinaten der  $\vec{a}_j$  sind, so gilt*

$$\mu(S) = \det A$$

Beweis. Wir zeigen

- Entsteht  $S'$  aus  $S$  durch Vertauschen zweier Vektoren, so gilt  $S' = S$ .
- Geht  $S'$  aus  $S$  durch Scherung hervor und ist  $S'$  messbar, so ist auch  $S$  messbar und es gilt  $\mu(S) = \mu(S')$ .

Das Erste ist Definition. Beim Zweitem wird in  $S'$  ein  $\vec{a}_i$  ersetzt durch  $\vec{a}'_i = \vec{a}_i + r\vec{a}_j$  mit einem  $j \neq i$ . O.B.d.A  $\vec{a}'_2 = \vec{a}_2 + r\vec{a}_1$ ,  $0 \leq r \leq \|\vec{a}_1\|$  und Winkel zwischen  $\vec{a}_1$  und  $\vec{a}_2$  höchstens  $90^\circ$ . Dann wird das Dreieck  $D O, r\vec{a}_1, \vec{a}_2$  bzw. (schiefe) Prisma mit Kante  $\vec{a}_3$  von  $S$  abgeschnitten und nach der Verschiebung  $\tau$  um  $\vec{a}_1$  auf der anderen Seite wieder angesetzt und  $S'$  zu erhalten. Also

$$S = (S' \setminus \tau(D)) \cup D$$

$D$  ist messbar (im Falle  $n = 3$  als Normalbereich bzgl. der  $\vec{v}_2, \vec{v}_3$ -Ebene eines kartesischen Koordinatensystem mit  $\vec{v}_1$  parallel  $\vec{a}_1$ ) und  $\mu(D) = \mu(\tau(D))$  und daher

$$\mu(S) = (\mu(S') - \mu(\tau(D))) + \mu(D) = \mu(S')$$

Ist nun  $A$  die Matrix, deren Koordinaten die Spalten der  $\vec{a}_i$  sind, so ist  $\det A \neq 0$  und  $A$  kann durch Spalten-Scherungen und Vertauschungen in Diagonalgestalt  $A'$  überführt werden mit  $|\det A| = |\det A'|$ .  $A'$  beschreibt dann einen achsenparallelen Quader  $Q$  mit Volumen

$$\mu(Q) = a'_{11} \cdot \dots \cdot a'_{nn} = \det A' = \det A$$

Führt man rückwärt diese Scherungen und Vertauschungen von  $Q$  ausgehend aus, so erhält man  $S$ , die Messbarkeit von  $S$  und  $\mu(S) = \mu(Q)$ .  $\square$

### 23.4.3 Affine Maßtransformation

**Satz 23.6** *Ist  $\sigma : \mathbb{R}^n \rightarrow \mathbb{R}^n$  eine affine Abbildung mit  $\det \sigma \neq 0$ , so ist  $B$  genau dann messbar, wenn  $\sigma(B)$  messbar ist und es gilt*

$$\mu(\sigma(B)) = \mu(B) |\det \sigma|$$

Dies kann man auch so interpretieren: misst man bzgl. des schiefwinkligen, durch die Vektoren  $\vec{v}_1, \dots, \vec{v}_n$  gegebenen Koordinatensystems  $\beta$ , so gilt

$$\mu_\beta(B) = \mu(B) \det(\vec{v}_1, \dots, \vec{v}_n)$$

Beweis. Wegen  $\det \sigma \neq 0$  ist Bild  $S = \sigma(C)$  eines Rechtecks bzw. Quaders ein Parallelogramm bzw. Spat der durch unabhängige Vektoren gegeben ist. Nach obigem Lemma ist  $\sigma(C)$  messbar mit  $\mu(\sigma(C)) = \mu(C) |\det \sigma|$ . Sind nun die  $Z_n$  Gitter-Zerlegungen von  $B$  mit Weite  $\rightarrow 0$ , so sind die  $\sigma(Z_n)$  Zerlegungen von  $\sigma(B)$  mit Weite  $\rightarrow 0$  und bei der Bestimmung von  $\mu(\sigma(B))$  bzgl. dieser Zerlegungen haben wir jeweils den Faktor  $|\det \sigma|$ .  $\square$

## 27 Eigenwerte und Jordan-Normalform

### 27.1 Eigenwerte und Eigenvektoren

#### 27.1.1 Motivation: Schwingungsgleichung

Eine freie Schwingung wird bekanntlich durch eine Differentialgleichung (zweiter Ordnung) der folgenden Form beschrieben

$$y'' + ay' + by = 0, \quad a, b \in \mathbb{R}$$

Setzen wir

$$\mathbf{y} = \begin{pmatrix} y \\ y' \end{pmatrix}, \quad \text{also} \quad \frac{\partial \mathbf{y}}{\partial t} = \begin{pmatrix} y' \\ y'' \end{pmatrix}$$

so erhalten wir daraus das gleichwertige System (erster Ordnung)

$$\frac{\partial \mathbf{y}}{\partial t} = A \begin{pmatrix} y \\ y' \end{pmatrix} \quad \text{mit} \quad A = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix}$$

Der Ansatz (mit komplexwertigen Lösungen)

$$\mathbf{y} = e^{\lambda t} \mathbf{v} \quad \text{mit} \quad \lambda \in \mathbb{C}, \quad \mathbf{v} \in \mathbb{C}^2$$

eingesetzt in die Systemgleichung ergibt

$$\frac{\partial \mathbf{y}}{\partial t} = A e^{\lambda t} \mathbf{v} = e^{\lambda t} A \mathbf{v}$$

Andererseits hat man die Ableitung

$$\frac{\partial \mathbf{y}}{\partial t} = \lambda e^{\lambda t} \mathbf{v} = e^{\lambda t} \lambda \mathbf{v}$$

Da  $|e^{\lambda t}| = |e^{\operatorname{Re} \lambda t}| \neq 0$ , liegt eine Lösung der angesetzten Form genau dann vor, wenn

$$A \mathbf{v} = \lambda \mathbf{v}$$

Den trivialen Fall  $\mathbf{v} = \mathbf{0}$  kann man von der Betrachtung ausschliessen. Mit den Mitteln der Linearen Algebra, die wir in diesem Kapitel diskutieren, lassen sich, zu gegebenen  $a, b$ , alle Paare  $\lambda, \mathbf{v}$  bestimmen, und damit alle Lösungen der angesetzten Form. Andererseits kann man zeigen, dass im Falle  $b \neq 0$  alle komplexwertigen Lösungen Linearkombinationen von Lösungen dieser Form sein müssen. Besonders einfach ist hier der Fall  $a = 0$ . Die Bedingung ist dann

$$\begin{pmatrix} v_2 \\ -bv_1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -b & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} \lambda v_1 \\ \lambda v_2 \end{pmatrix}$$

Es folgt

$$v_2 = \lambda v_1, \quad \lambda^2 v_1 = \lambda v_2 = -bv_1$$

Soll es überhaupt eine Lösung  $\mathbf{v} \neq \mathbf{0}$  geben, so muss gelten

$$\lambda^2 = -b \quad \text{d.h.} \quad \lambda^2 + b = 0$$

und hier hat man gerade die Lösungen

$$C \begin{pmatrix} 1 \\ \lambda \end{pmatrix}, \quad C \in \mathbb{C}$$

also

$$y(t) = Ce^{\lambda t}$$

(1) Sei  $b < 0$ . Dann  $\lambda_1 = \sqrt{-b}$ ,  $\lambda_2 = -\sqrt{-b}$  und man erhält die Lösungen

$$y_1(t) = C_1 e^{\sqrt{-b}t}, \quad y_2 = C_2 e^{-\sqrt{-b}t}$$

und somit mit Konstanten  $C_1, C_2$  alle Lösungen, nämlich

$$y(t) = C_1 e^{\sqrt{-b}t} + C_2 e^{-\sqrt{-b}t}$$

Die reellen Lösungen sind dann gerade die mit  $C_1, C_2 \in \mathbb{R}$

(2)  $b > 0$ . Dann  $\lambda_1 = e^{bj}$ ,  $\lambda_2 = e^{-bj}$  und man erhält die Lösungen

$$y(t) = C_1 e^{bjt} + C_2 e^{-bjt}$$

und damit die reellen Lösungen (vgl. Thm.10.6.(iii))

$$y(t) = A_1 \sin bt + A_2 \cos bt$$

### 27.1.2 Basistransformation

Zur Erinnerung: Sind  $\alpha, \beta$  Basen von  $V$  und  $\phi$  eine lineare Abbildung von  $V$  in  $V$  mit Matrix  $A$  bzgl.  $\alpha$  und  $B$  bzgl.  $\beta$ , so gilt

$$B = {}^\beta T_\alpha A {}^\alpha T_\beta = {}^\alpha T_\beta^{-1} A {}^\alpha T_\beta$$

Anwendung 1. Sei  $B$  von  $\phi$  bezüglich einer ‘günstigen’ Basis  $\beta$  bekannt. Man bestimme Matrix  $A$  von  $\phi$  bzgl. der kanonischen Basis  $\alpha$ .

Anwendung 2. Die Matrix  $A$  von  $\phi$  bezüglich der kanonischen Basis  $\alpha$  sei bekannt, sagt aber wenig über die Struktur. Man bestimme eine ‘günstige’ Basis  $\beta$ , sodass man der Matrix  $B$  von  $\phi$  bzgl.  $\beta$  geometrische Eigenschaften von  $\phi$  ansehen kann. Bzw. so, dass man das durch  $B$  gegebene, zu  $A$  gleichwertige, (Differential)Gleichungssystem lösen kann. Das ist das Thema der Eigenwerttheorie.

### 27.1.3 Definition

Definition. Sei  $V$  ein  $K$ -Vektorraum und  $\phi$  eine lineare Abbildung von  $V$  in  $V$ , z.B.  $V = K^n$  und  $\phi \mathbf{x} = A\mathbf{x}$ . Ein Skalar  $\lambda$  aus  $K$  heisst *Eigenwert*, kurz EW, von  $\phi$  bzw.  $A$ , wenn es einen Vektor  $\vec{v} \neq 0$  in  $V$  gibt, der die ‘Eigenwertgleichung’

$$\phi(\vec{v}) = \lambda \vec{v} \quad \text{bzw.} \quad A\mathbf{v} = \lambda \mathbf{v}$$

erfüllt. Jedes solche  $\vec{v}$  heisst ein *Eigenvektor*, kurz EV, von  $\phi$  bzw.  $A$  zum EW  $\lambda$ . Geometrisch bedeutet dies, dass  $\phi$  in Richtung von  $v$  eine Streckung um den Skalar  $\lambda$  ist.

Die Lösungen der Eigenwertgleichung zum EW  $\lambda$ , d.h. 0 und die EV, bilden offenbar einen Untervektorraum von  $V$ , den *Eigenraum*  $U_\lambda$ .

## 27.1.4 Beispiele

Bei der Drehung ( $\neq id$ ) im Raum ist die Drehachse (durch  $O$ !) Eigenraum zum Eigenwert 1. Weitere EW gibt es nicht, ausser bei der  $180^\circ$ -Drehung: hier ist die Drehebene (senkrecht zur Achse) Eigenraum zum Eigenwert  $-1$ .

Bei der Spiegelung an der Ursprungsebene  $E$  gilt  $U_1 = E$  und  $U_{-1}$  ist die Normale.

Bei der Parallelprojektion auf die Ursprungsebene  $E$  gilt  $U_1 = E$  und  $U_0$  ist die Ursprungsgerade in Projektionsrichtung.

Bei der zentrischen Streckung um den Faktor  $r$  ist der Raum Eigenraum zum EW  $r$ .

Der Eigenraum  $U_0$  ist  $\text{Kern}\phi$ , also 0 ein EW genau dann, wenn  $A$  bzw.  $\phi$  nicht invertierbar.

## 27.1.5 Ansatz

Die Koordinatenspalten (bzgl. einer Basis)  $\mathbf{v}$  der EV zum EW  $\lambda$  sind die von  $\mathbf{0}$  verschiedenen Lösungen des linearen Gleichungssystems

$$(A - \lambda E)\mathbf{x} = \mathbf{0}.$$

Beweis.  $A\mathbf{v} = \lambda\mathbf{v}$  ist gleichbedeutend mit  $(A - \lambda E)\mathbf{v} = \mathbf{0}$ . Die Existenz einer nichttrivialen Lösung bedeutet nach 12.4 dass  $(A - \lambda E)$  nicht invertierbar ist. Es stellt sich also die Aufgabe, die  $\lambda$  mit dieser Eigenschaft zu berechnen. Dazu ist, zumindest theoretisch, die Determinante geeignet.

## 27.1.6 Beispiel

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Für welche  $\lambda$  hat

$$(A - \lambda E)\mathbf{x} = \begin{pmatrix} 1 - \lambda & 1 & 0 \\ 0 & 1 - \lambda & 0 \\ 0 & 0 & 2 - \lambda \end{pmatrix} \mathbf{x} = \mathbf{0}$$

eine nichttriviale Lösung? Nur dann, wenn mindestens ein Diagonaleintrag Null ist, d.h.  $\lambda = 1$  oder  $\lambda = 2$ . Hier hat man die eindimensionalen Lösungsräume

$$U_1 = \text{Spann} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad U_2 = \text{Spann} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

## 27.1.7 Gekoppelte Schwingung

Für die gekoppelte Schwingung zweier Massen  $m_1$  und  $m_2$  und zwei Federn mit Konstanten  $c_1, c_2$  hat man für die Auslenkungen  $y_1, y_2$  die Differentialgleichungen

$$m_1 y_1'' = -c_1 y_1 + c_2 (y_2 - y_1), \quad m_2 y_2'' = -c_2 (y_2 - y_1).$$

Wählt man  $m_1 = m_2 = 1$ ,  $c_1 = 3$ ,  $c_2 = 2$  so hat man

$$\mathbf{y}'' = A\mathbf{y} \text{ mit } A = \begin{pmatrix} -5 & 2 \\ 2 & -2 \end{pmatrix}.$$

Der Ansatz

$$y_i = b_i \cos \omega t, \quad \text{also } \mathbf{y} = \cos \omega t \mathbf{v}$$

führt zu

$$\mathbf{y}'' = -\omega^2 \cos \omega t \mathbf{v}, \quad A\mathbf{v} = -\omega^2 \mathbf{v}.$$

Mit  $\lambda = -\omega^2$  kommen wir auf das Gleichungssystem für  $v_1, v_2$

$$\begin{aligned} (-5 - \lambda)b_1 + 2b_2 &= 0 \\ 2v_1 + (-2 - \lambda)v_2 &= 0 \end{aligned}.$$

Dieses hat eine nichttriviale Lösung genau dann, wenn die Determinante der Koeffizientenmatrix  $(-5 - \lambda)(-2 - \lambda) - 4 = 0$ , also  $\lambda_1 = -1$  oder  $\lambda_2 = -6$ . Damit erhalten wir  $\omega_1 = 1$  und  $\omega_2 = \sqrt{6}$ . Für den Vektor  $\mathbf{v}$  erhält man dann aus dem Gleichungssystem

$$\mathbf{v}_1 = r_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \text{bzw. } \mathbf{v}_2 = r_2 \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

Da eine Summe von Lösungen des Differentialgleichungssystems wieder eine ist, hat man dann Lösungen der Form (und man kann zeigen, dass das schon alle sind)

$$\mathbf{y} = r_1 \cos t \begin{pmatrix} 1 \\ 2 \end{pmatrix} + r_2 \cos \sqrt{6}t \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

## 27.2 Charakteristische Gleichung

### 27.2.1 Algebraische Bestimmung der Eigenwerte

**Satz 27.1** Die EW von  $\phi$  sind gerade die zu  $K$  gehörigen Nullstellen  $\lambda_i$  der Charakteristischen Gleichung

$$\det(\phi - \lambda \text{id}) = 0$$

Ist  $\phi$  bezüglich einer Basis  $\alpha$  durch die Matrix  $A$  gegeben, so

$$\det(\phi - \lambda \text{id}) = \det(A - \lambda E)$$

und man spricht auch von den EW der Matrix  $A$ .

Beweis. Nach dem Produktsatz hat man für invertierbare  $S$   $\det(A - \lambda E) = (\det S)^{-1} \det(A - \lambda E) (\det S) = \det(S^{-1}(A - \lambda E)S) = \det(S^{-1}AS - \lambda E)$ , also die Unabhängigkeit der charakteristischen Gleichung von der Basis. Die Existenz einer nichttrivialen Lösung von  $(A - \lambda E)\mathbf{x} = \mathbf{0}$  bedeutet aber, dass  $\det(A - \lambda E) = 0$ .

*Bemerkung.*  $\det(A - xE)$  ist ein Polynom in der Unbestimmten  $x$  vom Grade  $n$ .

### 27.2.2 Beispiel

$$A = \frac{1}{2} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix}.$$

Charakteristisches Polynom  $\lambda^3 + \frac{3}{4}\lambda - \frac{1}{4}$ , Eigenwerte  $\lambda_1 = -1$ ,  $\lambda_{2,3} = \frac{1}{2}$  Eigenräume

$$U_{-1} = \text{Spann} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}, \quad U_{\frac{1}{2}} = \text{Spann} \left( \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \right).$$

## 27.2.3 Mehr Beispiele

Beispiel 1 Für  $A = \begin{pmatrix} 3 & 4 \\ 5 & 2 \end{pmatrix}$  ist  $A - \lambda I = \begin{pmatrix} 3 - \lambda & 4 \\ 5 & 2 - \lambda \end{pmatrix}$  und daher

$$\det(A - \lambda I) = (3 - \lambda)(2 - \lambda) - 20 = \lambda^2 - 5\lambda - 14.$$

Die Nullstellen dieses Polynoms und damit die Eigenwerte von  $A$  sind also  $\lambda_1 = 7$  und  $\lambda_2 = -2$ . Analog erhält man für

$$B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{bzw.} \quad B - \lambda I = \begin{pmatrix} -\lambda & 1 \\ -1 & -\lambda \end{pmatrix}$$

$$\det(B - \lambda I) = (-\lambda)^2 + 1 = \lambda^2 + 1$$

und damit die Eigenwerte  $\lambda_1 = i$ ,  $\lambda_2 = -i$ . Eine reelle Matrix muss also keine reellen Eigenwerte besitzen. Nun sei

$$C = \begin{pmatrix} 0 & 1 & 1 \\ -3 & 2 & 3 \\ 1 & 1 & 0 \end{pmatrix} \quad \text{bzw.} \quad C - \lambda I = \begin{pmatrix} -\lambda & 1 & 1 \\ -3 & 2 - \lambda & 3 \\ 1 & 1 & -\lambda \end{pmatrix}.$$

Mit der Sarrusschen Regel findet man

$$\begin{aligned} \det(C - \lambda I) &= \lambda^2(2 - \lambda) + 3 - 3 - (2 - \lambda) + 3\lambda - 3\lambda \\ &= (\lambda^2 - 1)(2 - \lambda) = (\lambda - 1)(\lambda + 1)(2 - \lambda) \end{aligned}$$

Die Eigenwerte der Matrix  $C$  sind also  $\lambda_1 = 2$ ,  $\lambda_2 = -1$  und  $\lambda_3 = 1$ . Schließlich sei

$$D = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{bzw.} \quad D - \lambda I = \begin{pmatrix} -\lambda & 1 \\ 0 & -\lambda \end{pmatrix}.$$

Dann ist  $\lambda_1 = \lambda_2 = 0$  ein doppelter Eigenwert. ■

Zur Berechnung der zu einem Eigenwert  $\lambda$  von  $A$  gehörigen Eigenvektoren muss man das homogene Gleichungssystem  $(A - \lambda I)x = 0$  lösen.

Beispiel 2 Wir berechnen die Eigenvektoren der Matrizen  $C$  und  $D$  aus Beispiel 1. Zum Eigenwert  $\lambda_1 = 2$  von  $C$  erhält man die Matrix

$$C - \lambda_1 I = \begin{pmatrix} -2 & 1 & 1 \\ -3 & 0 & 3 \\ 1 & 1 & -2 \end{pmatrix}.$$

Vertauschen von 1. und 3. Zeile und Division der 2. Zeile durch 3 führen auf die folgende Matrix, die wir mit den Operationen des Gaußalgorithmus weiter vereinfachen:

$$\begin{pmatrix} 1 & 1 & -2 \\ -1 & 0 & 1 \\ -2 & 1 & 1 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 1 & -2 \\ 0 & 1 & -1 \\ 0 & 3 & -3 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 1 & -2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Die Lösungen dieses Gleichungssystems sind also alle Vektoren

$$t(1, 1, 1)^T \quad \text{mit } t \in \mathbb{R}, \quad (27.1)$$

und diese Vektoren bilden auch den Eigenunterraum von  $C$  zum Eigenwert  $\lambda_1 = 2$ . Die Eigenvektoren zu  $\lambda_1 = 2$  sind gerade die Vektoren (27.1) mit  $t \neq 0$ . Ganz analog findet man die Eigenvektoren

$$t(0, 1, -1)^T \quad \text{mit } t \in \mathbb{R} \setminus \{0\} \quad \text{bzw.} \quad t(1, 0, 1)^T \quad \text{mit } t \in \mathbb{R} \setminus \{0\}$$

zu den Eigenwerten  $\lambda_1 = -1$  bzw.  $\lambda_3 = 1$  von  $C$ . Für die Matrix  $D$  lautet das entsprechende Gleichungssystem zur Bestimmung der Eigenvektoren zum Eigenwert  $\lambda_1 = \lambda_2 = 0$

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \quad \text{bzw.} \quad x_2 = 0.$$

Alle Eigenvektoren haben also die Gestalt

$$t(1, 0) \quad \text{mit } t \in \mathbb{R} \setminus \{0\}.$$

#### 27.2.4 Tricks

Für Polynome vom Grad  $\geq 5$  gibt es keine Lösungsformel. Die für Grad 3 und 4 will oder kann auch kaum jemand anwenden. Übungs- und Klausuraufgaben sind aber meist mit der guten Absicht gestellt, dass sich die EW leicht bestimmen oder erraten lassen.

Trick Nr.1: Wenn man das Polynom  $\det(A - \lambda E)$  ausrechnet, sollte man nicht unbedacht ausmultiplizieren, sondern erst mal sehen, ob man nicht eine Zerlegung der Form  $p(\lambda)q(\lambda)$  finden kann. Die EW sind dann die Nullstellen von  $p(\lambda)$  und die von  $q(\lambda)$ .

Trick Nr.2: Soll man die (komplexen) Nullstellen von  $x^n + a_{n-1}x^{n-1} + \dots + a_0$  bestimmen und sind die  $a_i$  ganze Zahlen, so probiere man, ob nicht ein ganzzahliger Teiler von  $a_0$  Nullstelle ist.

Trick Nr.3: Hat man den Verdacht, dass  $\alpha$  Nullstelle von  $p(x)$  ist, so finde man eine Zerlegung  $p(x) = q(x)(x - \alpha)$  durch Polynomdivision oder nach dem Horner Schema.

### 27.3 Eigenräume und ihre Basen

#### 27.3.1 Vielfachheiten

Die *geometrische Vielfachheit* eines Eigenwerts  $\lambda_1$  von  $A$  ist die Dimension des Eigenraums  $U_{\lambda_1}$ , die *algebraische Vielfachheit* von  $\lambda_1$  ist die Zahl  $m$  derart, dass  $\lambda_1$  genau  $m$ -fache Nullstelle des Polynoms  $\det(A - \lambda E)$  ist, dh.  $\det(A - \lambda E) = (\lambda_1 - \lambda)^m q(\lambda)$  mit  $q(\lambda_1) \neq 0$ .

**Satz 27.2** *Geometrische Vielfachheit  $\leq$  algebraische Vielfachheit.*

Beweis: Sei  $\vec{v}_1, \dots, \vec{v}_k$  Basis von  $U_{\lambda_1}$ . Diese kann man zu einer Basis  $\beta$  von  $V = K^n$  ergänzen. Sei  $\phi : V \rightarrow V$  die lineare Abbildung mit Matrix  $A$  bezüglich der kanonischen Basis. Bezüglich  $\beta$  hat  $\phi$  eine Matrix  $A'$  so, dass in den ersten  $k$  Spalten gerade die  $\lambda_1 \varepsilon_j$ ,  $j = 1 \dots k$  stehen. Die Basistransformation liefert  $\det(A - \lambda E) = \det(A' - \lambda E) = (\lambda_1 - \lambda)^k \det \tilde{A}$ , wobei  $\tilde{A}$  die  $(n - k) \times (n - k)$  Untermatrix in der rechten unteren Ecke von  $A'$  ist. Also ist  $\lambda_1$  mindestens  $k$ -fache Nullstelle von  $\det(A - \lambda E)$ .

## 27.3.2 Unabhängigkeit von Eigenvektoren

**Satz 27.3** *EV  $\vec{v}_1, \dots, \vec{v}_k$  von  $\phi$  zu lauter verschiedenen EW  $\lambda_1, \dots, \lambda_k$  sind linear unabhängig. Insbesondere gibt es höchstens  $n = \dim V$  verschiedene EW zu  $\phi$*

**Hilfssatz.** *Ist  $\vec{v}$  EV zum EW  $\lambda$  und  $\mu \neq \lambda$  so ist  $(\phi - \mu \text{id})(\vec{v}) = (\lambda - \mu)\vec{v} \neq \vec{0}$  EV zum EW  $\lambda$ . Beweis*

$$\phi((\lambda - \mu)\vec{v}) = \lambda\phi\vec{v} - \mu\phi(\vec{v})\lambda^2\vec{v} - \mu\lambda\vec{v} = \lambda((\lambda - \mu)\vec{v})$$

Beweis durch Induktion über  $k$ . Sei eine Linearkombination

$$0 = r_1\vec{v}_1 + r_2\vec{v}_2 + \dots + r_k\vec{v}_k$$

gegeben. Wir haben zu zeigen, dass alle  $r_i = 0$  sind. Durch Anwenden von  $\psi = \phi - \lambda_1 \text{id}$  erhalten wir

$$\vec{0} = r_1\psi\vec{v}_1 + r_2\psi\vec{v}_2 + \dots + r_k\psi\vec{v}_k = \vec{0} + r_2\psi\vec{v}_2 + \dots + r_k\psi\vec{v}_k$$

Nach dem Hilfssatz sind die  $\psi\vec{v}_2, \dots, \psi\vec{v}_k$  Eigenvektoren zu  $\lambda_2, \dots, \lambda_k$  und nach Induktionsannahme linear unabhängig. Es folgt

$$r_2 = \dots = r_k = 0 \quad \text{also} \quad r_2 = \dots = r_k = 0$$

Somit  $r_1\vec{v}_1 = \vec{0}$  und  $r_1 = 0$  da  $\vec{v}_1 \neq \vec{0}$ .  $\square$

## 27.3.3 Diagonalisierbarkeit

$\phi$  bzw.  $A$  ist *diagonalisierbar*, wenn es eine Basis  $\vec{v}'_1, \dots, \vec{v}'_n$  gibt, bezüglich deren  $\phi$  durch eine Diagonalmatrix  $A'$  beschrieben wird -  $A' = S^{-1}AS$ , wobei die Transformationsmatrix  $S$  die Spalten  $\vec{v}'_1, \dots, \vec{v}'_n$  hat. Dies bedeutet, dass  $\vec{v}'_1, \dots, \vec{v}'_n$  eine Basis von EV  $\vec{v}'_i$  jeweils zum EW  $\lambda_i$  ist, wobei  $\lambda_1, \dots, \lambda_n$  die (nicht notwendig verschiedenen) Diagonaleinträge von  $A'$  sind.

**Satz 27.4**  *$\phi$  bzw.  $A$  ist genau dann diagonalisierbar, wenn  $n = \dim V$  die Summe  $n = \sum_{i=1}^k d_i$  der geometrischen Vielfachheiten der verschiedenen Eigenwerte  $\lambda_1, \dots, \lambda_k$  ist. Eine hinreichende Bedingung ist  $k = n$ , d.h. dass  $\phi$  bzw.  $A$  mindestens (also genau)  $n = \dim V$  verschiedene Eigenwerte hat.*

Beweis. Wähle für jeden Eigenraum eine Basis und liste diese Vektoren auf als  $\vec{v}_1, \dots, \vec{v}_m$  in der Reihenfolge der Eigenwerte. Wir haben  $m = n$  wegen der Voraussetzung über die Vielfachheiten. Wir behaupten, dass  $\vec{v}_1, \dots, \vec{v}_m$  unabhängig sind. In der Tat, sei  $\sum_i r_i\vec{v}_i = \vec{0}$ . Addieren wir Vektoren aus demselben Eigenraum jeweils zu  $\vec{w}_h$  zusammen, so bekommen wir  $\sum_{h=1}^k \vec{w}_h = \vec{0}$  und mit Satz 27.3  $\vec{w}_h = \vec{0}$ . Ist  $\vec{w}_i = r_l\vec{v}_l + \dots + r_{l+d_i}\vec{v}_{l+d_i}$ , so folgt  $r_i = 0$  für  $i = l, \dots, l + d_i$ , da  $\vec{v}_l, \dots, \vec{v}_{l+d_i}$  Basis von  $U_{\lambda_i}$ . Also sind die  $\vec{v}_1, \dots, \vec{v}_n$  unabhängig und bilden somit eine Basis von  $V$ .  $\square$

## 27.3.4 Potenzen und Inverse

**Korollar 27.5**  $A \in K^{n \times n}$  ist invertierbar genau dann, wenn 0 kein Eigenwert von  $A$  ist

**Satz 27.6** Sei  $k \geq 0$  eine ganze Zahl. Sind  $\lambda_1, \dots, \lambda_n$  die Eigenwerte von  $A \in K^{n,n}$ , so sind  $\lambda_1^k, \dots, \lambda_n^k$  die Eigenwerte von  $A^k$ . Ist  $A$  invertierbar, so sind  $\lambda_1^{-k}, \dots, \lambda_n^{-k}$  die Eigenwerte von  $A^{-k} = (A^{-1})^k$ .

Wir haben hier die Vereinbarung  $A^0 = I$  benutzt. Beweis. Ist  $A$  invertierbar so, auch  $A^k$  mit der Inversen  $A^{-k} := (A^{-1})^k$ . Ist  $A^k$  invertierbar,  $A^k B = E$ , so ist  $A$  invertierbar da  $AA^{k-1}B = E$ . Also dürfen wir  $\lambda_i \neq 0$  für alle  $i$  annehmen. Ist  $\mathbf{v} \neq \mathbf{0}$ , so  $\lambda_i^k \mathbf{v} \neq \mathbf{0}$  für alle  $k \in \mathbb{Z}$  und es folgt mit Induktion

$$A\mathbf{v} = \lambda\mathbf{v} \Leftrightarrow A^{k\pm 1}\mathbf{v} = \lambda^{\pm 1}A^k\mathbf{v} = \lambda^{\pm 1}\lambda^k\mathbf{v} = \lambda^{k\pm 1}\mathbf{v} \Leftrightarrow A^k\mathbf{v} = \lambda^k\mathbf{v}$$

## 27.4 Komplexe Eigenwerte und Eigenvektoren

Wir betrachten nun den Körper  $\mathbb{C}$  der komplexen Zahlen und den  $\mathbb{C}$ -Vektorraum  $\mathbb{C}^n$ .

## 27.4.1 Zerlegung des charakteristischen Polynoms

Nach dem Fundamentalsatz der Algebra zerfällt das charakteristische Polynom einer komplexen Matrix  $A \in \mathbb{C}^{n \times n}$  in Linearfaktoren

$$\det(A - xE) = \pm(\lambda_1 - x) \cdots (\lambda_n - x)$$

(zunächst wäre da noch ein konstanter Faktor, aber der muss  $\pm 1$  sein) und die  $\lambda_1, \dots, \lambda_n$  sind dann Eigenwerte, jeder kommt so oft in der Liste vor, wie seine algebraische Vielfach angibt.

**Satz 27.7** Sind  $\lambda_1, \dots, \lambda_n \in \mathbb{C}$  die Eigenwerte von  $A = (a_{ij})_{n,n}$ , so gilt

$$\lambda_1 + \lambda_2 + \dots + \lambda_n = a_{11} + a_{22} + \dots + a_{nn} \quad (27.2)$$

sowie

$$\lambda_1 \lambda_2 \cdots \lambda_n = \det A. \quad (27.3)$$

Die Summe  $a_{11} + a_{22} + \dots + a_{nn}$  der Hauptdiagonalelemente einer Matrix  $A$  heißt auch die *Spur* von  $A$ . Beweis. Für  $n = 3$ . Nach Sarrus

$$\det(A - xE) = -x^3 + (a_{11} + a_{22} + a_{33})x^2 - \dots + \det A$$

$$(\lambda_1 - x)(\lambda_2 - x)(\lambda_3 - x) = -x^3 + (\lambda_1 + \lambda_2 + \lambda_3)x^2 - \dots + \lambda_1 \lambda_2 \lambda_3$$

und die Behauptung folgt durch Koeffizientenvergleich. Für  $n > 3$  entsprechend durch komplettes Ausmultiplizieren.  $\square$

## 27.4.2 Komplexe Eigenwerte reeller Matrizen

Für eine Matrix  $A = (a_{ik})$  sei  $\bar{A} = (\bar{a}_{ik})$  die konjugierte (hier  $\bar{z} = a - jb$  falls  $z = a + jb$ ). Die Matrix über  $\mathbb{C}$  ist *reell*, wenn alle ihre Komponenten reell sind, d.h. wenn  $A = \bar{A}$ . Diese Notation überträgt sich auch auf Vektoren. **Achtung:**  $A^* = \bar{A}^t$  ist die zu  $A$  *adjungierte* Matrix. Da gibt es Verwechslungsgefahr, wenn man die Konjugation in  $\mathbb{C}$  als  $z^*$  schreibt.

**Satz 27.8** *Ist  $\mathbf{v}$  EV der reellen Matrix  $A$  zum EW  $\lambda$ , so ist der konjugierte Vektor  $\bar{\mathbf{v}}$  EV von  $A$  zum konjugierten EW  $\bar{\lambda}$ .*

Beweis.  $A\bar{\mathbf{v}} = \bar{A}\bar{\mathbf{v}} = \overline{A\mathbf{v}} = \overline{\lambda\mathbf{v}} = \bar{\lambda}\bar{\mathbf{v}}$ .  $\square$  Für

$$A = r \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}.$$

haben wir EW und EV

$$\lambda_{1/2} = r(\cos \alpha \pm j \sin \alpha), \quad \begin{pmatrix} \pm j \\ 1 \end{pmatrix}$$

In der Tat

$$A \begin{pmatrix} j \\ 1 \end{pmatrix} = r \begin{pmatrix} j \cos \alpha & -\sin \alpha \\ j \sin \alpha & \cos \alpha \end{pmatrix} = \lambda \begin{pmatrix} j \\ 1 \end{pmatrix}$$

*Bemerkung.* Der von  $\mathbf{v}$  und  $\bar{\mathbf{v}}$  aufgespannte  $\mathbb{C}$ -Untervektorraum  $U$  von  $\mathbb{C}^n$  hat eine Basis aus reellen Vektoren, z.B. wenn  $\mathbf{v}$  nicht reell

$$\mathbf{u}_1 = \frac{1}{2j}(\mathbf{v} - \bar{\mathbf{v}}) = \frac{-j}{2}(\mathbf{v} - \bar{\mathbf{v}}) = \text{Im}(\mathbf{v}), \quad \mathbf{u}_2 = \frac{1}{2}(\mathbf{v} + \bar{\mathbf{v}}) = \text{Re}(\mathbf{v}).$$

Dies ist auch eine Basis des Re-Vektorraums der reellen Vektoren aus  $U$ , der ein Untervektorraum von  $\mathbb{R}^n$  ist. Beweis:  $\mathbf{v} = j\mathbf{u}_1 + \mathbf{u}_2$  und  $\bar{\mathbf{v}} = -j\mathbf{u}_1 + \mathbf{u}_2$ .

Ist  $A$  eine reellen Matrix mit EV  $\mathbf{v} \neq \bar{\mathbf{v}}$  zum EW  $\lambda = r(\cos \alpha \pm j \sin \alpha)$ , so wird die Wirkung auf diesem Untervektorraum bezüglich dieser Basis  $\mathbf{u}_1, \mathbf{u}_2$  angegeben durch die Matrix

$$\begin{pmatrix} r \cos \alpha & -r \sin \alpha \\ r \sin \alpha & r \cos \alpha \end{pmatrix}.$$

Das gilt auch dann, wenn  $\mathbf{u}_1$  und  $\mathbf{u}_2$  nicht aufeinander senkrecht stehen. Man rechnet einfach nach

$$\begin{aligned} A\mathbf{u}_1 &= A \frac{1}{2j}(\mathbf{v} - \bar{\mathbf{v}}) = \frac{1}{2j}(A\mathbf{v} - A\bar{\mathbf{v}}) = \frac{1}{2j}(\lambda\mathbf{v} - \bar{\lambda}\bar{\mathbf{v}}) = \frac{1}{2j}[\lambda(j\mathbf{u}_1 + \mathbf{u}_2) - \bar{\lambda}(-j\mathbf{u}_1 + \mathbf{u}_2)] = \\ &= \frac{1}{2j}[(\lambda + \bar{\lambda})j\mathbf{u}_1 + (\lambda - \bar{\lambda})\mathbf{u}_2] = \frac{1}{2}(\lambda + \bar{\lambda})\mathbf{u}_1 + \frac{1}{2j}(\lambda - \bar{\lambda})\mathbf{u}_2 = \text{Re}(\lambda)\mathbf{u}_1 + \text{Im}(\lambda)\mathbf{u}_2 \\ &= r \cos \alpha \mathbf{u}_1 + r \sin \alpha \mathbf{u}_2. \end{aligned}$$

$$\begin{aligned} A\mathbf{u}_2 &= A \frac{1}{2}(\mathbf{v} + \bar{\mathbf{v}}) = \frac{1}{2}[\lambda(i\mathbf{u}_1 + \mathbf{u}_2) + \bar{\lambda}(-i\mathbf{u}_1 + \mathbf{u}_2)] = \frac{i}{2}(\lambda - \bar{\lambda})\mathbf{u}_1 + \frac{1}{2}(\lambda + \bar{\lambda})\mathbf{u}_2 = \\ &= -\text{Im}(\lambda)\mathbf{u}_1 + \text{Re}(\lambda)\mathbf{u}_2. \end{aligned}$$

Ohne Rechnen geht's so: Für  $A$  wie oben sind  $\mathbf{u}_1$  und  $\mathbf{u}_2$  gerade die kanonischen Basisvektoren, und für die haben wir den Zusammenhang ja schon hergestellt. Nun ist aber jede Basis kanonisch - wenn wir nur wollen.

Sind  $\mathbf{v}$  und  $\bar{\mathbf{v}}$  orthonormal bzgl. des komplexen Skalarprodukts, so sind  $\sqrt{2}\text{Im } \mathbf{v}$  und  $\sqrt{2}\text{Re } \mathbf{v}$  orthonormal.

## 27.5 Satz von Cayley-Hamilton

**Lemma 27.9** Sei  $A \in \mathbb{C}^{n \times n}$  mit EW  $\lambda_1, \dots, \lambda_n$  (alle entsprechend ihrer Vielfachheit aufgelistet). Dann gilt

$$(A - \lambda_1 E) \cdots (A - \lambda_n E) = O$$

Beweis durch Induktion über  $n$ . Zu  $\lambda_1$  gibt es EV  $\vec{v}_1$  und den kann man zu Basis ergänzen. Sei  $S$  die entsprechende Transformationsmatrix (die auf die kanonische Basis zurücktransformiert. Dann haben  $A$  und  $A' = S^{-1}AS$  dasselbe charakteristische Polynom. also dieselebe EW. Nun schaut  $A'$  so aus

$$A' = \begin{pmatrix} \lambda_1 & * \\ \mathbf{0} & B \end{pmatrix}$$

und Entwickeln von  $\det(A' - xE_n)$  nach der ersten Spalte ergibt

$$\det(A' - xE_n) = (\lambda_1 - x) \det(B - xE_{n-1})$$

d.h.  $B$  hat die EW  $\lambda_2, \dots, \lambda_n$ . Nach Induktionsannahme

$$(B - \lambda_2 E_{n-1}) \cdots (B - \lambda_n E_{n-1}) = O$$

Also  $\prod_{i=1}^n (A' - \lambda_i E_n)$

$$\begin{aligned} &= \begin{pmatrix} 0 & * \\ \mathbf{0} & B - \lambda_1 E_{n-1} \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 & * \\ \mathbf{0} & B - \lambda_2 E_{n-1} \end{pmatrix} \cdots \begin{pmatrix} \lambda_1 & + \\ \mathbf{0} & B - \lambda_n E_{n-1} \end{pmatrix} \\ &= \begin{pmatrix} 0 & + \\ \mathbf{0} & B - \lambda_1 E_{n-1} \end{pmatrix} \cdot \begin{pmatrix} \lambda_1^{n-1} & + \\ \mathbf{0} & \prod_{i=2}^n B - \lambda_i E_{n-1} \end{pmatrix} \\ &= \begin{pmatrix} 0 & + \\ \mathbf{0} & B - \lambda_1 E_{n-1} \end{pmatrix} \cdot \begin{pmatrix} \lambda_1^{n-1} & + \\ \mathbf{0} & O \end{pmatrix} = \begin{pmatrix} 0 & \mathbf{0}^t \\ \mathbf{0} & O \end{pmatrix} \end{aligned}$$

Durch Rücktransformation folgt

$$\prod_{i=1}^n (A - \lambda_i E) = \prod_{i=1}^n (SA'S^{-1} - \lambda_i E) =$$

$$S(A' - \lambda_1 E)S^{-1} \cdot S(A' - \lambda_2 E)S^{-1} \cdots S(A' - \lambda_n E)S^{-1} = S \prod_{i=1}^n (A' - \lambda_i E)S^{-1} = SOS^{-1} = O$$

## 27.6 Lineare Abbildungen in Ebene und Raum

## 27.6.1 Klassifikation ebener linearer Abbildungen

**Satz 27.10** Sei  $\phi$  eine lineare Abbildung der reellen Ebene mit charakteristischem Polynom

$$\chi(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \quad \lambda_i \in \mathbb{C}$$

Die  $\lambda_i$  sind bis auf die Reihenfolge eindeutig bestimmt und es tritt genau einer der folgenden Fälle ein ( $\phi_\beta$  bezeichnet die Matrix von  $\phi$  bzgl.  $\beta$ )

(i)	Zentrische Streckung $\lambda_1 = \lambda_2 \in \mathbb{R}$	für jede Basis $\beta$ $\phi_\beta = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1 \end{pmatrix}$
(ii)	Scherung $\lambda_1 = \lambda_2 \in \mathbb{R}$	es gibt Basis $\beta$ mit $\phi_\beta = \begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix}$
(iii)	(verzerrte) Achsenstreckung $\lambda_1, \lambda_2 \in \mathbb{R}, \lambda_1 \neq \lambda_2$	es gibt Basis $\beta$ mit $\phi_\beta = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$
(iv)	(verzerrte) Drehstreckung $\lambda_{1/2} = r(\cos \omega \pm j \sin \omega) \notin \mathbb{R}$	es gibt Basis $\beta$ mit $\phi_\beta = r \begin{pmatrix} \cos \omega & -\sin \omega \\ \sin \omega & \cos \omega \end{pmatrix}$

Beweis. Zu Fall (i) und (iii) siehe Satz 27.4. In Fall (iii) sind die  $\vec{b}_i$  Eigenvektoren zu den  $\lambda_i$  mit eindeutig bestimmten Richtungen. Zu Fall (ii) siehe Lemma 27.11 Zu Fall (iv) vgl. Satz 27.8.

**Lemma 27.11** Ist  $\dim V = 2$  und  $\phi : V \rightarrow V$  ( $V = \mathbb{R}^2$  und  $\phi \mathbf{x} = A\mathbf{x}$ ) linear mit EW  $\lambda$  von algebraischer Vielfachheit 2 und geometrischer Vielfachheit 1, und wählt man

$$\vec{b}_2 \notin U_\lambda = \text{Kern}(\phi - \lambda \text{id}), \quad \vec{b}_1 = (\phi - \lambda \text{id})(\vec{b}_2)$$

(z.B.  $\mathbf{b}_2$  als  $\mathbf{e}_i$ , der kein EV ist,  $\mathbf{b}_1$  als  $i$ -te Spalte von  $A - \lambda E$ ) so erhält man eine Basis  $\beta$  so, dass  $\phi$  folgende Matrix hat

$$B = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$

Beweis.  $\vec{b}_2$  kann man wählen, weil der Eigenraum  $U_\lambda$  eindimensional ist. Nach Cayley-Hamilton ist  $\vec{b}_1$  ein EV. Ist  $r_1 \vec{b}_1 + r_2 \vec{b}_2 = \vec{0}$ , so wendet man  $(\phi - \lambda \text{id})$  an und erhält  $r_2 \vec{b}_1 = \vec{0}$  also  $r_2 = 0$ . Dann aber auch  $r_1 = 0$ .  $\square$

Beispiel

$$A = \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix}, \quad \det(A - xE) = (3 - x)(1 - x) + 1 = x^2 - 4x + 4, \quad \text{EW } \lambda_1 = \lambda_2 = 2$$

$$A - \lambda_1 E = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}, \quad (A - \lambda E)\mathbf{e}_2 \neq \mathbf{0}$$

$$\mathbf{b}_2 = \mathbf{e}_2, \quad \mathbf{b}_1 = (A - \lambda_1 E)\mathbf{b}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$S = {}^\alpha T_\beta = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}, \quad J = S^{-1}AS = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$$

## 27.6.2 Schwingungsgleichung

Die Matrix zur Schwingungsgleichung hat das charakteristische Polynom

$$\lambda^2 + a\lambda + b$$

Haben wir EW  $\lambda_1 \neq \lambda_2$  so erhalten wir unabhängige Lösungen  $y_i(t) = e^{\lambda_i t}$  also alle Lösungen durch die

$$y(t) = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t}$$

Der Fall  $\lambda = \lambda_1 = \lambda_2$  tritt genau dann auf, wenn  $b = \frac{a^2}{4}$  und somit  $\lambda = -\frac{a}{2}$ . Da  $A \neq \lambda E$ , muss dann die geometrische Vielfachheit 1 sein. Hier führt der Ansatz nur zur Lösung  $y(t) = e^{\lambda t}$ . Betrachtet man die transformierte Matrix, so kommt man zu den System

$$\begin{pmatrix} z_1'(t) \\ z_2'(t) \end{pmatrix} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} z_1(t) \\ z_2(t) \end{pmatrix} = \begin{pmatrix} \lambda z_1(t) + z_2(t) \\ \lambda z_2(t) \end{pmatrix}$$

und errät schnell die Lösung

$$\begin{pmatrix} z_1(t) \\ z_2(t) \end{pmatrix} = e^{\lambda t} \begin{pmatrix} c_1 + c_2 t \\ c_2 \end{pmatrix}$$

Durch Rücktransformation sieht man, dass auch das Ausgangssystem Lösungen dieser Gestalt hat.

## 27.6.3 Lineare Abbildungen im Raum

Sei  $V$  ein 3-dimensionaler  $\mathbb{R}$ -Vektorraum und  $\phi : V \rightarrow V$  linear ( $V = \mathbb{R}^3$  und  $\phi \mathbf{x} = A\mathbf{x}$ ).  $\phi$  ist diagonalisierbar genau dann, wenn  $\phi$

- (1) 3 verschiedene EW hat, z.B. falls es nicht reellen EW  $\lambda$  gibt - dann ist auch  $\bar{\lambda}$  ein EW
- (2) 2 EW hat, einen davon von geometrischer Vielfachheit 2
- (3) nur einen EW  $\lambda$  hat - der ist reell und  $\phi = \lambda \text{id}$

Damit bleiben als nicht diagonalisierbare Fälle

- (4)  $\phi$  hat EW  $\lambda_1 \neq \lambda_2$  (beide reell) und zwar von algebraischer Vielfachheit 2 bzw. 1, beide von geometrischer Vielfachheit 1. Wähle  $\vec{b}_2 \in \text{Bild}(\phi - \lambda_2 \text{id}) \setminus \text{Kern}(\phi - \lambda_1 \text{id})$  (z.B.  $\mathbf{b}_2$  als Spalte von  $A - \lambda_2 E$ , die kein EV zu  $\lambda_1$  ist) und  $\vec{b}_1 = (\phi - \lambda_1 \text{id})(\vec{b}_2)$ . Wähle  $\vec{b}_3$  als EV zu  $\lambda_2$ . Dann hat  $\phi$  bzgl. der Basis  $\beta : \vec{b}_1, \vec{b}_2, \vec{b}_3$  die Matrix

$$\begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}$$

Beispiel.

$$A = \begin{pmatrix} 3 & 1 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \det(xE - A) = (x^2 - 4x + 4)(x - 1), \text{ EW } \lambda_1 = \lambda_2 = 2, \lambda_3 = 1$$

$$A - \lambda_1 E = \begin{pmatrix} 1 & 1 & 0 \\ -1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}, \quad A - \lambda_3 E = \begin{pmatrix} 2 & 1 & 0 \\ -1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$$(A - \lambda_1 E)(A - \lambda_3 E)\mathbf{e}_2 \neq \mathbf{0}, \quad \mathbf{b}_2 = (A - \lambda_3 E)\mathbf{e}_2 = \mathbf{e}_1$$

$$\mathbf{b}_1 = (A - \lambda_1 E)\mathbf{b}_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \quad (A - \lambda_3 E)\mathbf{b}_3 = \mathbf{0}, \quad \mathbf{b}_3 = \mathbf{e}_3$$

$${}^\alpha T_\beta = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

- (5)  $\phi$  hat nur einen EW  $\lambda$  und zwar von geometrischer Vielfachheit 2. Entsprechend Lemma 27.11 wähle  $\vec{b}_2 \notin \text{Kern}(\phi - \lambda \text{id})$  (z.B. als  $\mathbf{e}_i$ , der kein EV ist) und  $\vec{b}_1 = (\phi - \lambda \text{id})(\vec{b}_2)$ . Wähle  $\vec{b}_3$  als EV zu  $\lambda$ , aber nicht als Vielfaches von  $\vec{b}_1$ . Dann hat  $\phi$  bzgl. der Basis  $\beta : \vec{b}_1, \vec{b}_2, \vec{b}_3$  die Matrix

$$\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}$$

Beispiel.

$$A = \begin{pmatrix} 3 & 1 & 1 \\ -1 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}, \quad \det(xE - A) = (x^2 - 4x + 4)(x - 2), \quad \text{EW } \lambda_1 = \lambda_2 = \lambda_3 = 2$$

$$A - \lambda_1 E = \begin{pmatrix} 1 & 1 & 1 \\ -1 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

$$(A - \lambda_1 E)\mathbf{e}_1 \neq \mathbf{0}, \quad \mathbf{b}_2 = \mathbf{e}_1$$

$$\mathbf{b}_1 = (A - \lambda_1 E)\mathbf{b}_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad (A - \lambda_1 E)\mathbf{b}_3 = \mathbf{0}, \quad \mathbf{b}_3 \notin \text{Spann}\{\mathbf{b}_1\}, \quad \mathbf{b}_3 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

$${}^\alpha T_\beta = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

- (6)  $\phi$  hat nur einen EW  $\lambda$  und zwar von geometrischer Vielfachheit 1. Wähle  $\vec{b}_3 \in V \setminus \text{Kern}(\phi - \lambda \text{id})^2$  (z.B. als  $\mathbf{e}_i$ , wobei die  $i$ -te Spalte von  $(A - \lambda E)^2$  nicht  $\mathbf{0}$  ist) und

$$\vec{b}_2 = (\phi - \lambda \text{id})(\vec{b}_3), \quad \vec{b}_1 = (\phi - \lambda \text{id})(\vec{b}_2)$$

Dann hat  $\phi$  bzgl. der Basis  $\beta : \vec{b}_1, \vec{b}_2, \vec{b}_3$  die Matrix

$$\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$$

Beispiel

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}, \quad \det(xE - A) = (x - 2)^3, \quad \text{EW } \lambda_1 = \lambda_2 = \lambda_3 = 2$$

$$A - \lambda_1 E = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad (A - \lambda_1 E)^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$(A - \lambda_1 E)^2 \mathbf{e}_3 \neq \mathbf{0}, \quad \mathbf{b}_3 = \mathbf{e}_3, \quad \mathbf{b}_2 = (A - \lambda_1 E)\mathbf{b}_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{b}_1 = (A - \lambda_1 E)\mathbf{b}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$${}^\alpha T_\beta = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Die hier gefundenen Matrizen sind *Jordan-Matrizen* und bestimmen eine *Jordan-Normalform* für  $\phi$ .

Beweis. Es ist zu zeigen, dass die angegebene  $\beta$  Basen sind. Die Gestalt der Matrix ist dann klar. Zu (4).  $A - \lambda_2 E$  hat Rang 2, d.h.  $\text{Bild}(\phi - \lambda_2 \text{id})$  ist 2-dimensional, aber der Eigenraum zu  $\lambda_1$  ist eindimensional, also gibt es  $\vec{b}_2$ . Nach Lemma 27.9 gilt  $\vec{b}_2 \in \text{Kern}(\phi - \lambda_1)^2$  und die Behauptung folgt mit Lemma 27.11. Hier gilt  $\text{Bild}(\phi - \lambda_2 \text{id}) = \text{Kern}(\phi - \lambda_1)^2$ .

In (5) und (6) haben wir nach Lemma 27.9  $V = \text{Kern}(\phi - \lambda \text{id})^3$ . Ist  $V = \text{Kern}(\phi - \lambda \text{id})^2$ , so gehen wir wie in (5) beschrieben vor. Dass  $\vec{b}_1, \vec{b}_2$  unabhängig sind und  $\vec{b}_1$  ein EV, folgt wie in Lemma 27.11. Da  $\vec{b}_2$  kein EV ist enthält  $\text{Spann}\{\vec{b}_1, \vec{b}_2\}$  nur solche EV, die Vielfache von  $\vec{b}_1$  sind. Daher können wir mit einem EV  $\vec{b}_3$  zu Basis ergänzen und wir sind im Fall (5). Sei nun also  $V \neq \text{Kern}(\phi - \lambda \text{id})^2$  und  $\beta$  wie in (6) gewählt. Dann ist  $\vec{b}_1$  ein EV. Angenommen,  $r_1 \vec{b}_1 + r_2 \vec{b}_2 + r_3 \vec{b}_3 = \vec{0}$ . Anwenden von  $\phi - \lambda \text{id}$  ergibt  $r_2 \vec{b}_1 + r_3 \vec{b}_2 = \vec{0}$  und wie in Lemma 27.11  $r_2 = r_3 = 0$ . Somit auch  $r_1 = 0$ .  $\square$

## 27.7 Jordan-Normalform

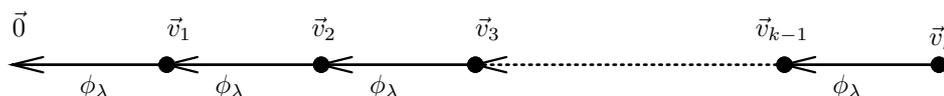
### 27.7.1 Jordan-Ketten

Sei  $\phi$  eine lineare Abbildung von  $V$  in  $V$  und  $\lambda$  ein EW von  $\phi$ . Dann bilden  $\vec{v}_1, \dots, \vec{v}_k$  aus  $V$  eine  $\lambda$ -*Jordankette* der Länge  $k$  mit *Startvektor*  $\vec{v}_k$  und *Eigenvektor*  $\vec{v}_1$  für  $\phi$  genau dann, wenn  $\vec{v}_1 \neq \vec{0}$  und

$$\vec{0} = \phi(\vec{v}_1) - \lambda \vec{v}_1, \quad \vec{v}_1 = \phi(\vec{v}_2) - \lambda \vec{v}_2, \quad \vec{v}_2 = \phi(\vec{v}_3) - \lambda \vec{v}_3, \dots, \vec{v}_{k-1} = \phi(\vec{v}_k) - \lambda \vec{v}_k$$

bzw. mit  $\phi_\lambda(x) = \phi(x) - \lambda x$ , d.h. mit  $\phi_\lambda = \phi - \lambda \text{id}$

$$0 = \phi_\lambda(\vec{v}_1), \quad \vec{v}_1 = \phi_\lambda(\vec{v}_2), \quad \vec{v}_2 = \phi_\lambda(\vec{v}_3), \dots, \vec{v}_{k-1} = \phi_\lambda(\vec{v}_k).$$



Ist  $V = K^n$ ,  $\mathbf{v}_i = v_i^\alpha$  und  $A = \phi_\alpha$  so schreibt sich das so

$$\mathbf{0} = (A - \lambda E)\mathbf{v}_1, \quad \mathbf{v}_1 = (A - \lambda E)\mathbf{v}_2, \quad \dots, \quad \mathbf{v}_{k-1} = (A - \lambda E)\mathbf{v}_k,$$

**Lemma 27.12** Die Vektoren  $\vec{v}_1, \dots, \vec{v}_k$  einer Jordan-Kette von  $\phi$  sind linear unabhängig.

Beweis durch Induktion. Sei  $\sum_{i=1}^k r_i \vec{v}_i = \vec{0}$ . Dann

$$\sum_{i=2}^k r_i \vec{v}_{i-1} = \sum_{i=2}^k r_i \phi_\lambda \vec{v}_i = \phi_\lambda \sum_{i=1}^k r_i \vec{v}_i = \vec{0}$$

also nach Induktion  $r_2 = \dots = r_k = 0$ . Somit  $r_1 \vec{v}_1 = \vec{0}$  also auch  $r_1 = 0$ .  $\square$

### 27.7.2 Jordan-Blöcke

Ein  $k$ -Jordanblock zum Eigenwert  $\lambda$  ist eine  $k \times k$ -Matrix

$$\begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & \lambda \end{pmatrix}.$$

**Korollar 27.13** Die Matrix einer Abbildung  $\phi$  von  $V$  in  $V$  bzgl. einer Basis  $\beta : \vec{v}_1, \dots, \vec{v}_n$  von  $V$  ist von der Form

$$\begin{pmatrix} J & + \\ O & + \end{pmatrix}$$

mit einem  $k$ -Jordanblock zum EW  $\lambda$  genau dann, wenn  $\vec{v}_1, \dots, \vec{v}_k$  eine  $\lambda$ -Jordankette für  $\phi$  ist.

### 27.7.3 Jordan-Basen

**Lemma 27.14** Sei hat man mehrere Jordanketten von  $\phi$  und sind die Eigenvektoren dieser Ketten unabhängig, so haben die Ketten keinen Vektor gemeinsam und alle ihre Vektoren zusammen sind unabhängig.

Eine Basis, die so durch Aneinanderreihung von Jordan-Ketten entsteht, ist eine *Jordan-Basis* von  $\phi$ .

Beweis durch Induktion über die Summe der Längen der Ketten. Sei eine Linearkombination des Nullvektors gegeben. Wir haben zu zeigen, dass alle Koeffizienten 0 sind. Sei  $\lambda$  EW einer der Ketten. Man wende  $\phi - \lambda \text{id}$  auf alle gegebenen Vektoren an. Dann werden die zu  $\lambda$  gehörigen Ketten alle um ihren EW verkürzt, wobei die Koeffizienten zum Start hin verschoben werden (vgl. Beweis von Lemma 27.12). Aus Ketten zu EW  $\mu \neq \lambda$  werden wieder Ketten zu  $\mu$  der gleichen Länge und mit denselben Koeffizienten, da

$$(A - \mu E)(A - \lambda E)\mathbf{v} = (A - \lambda E)(A - \mu E)\mathbf{v}$$

und da die EV nicht in den Nullvektor übergehen (Lemma 27.3). Nach Induktionsannahme sind nun alle Koeffizienten = 0, ausser denen bei den EW zu  $\lambda$ . Da diese aber als unabhängig vorausgesetzt waren, sind auch diese Koeffizienten = 0.  $\square$

## 27.7.4 Jordan-Matrizen

Eine *Jordanmatrix*  $J$  ist blockdiagonal mit Jordan Kästchen auf der Diagonale.

**Korollar 27.15** Die Matrix einer Abbildung  $\phi$  von  $V$  in  $V$  bzgl. einer Basis  $\beta$  ist eine Jordan-Matrix genau dann, wenn  $\beta$  eine Jordan-Basis ist.

Die einzelnen Jordan-Ketten entsprechen dabei den Blöcken.

## 27.7.5 Hauptsatz

**Satz 27.16** Sei  $V$  ein  $K$  Vektorraum und  $\phi : V \rightarrow V$  eine lineare Abbildung. Genau dann ist das charakteristische Polynom  $\det(\phi - \text{id})$  ein Produkt von Linearfaktoren  $\lambda_i - x$  mit  $\lambda_i \in K$ , wenn  $\phi$  eine Jordanbasis besitzt.

**Korollar 27.17** Sei  $A \in K^{n \times n}$  eine Matrix so, dass  $\det(A - xE) = \prod_{i=1}^n (\lambda_i - x)$  mit  $\lambda_i \in K$ , d.h. das charakteristische Polynom zerfällt in  $K$  in Linearfaktoren (z.B. wenn  $K = \mathbb{C}$  - Fundamentalsatz der Algebra). Dann gibt es eine invertierbare Matrix  $S \in K^{n \times n}$  so, dass  $A' = S^{-1}AS$  Jordanmatrix ist, die Jordansche Normalform von  $A$ .

Im Folgenden betrachten wir nur solche  $\phi$  bzw.  $A$ , z.B. wir lassen wir nur den Körper  $K = \mathbb{C}$  zu. Es stellen sich drei Fragen

- Wie eindeutig ist die Jordan-Normalform einer Matrix? Antwort: bis auf Vertauschung der Blöcke.
- Wie bestimmt man die Jordan-Normalform einer Matrix? Antwort: Indem man die Zahlen  $\text{Rang}(A - \lambda E)$  berechnet.
- Wie bestimmt man eine Jordan-Basis für eine lineare Abbildung bzw. ein  $S$  so, dass  $S^{-1}AS$  Jordan-Matrix ist? Antwort:
  - Indem man die Startvektoren von Jordan Ketten so sucht, dass die zugehörigen EW unabhängig sind und die Summe der Kettenlängen  $n = \dim V$  ergibt.
  - Das kann man für jeden EW separat tun.
  - Für jeden einzelnen EW  $\lambda$  bestimmt man
    - \*  $m$  minimal mit  $\text{Kern}(A - \lambda E)^m = \text{Kern}(A - \lambda E)^n$ ,  
d.h.  $\text{Rang}(A - \lambda E)^m = \text{Rang}(A - \lambda E)^{m+1}$
    - \* und sucht dann zunächst Startvektoren  $\mathbf{v}$  für Ketten der (maximalen) Länge  $m$ , also

$$(A - \lambda E)^m \mathbf{v} = \mathbf{0}, \quad (A - \lambda E)^{m-1} \mathbf{v} \neq \mathbf{0}$$

99,99% der Vektoren in  $\text{Kern}(A - \lambda E)^m$  sind solche  $\mathbf{v}$

- \* die Anzahl der zu findenden maximalen Jordan-Ketten zu  $\lambda$  ist

$$\text{Rang}(A - \lambda E)^{m+1} - \text{Rang}(A - \lambda E)^m$$

man muss nur darauf achten, dass die zugehörigen EW unabhängig sind

- \* dann macht man in  $\text{Kern}(A - \lambda E)^{m-1}$  mit den schon gefundenen Ketten (ohne ihre Startvektoren) weiter

## 27.7.6 Tricks

**Lemma 27.18** Schiebung. Sei  $\mu \in \mathbb{C}$ . Dann gilt  $S^{-1}(A - \mu E)S = S^{-1}AS - \mu E$   
 $S^{-1}AS$  in Jordan-NF  $\Leftrightarrow S^{-1}(A - \mu E)S$  in Jordan-NF.

Beweis klar.  $\square$

**Lemma 27.19** Dreh. Sei  $\phi : V \rightarrow V$  linear. Seien  $\vec{v}_1, \dots, \vec{v}_k$  und  $\vec{w}_1, \dots, \vec{w}_l$  Jordan-Ketten von  $\phi$  zum EW  $\lambda$ ,  $l \leq k$  und  $\vec{w}_1 = r\vec{v}_1$  mit  $r \neq 0$  und  $\vec{v}_1, \dots, \vec{v}_k, \vec{w}_2, \dots, \vec{w}_l$  unabhängig, ihr Spann sei  $U$ . Dann hat  $U$  Jordan Basis bestehend aus den zwei  $\lambda$ -Jordan-Ketten

$$\vec{v}_1, \dots, \vec{v}_k, \vec{w}_2 - r\vec{w}_2, \dots, \vec{w}_l - r\vec{v}_l$$

Beweis, Die neuen Vektoren spannen  $U$  auf und sind genau so viele, also Basis. O,B,d.A  $\lambda = 0$  wegen Schiebung. Es gilt

$$\phi(\vec{w}_i - r\vec{v}_i) = \phi\vec{w}_i - r\phi\vec{v}_i = \vec{w}_{i-1} - r\vec{v}_{i-1} \quad (= \vec{0} \text{ falls } i = 2) \square$$

## 27.7.7 Beispiel: Jordan-Basis bei nur einem Eigenwert

$$A = \begin{pmatrix} 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad A^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 2 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$A^3 = O$ ,  $\text{Rang}(A^2) = 2$ .  $A^2\mathbf{e}_5 = 2\mathbf{e}_1$  und  $A^2\mathbf{e}_6 = 2\mathbf{e}_2$  unabhängig, also  $\mathbf{e}_5$  and  $\mathbf{e}_6$  geeignete Startvektoren. Zugehörige Jordan-Ketten  $A\mathbf{e}_5 = \mathbf{e}_3 + \mathbf{e}_4$  und  $A\mathbf{e}_6 = \mathbf{e}_3 - \mathbf{e}_4$ , spannen  $U$  mit Basis  $\mathbf{e}_1, \dots, \mathbf{e}_6$  auf. Basisergänzung mit Vektoren aus  $\ker A^2$  z.B.  $\mathbf{e}_7, \mathbf{v} = 2\mathbf{e}_8 - 3\mathbf{e}_5$ . Nun  $A\mathbf{e}_7 = \mathbf{e}_1 + 2\mathbf{e}_2 \in U$  aber  $A\mathbf{v} = -3\mathbf{e}_2 + \mathbf{e}_4 + 2\mathbf{e}_7 \notin U$ , also Startvektor  $\mathbf{v}$  wählen.

## 27.7.8 Beispiel: Verallgemeinerte Eigenräume

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

$$A - 3E = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix} \quad (A - 3E)^2 = \begin{pmatrix} 0 & 0 & 0 & -1 & 2 & 2 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & -2 & -2 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\text{Rang}(A - 3E) = 2, \text{Rang}(A - 3E)^2 = \text{Rang}(A - 3E)^3 = 3$$

Basis  $\beta_3$  von  $\text{Kern}(A - 3E)^2$ :  $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$

$$A - 2E = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad (A - 2E)^2 = \begin{pmatrix} 1 & 2 & 2 & 1 & 2 & 2 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\text{Rang}(A - 2E) = \text{Rang}(A - 3E)^2 = 3$$

Basis  $\beta_2$  von  $\text{Kern}(A - 2E)^2$ :

$$-\mathbf{e}_3 + \mathbf{e}_6, -\mathbf{e}_2 + \mathbf{e}_5, -\mathbf{e}_1 + \mathbf{e}_4$$

$${}^\alpha T_\beta = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}, \quad A - {}^\alpha T_\beta^{-1} A {}^\alpha T_\beta = \begin{pmatrix} 3 & 1 & 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 1 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

$$= \begin{pmatrix} A'|V_3 & O \\ O & A'|V_2 \end{pmatrix}$$

$$A'|V_3 = \begin{pmatrix} 3 & 1 & 1 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \quad A'|V_2 = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

J-Basis  $V_3$ :  $\gamma_3$ :  $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 - \mathbf{e}_2$

J-Basis:  $V_2$ :  $\gamma_2$

$$-\mathbf{e}_3 + \mathbf{e}_6, -\mathbf{e}_2 + \mathbf{e}_5, -\mathbf{e}_1 + \mathbf{e}_4 - (-\mathbf{e}_2 + \mathbf{e}_5)$$

$${}^\beta T_\gamma = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad {}^\alpha T_\gamma = {}^\alpha T_\beta {}^\beta T_\gamma = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$J = {}^\alpha T_\gamma^{-1} A {}^\alpha T_\gamma = \begin{pmatrix} 3 & 1 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$



## 27.7.10 Warnung

Ein weit verbreiteter Irrtum über die Jordan-Normalform ist, dass man zuerst eine Maximalzahl unabhängiger Eigenvektoren bestimmt und dann dazu Hauptvektoren, indem man Gleichungssystem  $(A - \lambda E)\mathbf{x} = \mathbf{v}$  löst. Dieses Vorgehen führt aber nur in den im nächsten Abschnitt angeführten Ausnahmefällen zum Erfolg. Für die Eigenvektoren einer Jordanbasis gelten nämlich gravierende Einschränkungen, exemplarisch

- Hat man zum Eigenwert  $\lambda$  Jordanketten der maximalen Länge  $m$ , so muss man aus den Eigenvektoren der Jordanbasis eine Basis von  $\text{Bild}(A - \lambda E)^{m-1}$  auswählen können. Gibt es in jeder Jordan-Basis nur eine maximale  $\lambda$ -Kette, so ist der entsprechende Eigenvektor bis auf ein skalares Vielfaches eindeutig bestimmt und in jeder Jordan-Basis enthalten.

Es folgt

- Zu einem Eigenwert  $\lambda$  gebe es Jordan-Basis-Ketten unterschiedlicher Länge. Mit welcher Wahrscheinlichkeit kann man eine nach dem Gauss-Algorithmus bestimmte Basis des Eigenraums zu einer Jordan-Basis ergänzen? Die Wahrscheinlichkeit ist 0, in Worten Null. Auf jeden Fall kleiner, als die Wahrscheinlichkeit, eine Nadel in einem Heuhaufen zu finden.

Andererseits gilt

- Ein zufällig gewählter Vektor in  $\text{Kern}(A - \lambda E)^n$  ist mit Wahrscheinlichkeit 1 Startvektor einer maximalen Jordan-Kette.

## 27.7.11 Ausnahmefälle

Die folgenden Aussagen für eine Matrix  $A \in \mathbb{C}^{n \times n}$  sind äquivalent

- (i) Das folgende Verfahren liefert stets eine Jordan-Basis von  $A$ : Man bestimmt zunächst zu jedem Eigenwert  $\lambda$  eine Basis des Eigenraums und dann zu jedem gewählten Eigenvektor  $\mathbf{v}$  eine Jordankette

$$\mathbf{v} = \mathbf{v}_1 = (A - \lambda E)\mathbf{v}_2 = (A - \lambda E)\mathbf{v}_3, \dots, \mathbf{v}_{k-1} = (A - \lambda E)\mathbf{v}_k$$

indem man nacheinander die Gleichungssysteme  $\mathbf{v}_i = (A - \lambda E)\mathbf{x}$  löst, die Lösung  $\mathbf{v}_{i+1}$  nennt und das solange fortsetzt, bis man auf ein unlösbares System stößt.

- (ii) Für jeden gegebenen Eigenwert  $\lambda$  haben die  $\lambda$ -Jordan-Ketten einer Jordan-Basis die gleiche Länge, d.h.

$$n - \text{rang}(A - \lambda E) = \text{Rang}(A - \lambda E)^{m+1} - \text{Rang}(A - \lambda E)^m \quad \text{für alle } m < n$$

(ii) ist für  $\lambda$  insbesondere dann erfüllt, wenn die algebraische Vielfachheit  $\leq 2$  oder die geometrische Vielfachheit = 1 ist.

## 28 Quadratische Formen und Hauptachsen

### 28.1 Motivation

#### 28.1.1 Extrema von Funktionen zweier Veränderlicher

Eine reelle Funktion  $f(x_1, x_2) = z$  zweier reeller Variablen  $x_1, x_2$  ordnet jedem Paar reeller Zahlen ein reelle Zahl zu. Beispiel: Jedem Koordinatenpaar von Punkten einer Karte wird die entsprechende Höhe im Gelände zugeordnet. Man kann sich dann den Graphen der Funktion als Fläche im Raum denken - die Geländeoberfläche. Zu einer vorgegebenen Höhe  $h$  hat man in der Karte die *Höhenlinie*, angegeben bzgl. eines kartesischen Koordinatensystems  $\alpha$  als

$$\{P \mid P^\alpha = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, f(x_1, x_2) = h\}$$

Ist die Funktion genügend glatt, insbesondere ohne senkrechte oder gar überhängende Abstürze und scharfe Grate, so kann man sie nach Taylor in der Nähe jedes Punktes  $(p_1, p_2)$  durch ein Polynom ausdrücken. Insbesondere in der Nähe des Ursprungs

$$f(x_1, x_2) \approx \sum_{i+j \leq N} b_{ij} x_1^i x_2^j \quad \text{für } x_1 \approx 0, x_2 \approx 0$$

Dabei ist im allgemeinen  $N = \infty$ , häufig darf man jedoch die höheren Potenzen vernachlässigen und kommt mit  $N = 1$  bzw.  $N = 2$  aus. In letzteren Falle haben wir dann

$$f(x_1, x_2) \approx b_{00} + b_{10}x_1 + b_{01}x_2 + b_{20}x_1^2 + b_{20}x_2^2 + b_{11}x_1x_2$$

Die Koeffizienten erhalten wir durch partielles Ableiten an der Stelle - hier  $(0, 0)$

$$b_{00} = f(0, 0), \quad b_{10} = \frac{\partial z}{\partial x_1}, \quad b_{01} = \frac{\partial z}{\partial x_2}, \quad b_{20} = \frac{1}{2} \frac{\partial^2 z}{\partial x_1^2}, \quad b_{02} = \frac{1}{2} \frac{\partial^2 z}{\partial x_2^2}, \quad b_{11} = \frac{\partial^2 z}{\partial x_1 \partial x_2}$$

Im Allgemeinen muss man  $x_i$  durch  $x_i - p_i$  ersetzen. Hat man in einem Punkt (o.B.d.A. dem Punkte  $(0, 0)$ ) einen (relativen) Extremwert, so muss die Tangentialebene an die Fläche in diesem Punkt horizontal sein, was sich in  $b_{10} = b_{01} = 0$  ausdrückt. Ausserdem ist der Wert von  $b_{00}$  belanglos. Also geht es um die *quadratische Form*

$$(1) \quad Q(x_1, x_2) = b_{20}x_1^2 + b_{20}x_2^2 + b_{11}x_1x_2$$

Die Punkte mit horizontaler Tangentialebene haben stets ein bestimmtes Muster von Höhenlinien und durch Wahl eines geeigneten neuen Koordinatensystems  $x'_1, x'_2$  und reeller Zahlen  $\lambda_1, \lambda_2$  kann man immer erreichen, dass das Gelände beschrieben wird durch

$$(2) \quad \tilde{Q}(x'_1, x'_2) = \lambda_1 x'^2_1 + \lambda_2 x'^2_2$$

Berggipfel	Maximum	konzentrische Ellipsen	$\lambda_1 < 0, \lambda_2 < 0$
Senke	Minimum	konzentrische Ellipsen	$\lambda_1 > 0, \lambda_2 > 0$
Sattel		Hyperbelpaar	$\lambda_1 \cdot \lambda_2 < 0$
Rinne/Rücken		parallele Geraden	ein $\lambda_i = 0$
Ebene		$\emptyset$	$\lambda_1 = \lambda_2 = 0$

Bei einer Wanderung auf einem in der Karte als Kreis um den betreffenden Punkt eingetragenen Weg, wird die maximale bzw. minimale Höhe genau dann erreicht, wenn der Weg eine dieser neuen Koordinatenachsen kreuzt: Ist  $\lambda_i \geq \lambda_j$ , so wird das Maximum bei der Achse  $x'_i$ , das Minimum bei  $x'_j$  angenommen.

### 28.1.2 Flächen zweiter Ordnung

Beispiele von Kurven zweiter Ordnung hatten wir gerade: die Höhenlinien von quadratischen Formen wie in (1). Bei den Flächen wollen wir uns auch beschränken auf die, die von quadratischen Formen herrühren und sich bezüglich eines kartesischen Koordinatensystems  $\alpha$  beschreiben lassen als

$$\{P \mid P^\alpha = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, Q(x_1, x_2, x_3) = c\}$$

$$Q(x_1, x_2, x_3) = ax_1^2 + bx_2^2 + cx_3^2 + dx_1x_2 + ex_1x_3 + fx_2x_3$$

Vorstellen kann man sich dabei Isothermen bzw. Isobaren, nicht auf einer simplen Wetterkarte, sondern mit Berücksichtigung der Höhe. Um Beispiele diskutieren zu können, nehmen wir erst mal  $c > 0$  und

$$Q(x_1, x_2, x_3) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \lambda_3 x_3^2$$

$\lambda_1 > 0$	$\lambda_2 > 0$	$\lambda_3 > 0$	Ellipsoid
$\lambda_1 > 0$	$\lambda_2 > 0$	$\lambda_3 < 0$	einschaliges Hyperboloid
$\lambda_1 > 0$	$\lambda_2 < 0$	$\lambda_3 < 0$	einschaliges Hyperboloid
$\lambda_1 > 0$	$\lambda_2 > 0$	$\lambda_3 = 0$	elliptischer Zylinder
$\lambda_1 > 0$	$\lambda_2 < 0$	$\lambda_3 = 0$	hyperbolischer Zylinder
$\lambda_1 > 0$	$\lambda_2 = 0$	$\lambda_3 = 0$	2 parallele Ebenen

Entsprechend bei Vertauschen der Indices bzw.  $c < 0$  (bei den Ellipsoiden bzw. elliptischen Zylindern wird's dann leer). Auch in diesem Falle sind die Koordinatenachsen dadurch ausgezeichnet, dass bei ihnen Maximum bzw. Minimum von  $Q(x_1, x_2, x_3)$  unter der Nebenbedingung  $x_1^2 + x_2^2 + x_3^2 = 1$  angenommen wird, d.h. wenn man  $Q$  als Funktion nur auf der Einheitskugel betrachtet. Euler hat (fast) gezeigt, dass man stets ein Koordinatensystem findet so, dass es genauso wie in den Beispielen aussieht.

### 28.1.3 Metrischer Tensor

Physiker sprechen statt von einer symmetrischen Bilinearform auch von einem *symmetrischen Tensor*. Wichtigstes Beispiel ist der *metrische Tensor*  $\langle \vec{x} \mid \vec{y} \rangle$ . Für ihn gilt

$$\langle \vec{x} \mid \vec{x} \rangle = \|\vec{x}\|^2$$

Hat  $\vec{x}$  Länge 1, so ist  $\langle \vec{x} \mid \vec{y} \rangle$  gerade die Projektion von  $\vec{y}$  auf  $\vec{x}$ , aufgefasst als Skalar. Und  $\langle \vec{x} \mid \vec{y} \rangle = 0$  genau dann, wenn  $\vec{x}$  und  $\vec{y}$  aufeinander senkrecht stehen. D.h. dieser Tensor definiert die Begriffe Länge, Orthogonalität und Winkel und kann umgekehrt schon aus

der Länge definiert werden (wobei es dann auf die Längeneinheit ankommt). Er ist durch seine *positive Definitheit* ausgezeichnet

$$\langle \vec{x} | \vec{x} \rangle > 0 \quad \text{für } \vec{x} \neq \vec{0}$$

Abstrakt gesehen ist jeder positiv definite symmetrische Tensor genauso gut wie der metrische - wenn es einem nichts ausmacht, in einer Richtung in Meter, in der zweiten in Fuss, in der dritten in Ellen zu messen.

#### 28.1.4 Trägheitstensor +

Wie Sie alle vom Karusellfahren her wissen, ergibt sich das Trägheitsmoment  $J(\vec{a})$  eines starren Körpers mit der Massendichte  $m(\vec{x})$  (wir benutzen Ortsvektoren bzgl. Ursprung  $O$ ) bzgl. einer durch  $O$  gehenden Achse  $\vec{a}$  (mit  $\|\vec{a}\| = 1$ ) als

$$J(\vec{a}) = \int m(\vec{x}) \cdot \|\vec{x} \times \vec{a}\|^2 d\vec{x}$$

dabei ist  $\|\vec{x} \times \vec{a}\|$  der Abstand des durch  $\vec{x}$  gegebenen Punktes  $P$  von der Achse  $\vec{a}$ . Statt des Integrals dürfen Sie sich eine Summation über Würfelchen  $d\vec{x}$  denken. In der Tat, denkt man sich die Masse  $M$  im Schwerpunkt  $P$  des Karusellfahrers konzentriert und als  $\vec{a}$  die Achse des Karusells, so ist  $J(\vec{a}) = M\|\vec{x} \times \vec{a}\|^2$  und das wird kleiner, wenn  $P$  näher zur Achse rückt. Da  $M$  und die Energie erhalten bleiben, muss die Winkelgeschwindigkeit zunehmen. Euler hat gezeigt, dass  $J$  in der Tat eine quadratische Form ist. Den Tensor, der hinter dem Trägheitsmoment steckt, erhält man nämlich ganz einfach als

$$\Phi(\vec{a}, \vec{b}) = \int m(\vec{x}) \cdot \langle \vec{x} \times \vec{a} | \vec{x} \times \vec{b} \rangle d\vec{x}$$

Bezüglich irgendeiner Basis  $\vec{v}_1, \vec{v}_2, \vec{v}_3$  hat man also für  $\vec{a} = \sum_i a_i \vec{v}_i$

$$J(\vec{a}) = J(\vec{v}_1)a_1^2 + J(\vec{v}_2)a_2^2 + J(\vec{v}_3)a_3^2 + 2d_{12}a_1a_2 + 2d_{13}a_1a_3 + 2d_{23}a_2a_3$$

Hat man eine Orthonormalbasis gewählt und den Ursprung im Schwerpunkt des Körpers, so kann man die  $d_{ij}$  als *Deviationsmomente* deuten: sie beschreiben die Ablenkung in Richtung auf die Achse  $\vec{v}_j$  bei Rotation um die Achse  $\vec{v}_i$ .

Die *Hauptachsen* sind nach Euler (oder J.A. von Segner) diejenigen  $\vec{a}$ , für die  $J(\vec{a})$  maximal bzw. minimal ist (cum grano salis). Euler hat gezeigt, dass es sowas gibt und dass die auch noch aufeinander senkrecht stehen, also eine Orthonormalbasis bilden. Wenn man eine Basis aus Hauptachsen zugrundelegt, gilt

$$J(\vec{a}) = J(\vec{v}_1)a_1^2 + J(\vec{v}_2)a_2^2 + J(\vec{v}_3)a_3^2 \quad \text{für } \vec{a} = \sum_i a_i \vec{v}_i$$

Die  $J(\vec{v}_i)$  heissen *Hauptträgheitsmomente*. Insbesondere tritt bei Rotation um eine Hauptachse keine Deviation auf, am stabilsten ist die Lage bei Rotation auf die Achse mit maximalem Trägheitsmoment.

## 28.2 Bilineare und quadratische Formen

Im Folgenden sei  $V$  ein endlichdimensionaler reeller Vektorraum. Seine Elemente schreiben wir als  $\vec{v}$ .

## 28.2.1 Bilinearformen

Eine *Bilinearform* auf  $V$  ist eine Abbildung  $\Phi : V \times V \rightarrow \mathbb{R}$  mit

$$\begin{aligned}\Phi(\vec{v} + \vec{w}, \vec{u}) &= \Phi(\vec{v}, \vec{u}) + \Phi(\vec{w}, \vec{u}) & \Phi(\vec{u}, \vec{v} + \vec{w}) &= \Phi(\vec{u}, \vec{v}) + \Phi(\vec{u}, \vec{w}) \\ \Phi(r\vec{v}, \vec{w}) &= r\Phi(\vec{v}, \vec{w}) & \Phi(\vec{v}, r\vec{w}) &= r\Phi(\vec{v}, \vec{w})\end{aligned}$$

Sei  $\alpha : \vec{v}_1, \dots, \vec{v}_n$  eine Basis von  $V$ . Dann gibt es eine bijektive Entsprechung zwischen Bilinearformen  $\Phi$  auf  $V$  und Matrizen  $A \in \mathbb{R}^{n \times n}$  vermöge

$$\boxed{\Phi(\vec{x}, \vec{y}) = (\vec{x}^\alpha)^t A \vec{y}^\alpha, \quad A = (\Phi(\vec{v}_i, \vec{v}_j))_{n \times n}}$$

(die transponierte Matrix zu  $X$  schreiben wir als  $X^t$ ).  $A = \Phi^\alpha$  heisst dann die (*Gram*)-*Matrix* von  $\Phi$  bzgl.  $\alpha$ . Die Form  $\Phi$  ist *symmetrisch*, d.h.

$$\Phi(x, y) = \Phi(y, x) \text{ für alle } x, y \in V \quad \Leftrightarrow \quad A^t = A$$

## 28.2.2 Quadratische Formen

Die zugehörige *quadratische Form*  $Q : V \rightarrow \mathbb{R}$  ist gegeben durch

$$\begin{aligned}Q(\vec{x}) &= \Phi(\vec{x}, \vec{x}) = (\vec{x}^\alpha)^t A \vec{x}^\alpha = \sum_{ij} a_{ij} x_i x_j & \text{mit } \vec{x}^\alpha &= (x_1, \dots, x_n)^t \\ &= \sum_i q_{ii} x_i^2 + \sum_{i < j} q_{ij} x_i x_j & \text{mit } q_{ii} &= a_{ii} \text{ und } q_{ij} = 2a_{ij}\end{aligned}$$

und muss dann, wenn man die  $q_{ij}$  und die Matrix  $A$  angeben will beachten, dass

$$a_{ij} = \begin{cases} q_{ii} & \text{falls } i = j \\ \frac{1}{2}q_{ij} & \text{falls } i < j \text{ bzw. } j < i \end{cases}$$

Und kann man  $\Phi$  aus  $Q$  zurückgewinnen:  $\Phi(\vec{x}, \vec{y}) = \frac{1}{2}(Q(\vec{x} + \vec{y}) - Q(\vec{x}) - Q(\vec{y}))$ . Eine reelle quadratische Form auf  $\mathbb{R}^n$  ist natürlich eine reelle Funktion in  $n$  Variablen und man veranschaulicht sie durch Niveau-Hyperflächen, für  $n = 2$  also durch Höhenlinien.

## 28.2.3 Transformation

Ist  $\beta$  eine weitere Basis, so werden die Koordinaten von Vektoren transformiert nach dem Ansatz

$$\vec{x}^\alpha = {}^\alpha T_\beta \vec{x}^\beta$$

wobei in den Spalten der Transformationsmatrix  ${}^\alpha T_\beta$  die Koordinaten der neuen Basisvektoren bzgl. der alten Basis  $\alpha$  stehen. Wer hier einen Isomorphismus sehen will, nehme den von  $\mathbb{R}^n \rightarrow \mathbb{R}^n$  mit  $\vec{x}^\beta \mapsto \vec{x}^\alpha$  - d.h. die Umrechnung der Koordinaten. Für die Basisbeschreibungen der Form gilt dann

$$\boxed{\Phi^\beta = {}^\alpha T_\beta^t \Phi^\alpha {}^\alpha T_\beta}$$

Beweis.  $\vec{x}^\alpha = {}^\alpha T_\beta \vec{x}^\beta$ ,  $\vec{y}^\alpha = {}^\alpha T_\beta \vec{y}^\beta$ , also

$$(\vec{x}^\beta)^t {}^\alpha T_\beta^t \Phi^\alpha {}^\alpha T_\beta \vec{y}^\beta = ({}^\alpha T_\beta \vec{x}^\beta)^t \Phi^\alpha \vec{y}^\alpha = (\vec{x}^\alpha)^t \Phi^\alpha \vec{y}^\alpha = \Phi(\vec{x}, \vec{y})$$

### 28.2.4 Definitheit

Eine quadratische Form  $Q$  heisst

$$\begin{aligned} \text{positiv definit} &\Leftrightarrow Q(\vec{x}) > 0 \quad \forall \vec{x} \neq \vec{0} & \text{positiv semi-definit} &\Leftrightarrow Q(\vec{x}) \geq 0 \quad \forall \vec{x} \\ \text{negativ definit} &\Leftrightarrow Q(\vec{x}) < 0 \quad \forall \vec{x} \neq \vec{0}, & \text{negativ semi-definit} &\Leftrightarrow Q(\vec{x}) \leq 0 \quad \forall \vec{x} \\ \text{indefinit} &\Leftrightarrow \text{sonst} \end{aligned}$$

Eine symmetrische Matrix heisst *X-definit*, falls sie bzgl. einer (und dann jeder) Basis eine X-definite Form definiert. Natürlich ist  $A$  genau dann negativ (semi) definit, wenn  $-A$  positiv (semi) definit ist.

### 28.3 Hauptachsentransformation

Sei  $V$  nun auch noch ein euklidischer Raum, d.h. mit einem Skalarprodukt ausgestattet. ON-Basis bedeutet Orthonormalbasis bzgl. dieses Skalarprodukts.

#### 28.3.1 Hauptsatz und Korollare

**Satz 28.1** *Zu jeder reellen symmetrischen Bilinearform  $\Phi$  gibt es eine ON-Basis  $\beta$  so, dass  $\Phi^\beta$  (reelle) Diagonalmatrix ist.*

Die Basisvektoren  $\vec{v}_i$  aus  $\beta$  bzw. die durch sie bestimmten Achsenrichtungen bilden dann ein *Hauptachsensystem* für  $Q$  und die Diagonaleinträge  $\lambda_i = Q(\vec{v}_i)$  die zugehörigen *Hauptmomente*. Alle zu einem Hauptmoment  $\lambda$  gehörigen Hauptachsenvektoren  $\vec{v}_i$  (d.h. mit  $\lambda_i = \lambda$ ) spannen den *Eigenraum*  $E_\lambda$  auf. Sind nun die  $z_i$  die Koordinaten bzgl.  $\beta$ , so können wir  $Q$  in der *Hauptachsenform* schreiben

$$Q(\vec{x}) = \lambda_1 z_1^2 + \dots + \lambda_n z_n^2 \quad \text{für } \vec{x} = \sum_i z_i \vec{v}_i$$

**Korollar 28.2** • *Äquivalent sind*

- $Q$  hat an  $\vec{0}$  ein (striktes) Minimum nämlich 0,
- alle Hauptmomente  $\lambda_i \geq 0$  ( $> 0$ ).
- $Q$  ist positiv semidefinit (positiv definit)

• *Äquivalent sind*

- $Q$  hat an  $\vec{0}$  ein (striktes) Maximum nämlich 0,
- alle Hauptmomente  $\lambda_i \leq 0$  ( $< 0$ ),
- $Q$  ist negativ semidefinit (negativ definit)

*In beiden strikten Fällen, und nur in diesen, hat man elliptische Niveau-Hyperflächen.*

**Korollar 28.3** *Das Maximum von  $Q$  unter der Nebenbedingung  $\|\vec{x}\| = 1$  ist der grösste Hauptmoment  $\lambda_{\max}$ , das Minimum der kleinste Hauptmoment  $\lambda_{\min}$ . Sie werden gerade auf den zugehörigen Eigenräumen angenommen.*

**Zusatz 28.4** *Die Hauptmomente und Eigenräume sind eindeutig bestimmt. Vektoren aus Eigenräumen zu verschiedenen Hauptmomenten stehen aufeinander senkrecht. ON-Basen der Eigenräume ergänzen sich zu einem Hauptachsensystem.*

## 28.3.2 Beweis

Die Korollare liest man sofort ab. Den Beweis von Satz und Zusatz führen wir über die zugehörige symmetrische Bilinearform  $\Phi$  und ihre Matrix  $A$ . Wir beginnen mit  $n = 2$ . Sei eine ON-Basis  $\alpha$  gegeben. Wir suchen ON-Basis

$$\vec{v}_1, \vec{v}_2 \quad \text{mit} \quad \Phi(\vec{v}_1, \vec{v}_2) = 0$$

Dazu der Ansatz

$$\vec{v}_1^\alpha = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \vec{v}_2^\alpha = \begin{pmatrix} x_2 \\ -x_1 \end{pmatrix}, \quad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^t A \begin{pmatrix} x_2 \\ -x_1 \end{pmatrix} = 0$$

Also

$$(a_{11} - a_{22})x_1x_2 - a_{12}x_1^2 + a_{21}x_2^2 = 0$$

$$at^2 + bt - a = 0 \quad \text{mit} \quad t = \frac{x_2}{x_1}, \quad a = a_{12} = a_{21}, \quad b = a_{11} - a_{22}$$

$$t_{1/2} = -\frac{b}{2a} \pm \sqrt{\frac{b^2}{4a^2} + 1}$$

Daraus  $x_1$  und  $x_2$  und die  $\vec{v}_i$  durch Normieren. Man kann auch gleich  $(1, t)A(-, 1)^t = 0$  ansetzen. Ganz ohne Rechnen gehts so:

$$f(t) = \Phi\left(\begin{pmatrix} \cos t \\ \sin t \end{pmatrix}, \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}\right)$$

ist stetig und

$$f(\pi) = \Phi\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix}\right) = -\Phi\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = f(0)$$

also gibt es nach dem Zwischenwertsatz  $t_0 \in [0, \pi]$  mit  $f(t_0) = 0$ . Die Hauptachsen sind dann

$$\begin{pmatrix} \cos t_0 \\ \sin t_0 \end{pmatrix}, \quad \begin{pmatrix} -\sin t_0 \\ \cos t_0 \end{pmatrix}$$

Hat man  $\lambda = \lambda_1 = \lambda_2$  so ist der Eigenraum die ganze Ebene und auch  $\lambda$  eindeutig, nämlich  $\lambda = Q(\vec{x})$  für alle  $\vec{x}$  auf dem Einheitskreis. Andernfalls sind  $\lambda_1$  und  $\lambda_2$  als maximaler bzw. minimaler Wert von  $Q$  auf dem Einheitskreis bestimmt und werden gerade an den Hauptachsen angenommen - weshalb die auch eindeutig sind.

Nun gehts weiter mit Induktion. Da die Einheitshyperkugel  $\{\vec{x} \mid \|\vec{x}\| = 1\}$  kompakt und  $\vec{v} \mapsto Q(\vec{x})$  stetig ist, gibt es mindestens ein  $\vec{v}_1$  mit  $\|\vec{v}_1\| = 1$  und  $Q(\vec{v}_1) =: \mu$  maximal. Die darauf orthogonalen Vektoren bilden einen Untervektorraum  $U = \vec{v}_1^\perp$  und für die Einschränkung von  $\Phi$  auf  $U$  können wir's schon nach Induktionsannahme. d.h. es gibt ON-Basis  $\vec{v}_2, \dots, \vec{v}_n$  von  $U$  mit

$$\Phi(\vec{v}_i, \vec{v}_j) = 0 \quad \text{für} \quad i \neq j, \quad i, j > 1$$

Für jedes  $i > 1$  haben wir aber den von  $\vec{v}_1, \vec{v}_i$  aufgespannten 2-dimensionalen Untervektorraum  $U_i$  und da haben wir's ehrlich schon gemacht, sogar mit Zusatz. Weil  $Q(\vec{v}_1)$  maximal

ist und  $\vec{v}_1, \vec{v}_i$  ON-Basis von  $U_i$  ist's ein Hauptachsensystem für die Einschränkung von  $\Phi$  auf  $U_i$ , also

$$\Phi(\vec{v}_1, \vec{v}_i) = 0 \quad \text{für } i > 1$$

Das war's für den Satz. Offensichtlich ist  $\mu$  als Maximum eindeutig bestimmt und der zugehörige Eigenraum  $E_\mu$  wird gerade von den  $\vec{v}$  mit  $\|\vec{v}\| = 1$  und  $Q(\vec{v}) = \mu$  erzeugt (und ist damit auch eindeutig bestimmt). Weil ein  $\vec{v}$ , das auch Anteile in anderen Achsenrichtungen hat, den Maximalwert nicht annehmen kann. Hat man nun irgendein Hauptachsensystem, so liegen die Achsen zu Hauptmomenten  $\lambda_i < \mu$  in dem Raum  $E_\mu^\perp$  aller zu  $E_\mu$  orthogonalen Vektoren. Der und die Einschränkung von  $\Phi$  sind eindeutig bestimmt und es folgt alles weitere mit Induktion.  $\square$

### 28.3.3 Ausartung

Die symmetrische Bilinearform  $\Phi$  bzw. die zugehörige quadratische Form  $Q$  heisst *ausgeartet*, wenn eine der folgenden äquivalenten Bedingungen gilt

- (1) 0 ist ein Hauptmoment
- (2) Es gibt  $\vec{v} \neq 0$  mit  $\Phi(\vec{v}, \vec{x}) = 0$  für alle  $\vec{x}$
- (3)  $\det \Phi^\alpha = 0$  für eine/jede Basis  $\alpha$

Beweis. 1  $\Rightarrow$  2: In einem Hauptachsensystem wähle  $\vec{v}$  zum Hauptmoment 0. 2  $\Rightarrow$  3. Bei Basistransformation  $\det S^t A S = (\det S)^2 \det A$ . 3  $\Rightarrow$  1: Für ein Hauptachsensystem  $\beta$  ist  $\det \Phi^\beta = \prod_i \lambda_i$ .  $\square$

**Korollar 28.5** *Eine nicht ausgeartete quadratische Form ist pos./neg. definit genau dann, wenn sie pos./neg. semidefinit ist.*

### 28.3.4 Klassifikation für $n = 2, 3$

Nach dem Satz über die Hauptachsentransformation ergeben sich für  $n = 2$ , neben dem uninteressanten Fall  $\lambda_1 = \lambda_2 = 0$ , die folgenden Fälle ( $\lambda_1 \geq \lambda_2$ , Niveau  $\neq 0$ )

Hauptmomente	Definitheit	Höhenlin	Fläche	offen	lok. Verh.
$\lambda_1 > 0, \lambda_2 > 0$	positiv	Ellipse	ellipt. Paraboloid	oben	Min.echt
$\lambda_1 < 0, \lambda_2 < 0$	negativ	Ellipse	ellipt. Paraboloid	unten	Max.echt
$\lambda_1 > 0, \lambda_2 < 0$	indefinit	Hyperbel	hyperbol.Parabol.		Sattelpkt
$\lambda_1 > 0, \lambda_2 = 0$	pos.semidef. ausg.	Ger.	parabol.Zylinder	oben	Minimum
$\lambda_1 < 0, \lambda_2 = 0$	neg.semidef. ausg.	Ger.	parabol.Zylinder	unten	Maximum

Für Niveau=0 hat man Punkt, 2 schneidende Geraden, bzw. Doppelgerade als Höhenlinie.

Für  $n = 3$  wurden die wesentlichen Fälle schon in 27.1.2 aufgeführt. Definitheit liegt in Falle der Ellipsoide vor.

## 28.3.5 Geometrisches Beispiel

Gegeben sei bzgl. einer ON-Basis  $\alpha : \vec{e}_1, \vec{e}_2, \vec{e}_3$  die Form

$$Q(\vec{x}) = 2x_2(x_1 + x_3) \quad \text{für } \vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3$$

Gesucht ist unter den Vektoren mit  $\|\vec{x}\| = 1$  einer mit  $Q(\vec{x})$  maximal. Bei festem  $x_2$  liegen die  $(x_1, x_3)^t$  auf einem Kreis und man hat  $x_1 + x_3$  zu maximieren. Aus Symmetriegründen muss dann  $x_1 = x_3$  sein. Das gesuchte Maximum ist also  $\max 4x_2x_1$  auf dem Kreis  $\{\vec{x} \mid \|\vec{x}\| = 1, x_1 = x_3\}$ . Wieder aus Symmetriegründen wird das Maximum bei  $x_2 = 1/\sqrt{2}$  angenommen. Ergänzung zu ON-Basis.

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \vec{v}_1^\alpha = \frac{1}{2} \begin{pmatrix} 1 \\ \sqrt{2} \\ 1 \end{pmatrix}, \quad \vec{w}_2^\alpha = \frac{1}{\sqrt{3}} \begin{pmatrix} \sqrt{2} \\ -1 \\ 0 \end{pmatrix}, \quad \vec{w}_3^\alpha = \frac{1}{\sqrt{12}} \begin{pmatrix} 1 \\ \sqrt{2} \\ -3 \end{pmatrix}$$

$$B = \begin{pmatrix} \Phi(\vec{w}_2, \vec{w}_2) & \Phi(\vec{w}_2, \vec{w}_3) \\ \Phi(\vec{w}_3, \vec{w}_2) & \Phi(\vec{w}_3, \vec{w}_3) \end{pmatrix} = \begin{pmatrix} -\frac{2\sqrt{2}}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{4\sqrt{2}}{12} \end{pmatrix}. \quad t := x'_2/x'_3, \quad t^2 - \frac{1}{\sqrt{2}}t - 1 = 0,$$

$$t_1 = \sqrt{2}, \quad t_2 = -\frac{1}{\sqrt{2}}, \quad \vec{w}_2 + t_1\vec{w}_3 = \frac{1}{\sqrt{6}} \begin{pmatrix} 3 \\ 0 \\ -3 \end{pmatrix}, \quad \vec{w}_2 + t_2\vec{w}_3 = \frac{1}{\sqrt{6}} \begin{pmatrix} \frac{3}{2} \\ -\frac{3\sqrt{2}}{2} \\ \frac{3}{2} \end{pmatrix}$$

Hauptachsenrichtungen von  $\Phi(\mathbb{R}\vec{w}_2 + \mathbb{R}\vec{w}_3)$ . Die Hauptachsen durch Normierung.

$$\vec{v}_2^\alpha = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \vec{v}_3^\alpha = \frac{1}{2} \begin{pmatrix} 1 \\ -\sqrt{2} \\ 1 \end{pmatrix}, \quad \text{EW } \lambda_i = \Phi(\vec{v}_i, \vec{v}_i), \quad \lambda_1 = \sqrt{2}, \quad \lambda_2 = 0, \quad \lambda_3 = -\sqrt{2}.$$

Die Kennfläche ist ein hyperbolischer Zylinder.

## 28.4 Definitheit

## 28.4.1 Trägheitssatz

**Satz 28.6** *Zu jeder symmetrischen Bilinearform  $\Phi$  gibt es eine Basis  $\beta$  so, dass*

$$\Phi^\beta = \begin{pmatrix} E_p & O & O \\ O & -E_q & O \\ O & O & O \end{pmatrix}$$

*Dabei ist die Signatur  $p, q$  der Form  $\Phi$  eindeutig bestimmt. (Sylvester)*

Beweis. Mit Diagonalisierung erhält man eine reelle Diagonalmatrix  $D = T^t A T$  mit Diagonaleinträgen  $d_{ii}$ . Setze  $S = T \tilde{D} P$ , wobei  $\tilde{D}$  Diagonalmatrix mit Einträgen  $\tilde{d}_{ii} = 1/\sqrt{|d_{ii}|}$  bzw. 0 ist und  $P$  passende Permutationsmatrix. Natürlich kann auch schon vorher permutiert werden.

Zur Eindeutigkeit genügt es, den Fall zu betrachten, dass  $A' = S^t A S$  und  $A$  Diagonalmatrizen sind, und die ersten  $p$  bzw.  $k$  Diagonaleinträge 1, die nächsten  $q$  bzw.  $l - 1$  und

die restlichen 0 sind.  $S^t A = A' S^{-1}$  hat offenbar Spaltenrang  $p + q$  und Zeilenrang  $k + l$ , also  $r = p + q = k + l$ . Mit der Koordinatentransformation  $\mathbf{y} = S^{-1} \mathbf{x}$  gilt

$$|x_1|^2 + \dots + |x_p|^2 - |x_{p+1}|^2 - \dots - |x_r|^2 = \mathbf{x}^t A \mathbf{x} = |y_1|^2 + \dots + |y_k|^2 - |y_{k+1}|^2 - \dots - |y_r|^2.$$

Angenommen  $p > k$ . Dann wird durch  $x_{p+1} = 0, \dots, x_n = 0$  und  $y_1 = [S^{-1} \mathbf{x}]_1 = 0, \dots, y_k = [S^{-1} \mathbf{x}]_k = 0$  ein homogenes lineares Gleichungssystem von weniger als  $n$  Gleichungen in den Variablen  $x_1, \dots, x_n$  gegeben, also hat man eine Lösung  $\mathbf{x} \neq \mathbf{0}$ . Dann folgt aber  $|x_1|^2 + \dots + |x_p|^2 = -(|y_{k+1}|^2 + \dots + |y_r|^2) \leq 0$ , und damit  $x_1 = \dots = x_p = 0$  und doch  $\mathbf{x} = \mathbf{0}$ . Den Fall  $p < k$  schliesst man durch Vertauschen der Rollen von  $A$  und  $A'$  aus.

#### 28.4.2 Zerlegung

**Satz 28.7** *Zu jeder quadratischen Form  $Q$  auf  $V$  gibt es eine orthogonale Zerlegung*

$$V = V_+ \oplus^\perp V_- \oplus^\perp V_0 \text{ und } \lambda_+ > 0 > \lambda_-$$

*d.h. jedes  $\vec{x} \in V$  hat eindeutige Darstellung  $\vec{x} = \vec{x}_+ + \vec{x}_- + \vec{x}_0$  mit  $\vec{x}_+ \in V_+$ ,  $\vec{x}_- \in V_-$  und  $\vec{x}_0 \in V_0$  und Vektoren aus zwei verschiedenen dieser Untervektorräume sind zueinander senkrecht) so, dass  $\dim V_+, \dim V_-$  die Signatur von  $Q$  ist und*

$$Q(\vec{v}) \geq \lambda_+ |\vec{v}|^2 \quad \forall \vec{v} \in V_+, \quad Q(\vec{v}) \leq \lambda_- |\vec{v}|^2 \quad \forall \vec{v} \in V_-, \quad Q(\vec{v}) = 0 \quad \forall \vec{v} \in V_0$$

Beweis. Seien  $V_0 = E_0$  und  $V_+$  bzw.  $V_-$  Summe der Eigenräume zu positiven bzw. negativen Hauptmomenten. Sei  $\lambda_+$  das kleinste positive,  $\lambda_-$  das grösste negative Hauptmoment. Sind die  $\vec{v}_i, (i \in I)$  die Hauptachsenvektoren zu den  $\lambda_i > 0$  so

$$Q(\vec{v}) = Q\left(\sum_{i \in I} x_i \vec{v}_i\right) = \sum_{i \in I} \lambda_i x_i^2 \geq \sum_{i \in I} \lambda_+ x_i^2 = \lambda_+ \sum_{i \in I} x_i^2 = \lambda_+ |\vec{v}|^2 \quad \text{für } \vec{v} \in V_+$$

Entsprechend für  $V_-$ .  $\square$

#### 28.4.3 Symmetrischer Gauss

Die Hauptmomente kann man im wirklichen Leben nur numerisch bestimmen - und auch das ist ziemlich aufwendig. Oft braucht man aber nur ihre Vorzeichen und die können wir, dank Sylvester, ablesen, wenn wir  $S^t A S$  diagonal haben mit invertierbarem  $S = {}^\alpha T_\beta$ .

**Satz 28.8** *Zu jeder reellen symmetrischen Bilinearform  $\Phi$  auf einem endlichdimensionalen  $\mathbb{R}$ -Vektorraum gibt es eine Basis  $\beta$  so, dass  $\Phi^\beta$  Diagonalmatrix ist*

$$\Phi(\vec{x}, \vec{y}) = \sum_i d_i x_i y_i$$

Zu jedem symmetrischen  $A \in \mathbb{R}^{n \times n}$  gibt es invertierbares  $S$  so, dass  $S^t A S$  diagonal

Sind alle Hauptminoren  $\det A_{\leq k} \neq 0$ , so kann man  $S$  als obere Dreiecksmatrix wählen.

Dabei ist der *Hauptminor*  $A_{\leq k}$  die durch die ersten  $k$  Zeilen und Spalten gegebene  $k \times k$ -Untermatrix von  $A$ .

Um ein  $S$  zu finden, erinnern wir uns, dass nach Gauss jede invertierbare Matrix ein Produkt von Elementarmatrizen ist (solchen die die Zeilenumformungen beschreiben). Wir machen also jeweils eine Zeilenumformung gefolgt von der analogen Spaltenumformung, um die Symmetrie zu erhalten, bis wir am Ziel sind.

$$\begin{pmatrix} 4 & -2 & 2 \\ -2 & 10 & 2 \\ 2 & 2 & 4 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 4 & 0 & 0 \\ 0 & 9 & 3 \\ 0 & 3 & 3 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 4 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

**Symmetrischer Gaussalgorithmus:** Man kombiniere das Transformationsschema mit dem Zeilen-Gaussalgorithmus zur Bestimmung einer oberen Stufenform. Man erhält symmetrische  $A_k$ , also

$$A_k = \begin{pmatrix} D_k & O \\ O & + \end{pmatrix}, \quad A_{k+1} = T_{k+1} A_k T_{k+1}^t, \quad S_0 = E, \quad S_{k+1} = S_k T_{k+1}^t$$

wobei  $D_k$  Diagonalmatrix. Sind alle Hauptminoren  $\neq 0$ , so gilt stets  $a_{k+1,k+1}^{(k)} \neq 0$ , d.h. man kann die  $T_{k+1}$  als untere und somit alle  $S_k$  als obere Dreiecksmatrizen wählen.

► Gegeben sei  $A_k = \begin{pmatrix} D_k & & O \\ O & \begin{pmatrix} b_{k+1,k+1} & \cdots & b_{k+1,n} \\ \vdots & \ddots & \vdots \\ b_{n,k+1} & \cdots & b_{nn} \end{pmatrix} & \end{pmatrix}$

- Ist  $b_{k+1,k+1} \neq 0$ , so bewirke durch Zeilenscherungen, dass die neue  $k+1$ -te Zeile von 'A' ausserhalb der Diagonalen nur noch Nullen enthält. Ändere die Werte über der Diagonalen von 'A' so, dass wieder eine symmetrische Matrix entsteht
- Ist  $b_{k+1,k+1} = 0$  so suche vorher  $i > k+1$  mit  $b_{i,i} \neq 0$  und vertausche  $k+1$ -te und  $i$ -te Zeile in 'A', ebenso für Spalten von 'A'
- Ist  $b_{ii} = 0$  für alle  $i > k$ , so suche vorher  $j > i > k$  mit  $b_{ij} \neq 0$ . Addiere in 'A' die  $j$ -Zeile zur  $i$ -ten und die  $j$ -te Spalte zur  $i$ -ten. Das ergibt  $i$ -ten Diagonaleintrag  $b_{ij} + b_{ij} \neq 0$ .
- Sind alle  $b_{ij} = 0$  für  $j > i > k$ , so rufe 'Gott sei Dank'

## 28.5 Hauptminorenkriterium

**Satz 28.9 Hurwitz.** Für eine symmetrische Matrix  $A$  sind gleichwertig:

- $A$  ist positiv definit.
- Es gibt eine invertierbare Matrix  $W$  mit  $A = WW^t$ .
- Die Hauptminoren  $A_{\leq k} = (a_{ij})_{1 \leq i,j \leq k}$  von  $A$  haben  $\det A_{\leq k} > 0$ .

Beweis. Die Äquivalenz von (1) und (2) ist klar nach dem Trägheitssatz. Also ist für positiv definites  $A$  die Determinante  $\det A = \det(W^t W) = \det W^t \det W = |\det W|^2 > 0$ . Andererseits sind für positiv definites  $A$  alle Hauptminoren  $A_{\leq k}$  positiv definit: für die durch  $A_{\leq k}$  definierte quadratische Form  $Q_{\leq k}$  gilt:  $Q_{\leq k}(x_1, \dots, x_k) = Q(x_1, \dots, x_k, 0, \dots, 0) > 0$  falls ein  $x_i \neq 0$ . Also  $\det A_{\leq k} > 0$  für alle  $k$ .

Sei nun (3) vorausgesetzt, insbesondere  $a_{11} \neq 0$ . Also hat man im symmetrischen Gaußalgorithmus sofort  $A \rightsquigarrow B = T A T^t$  und es bleibt nur zu zeigen, dass  $B$  wieder die Voraussetzung erfüllt. Man beachte, dass für die Hauptminoren gilt  $B_{\leq k} = T_{\leq k} A_{\leq k} T_{\leq k}^t$  d.h. sie haben Determinante  $|\det T_{\leq k}|^2 \det A_{\leq k} > 0$ . Ist  $C$  der  $(n-1) \times (n-1)$ -Minor  $C = (b_{ij})_{1 < i, j \leq k}$  so ist  $C_{\leq k}$  Minor von  $B_{\leq k+1}$  und es gilt  $\det(B_{\leq k+1}) = a_{11} \det C_{\leq k}$ . Also hat  $C$  nur positive Hauptminoren und das Verfahren führt zu einer Diagonalmatrix mit nur positiven Einträgen, also schliesslich mit dem Produkt  $S$  der Matrizen aus den Einzelschritten zu  $S A S^t = E_n$  und  $A = W W^t$  mit  $W = S^{-1}$ .

## 28.6 Verwechslung

Für Matrizen  $A \in \mathbb{R}^{n \times n}$  haben wir zwei grundsätzlich unterschiedliche Bedeutungen

- Ist  $V$  ein  $n$ -dimensionaler Vektorraum mit Basis  $\alpha$ , so gibt es eine 1-1-Entsprechung zwischen Matrizen  $A$  und linearen Abbildungen  $\phi : V \rightarrow V$  gegeben durch  $\phi(\vec{x})^\alpha = A \vec{x}^\alpha$ . Die Transformation erfolgt nach

$$A' = {}^\beta T_\alpha A {}^\alpha T_\beta = S^{-1} A S \text{ mit } S = {}^\alpha T_\beta$$

- Ist  $V$  ein  $n$ -dimensionaler Vektorraum mit Basis  $\alpha$ , so gibt es eine 1-1-Entsprechung zwischen Matrizen  $A$  und Bilinearformen  $\Phi : V \rightarrow \mathbb{R}$  gegeben durch  $\Phi(\vec{x}, \vec{y}) = (\vec{x}^\alpha)^t A \vec{y}^\alpha$ . Die Transformation erfolgt nach

$$A' = {}^\alpha T_\beta^t A {}^\alpha T_\beta = S^t A S \text{ mit } S = {}^\alpha T_\beta$$

Hier ist  $\Phi$  symmetrisch genau dann, wenn  $A$  symmetrisch ist. In diesem Falle betrachtet man die quadratische Form  $Q(\vec{x}) = (\vec{x}^\alpha)^t A \vec{x}^\alpha$ .

Betrachtet man nur symmetrische Matrizen und lässt man nur orthogonale Transformationen zu, so erweisen sich die beide Konzepte als äquivalent (weshalb man auch bei quadratischen Formen von Eigenwerten und Eigenvektoren sprechen darf)

**Zusatz 28.10** Sei  $A \in \mathbb{R}^{n \times n}$  symmetrisch,  $\alpha$  und  $\beta : \vec{v}_1, \dots, \vec{v}_n$  eine ON-Basen des euklidischen Vektorraums  $V$ , und  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ . Dann sind äquivalent

- Für die durch  $A$  bzgl.  $\alpha$  gegebene quadratische Form  $Q$  auf  $V$  ist  $\beta$  Hauptachsensystem mit Hauptmomenten  $\lambda_i = Q(\vec{v}_i)$
- Für die durch  $A$  bzgl.  $\alpha$  definierten linearen Abbildung  $\phi : V \rightarrow V$ ,  $\phi(\vec{x})^\alpha = A \vec{x}^\alpha$ , ist  $\beta$  ist eine Basis von Eigenvektoren und die  $\lambda_i$  sind Eigenwerte:  $\phi(\vec{v}_i) = \lambda_i \vec{v}_i$

Beweis.  ${}^\alpha T_\beta$  ist orthogonal, also  ${}^\alpha T_\beta^t = {}^\alpha T_\beta^{-1}$  und es geht beidesmal um dieselbe Aufgabe für Matrizen.  $\square$

Die linearen Abbildungen  $\phi : V \rightarrow V$ , die bzgl. einer/jeder ON-Basis eine symmetrische Matrix haben, heissen *selbstadjungiert* und sind durch die folgende Eigenschaft charakterisiert

$$\langle \phi(\vec{x}) | \vec{y} \rangle = \langle \vec{x} | \phi(\vec{y}) \rangle \quad \text{für alle } \vec{x}, \vec{y} \in V$$

## 28.7 Quadriken

## 28.7.1 Quadratische Gleichungen

In einem  $n$ -dimensionalen euklidischen Vektorraum  $V$  sei ein kartesisches Koordinatensystem  $\alpha$  mit Ursprung  $O$  gegeben. Sei  $A \in \mathbb{R}^{n \times n}$ ,  $\mathbf{b} \in \mathbb{R}^n$  und  $c \in \mathbb{R}$ .

$$\mathcal{Q} = \{\vec{x} + O \mid \vec{x}^\alpha = \mathbf{x}, \mathbf{x}^t A \mathbf{x} + \mathbf{b}^t \vec{x} + c = 0\}$$

ist dann durch eine *quadratische Gleichung* definiert oder auch eine *quadratische Menge*. Liegt sie nicht ganz in einem echten affinen Teilraum, so ist sie eine *Quadrik*. Beschreibung und Transformation vereinfachen sich, wenn wir setzen

$$\tilde{\mathbf{x}} = \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^{n+1}, \quad \tilde{A} = \begin{pmatrix} c & \frac{1}{2}\mathbf{b}^t \\ \frac{1}{2}\mathbf{b} & A \end{pmatrix} \in \mathbb{R}^{(n+1) \times (n+1)}$$

Nämlich

$$\mathcal{Q} = \{\vec{x} + O \mid \tilde{\mathbf{x}}^t \tilde{A} \tilde{\mathbf{x}} = 0\}$$

Man spricht hier von *homogenen Koordinaten* - statt den nullten Eintrag auf 1 zu normieren, könnte man  $\tilde{\mathbf{x}}$  auch mit beliebigen Skalaren  $\neq 0$  multiplizieren.

## 28.7.2 Transformation

Wie bei den quadratischen Formen erlauben wir Transformationen  $S \mapsto S^t \tilde{A} S$  aber hier nur mit "affinen Matrizen"

$$S = \begin{pmatrix} 1 & O \\ \mathbf{t} & T \end{pmatrix}, \quad T \in \mathbb{R}^{n \times n} \text{ orthogonal}, \quad \mathbf{t} \in \mathbb{R}^n$$

Ist  $\mathbf{t} = \mathbf{0}$  so handelt es sich um eine Basistransformation und wir erhalten

$$S^t \tilde{A} S = \begin{pmatrix} c & \frac{1}{2}\mathbf{b}^t T \\ T^t \frac{1}{2}\mathbf{b} & T^t A T \end{pmatrix}$$

und bzgl. des neuen Koordinatensystems mit dem alten Ursprung  $O$  und der neuen Basis  $\beta$  mit  $T = {}^\alpha T_\beta$  haben wir

$$\mathcal{Q} = \{\vec{y} + O \mid \tilde{\mathbf{y}}^t S^t \tilde{A} S \tilde{\mathbf{y}} = 0\}$$

Ist  $T = E$  so haben wir das neue Koordinatensystem  $\beta$  mit der alten Basis aber neuem Ursprung  $\vec{t} + O$ , wobei  $\mathbf{t} = \vec{t}^\alpha$  und

$$S^t \tilde{A} S = \begin{pmatrix} c + \mathbf{b}^t \mathbf{t} + \mathbf{t}^t A \mathbf{t} & \frac{1}{2}\mathbf{b}^t + \mathbf{t}^t A \\ \frac{1}{2}\mathbf{b} + A \mathbf{t} & A \end{pmatrix}$$

$$\mathcal{Q} = \{\vec{y} + \vec{t} + O \mid \tilde{\mathbf{y}}^t S^t \tilde{A} S \tilde{\mathbf{y}} = 0\}$$

Im allgemeinen Fall gibt  $\vec{t} + O$  den neuen Ursprung an an die neue Basis  $\beta$  ist durch  $T = {}^\alpha T_\beta$  bestimmt.

Wir versuchen nun, durch Transformationen von  $\tilde{A}$  zu einer einfacheren Matrix zu kommen. Man kann einfacher transformieren, wenn man beliebige invertierbare  $T$  erlaubt. Am Ergebnis kann man dann immer noch erkennen, von welchem Typ die Quadrik ist - das entspricht dem Satz von Sylvester, ist aber etwas aufwendiger zu beweisen.

## 28.7.3 Rezept

1.  $\mathbf{t} = \mathbf{0}$  und  $T$  orthogonal so, dass  $T^{-1}AT$  diagonal, und Nullen nur am Ende der Diagonalen (Hauptachsentransformation, EW  $\lambda_i \neq 0$  für  $i \leq k$ ,  $\lambda_i = 0$  für  $i > k$ ). Das ergibt (und wir können auch  $c_n \neq 0$  erreichen, falls ein  $c_i \neq 0$  für  $i > k$ )

$$S_1 = \begin{pmatrix} 1 & \mathbf{0}^t \\ \mathbf{0} & T \end{pmatrix}, \quad \tilde{A}_1 = \begin{pmatrix} c & \frac{1}{2}c_1 & \dots & \frac{1}{2}c_k & \frac{1}{2}c_{k+1} & \dots & \frac{1}{2}c_n \\ \frac{1}{2}c_1 & \lambda_1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & 0 & \ddots & 0 & 0 & \dots & 0 \\ \frac{1}{2}c_k & 0 & \dots & \lambda_k & 0 & \dots & 0 \\ \frac{1}{2}c_{k+1} & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & 0 & \dots & 0 & 0 & \dots & 0 \\ \frac{1}{2}c_n & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}$$

2. Wähle nun

$$\mathbf{t} = \begin{pmatrix} t_1 \\ \vdots \\ t_k \\ \mathbf{0} \end{pmatrix} \quad \text{mit} \quad \frac{1}{2}c_i + \lambda_i t_i = 0 \quad \text{für} \quad i \leq k, \quad S_2 = \begin{pmatrix} 1 & \mathbf{t}^t \\ \mathbf{t} & E \end{pmatrix}$$

Das ergibt

$$\tilde{A}_2 = \begin{pmatrix} c' & 0 & \dots & 0 & \frac{1}{2}c_{k+1} & \dots & \frac{1}{2}c_n \\ 0 & \lambda_1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & 0 & \ddots & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & \lambda_k & 0 & \dots & 0 \\ \frac{1}{2}c_{k+1} & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & 0 & \dots & 0 & 0 & \dots & 0 \\ \frac{1}{2}c_n & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}$$

Ist  $k = n$  oder  $c_i = 0$  für alle  $i > k$ , so ist's hier zu Ende.

3. Wähle  $U$  orthogonal mit (ist  $c_i = 0$  für  $k < i < n$ , so ist dieser Schritt trivial)

$$U^t \begin{pmatrix} \frac{1}{2}c_{k+1} \\ \vdots \\ \frac{1}{2}c_{n-1} \\ \frac{1}{2}c_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \frac{1}{2}d \end{pmatrix}, \quad T = \begin{pmatrix} E_k & O \\ O & U \end{pmatrix}, \quad S_3 = \begin{pmatrix} 1 & \mathbf{0}^t \\ \mathbf{0} & T \end{pmatrix}$$

Das ergibt

$$\tilde{A}_3 = \begin{pmatrix} c' & 0 & \dots & 0 & 0 & \dots & \frac{1}{2}d \\ 0 & \lambda_1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & 0 & \ddots & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & \lambda_k & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & 0 & \dots & 0 & 0 & \dots & 0 \\ \frac{1}{2}d & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}$$

4. Wenn man bis hier kommt, ist  $d \neq 0$ . Wähle  $t_n$  mit

$$c' + dt_n = 0, \quad \mathbf{t} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ t_n \end{pmatrix}, \quad S_4 = \begin{pmatrix} 1 & O \\ \mathbf{t} & E \end{pmatrix}$$

Das ergibt

$$\tilde{A}_4 = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 & \dots & \frac{1}{2}d \\ 0 & \lambda_1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & 0 & \ddots & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & \lambda_k & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & 0 & \dots & 0 & 0 & \dots & 0 \\ \frac{1}{2}d & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}$$

#### 28.7.4 Beispiel

$$x_1^2 + x_2^2 + 2x_1x_2 - 6x_1 - 5x_2 - \frac{1}{\sqrt{2}}x_3 + \frac{89}{16} = 0$$

$$\tilde{A} = \begin{pmatrix} \frac{89}{16} & -3 & -\frac{5}{2} & -\frac{1}{2\sqrt{2}} \\ -3 & 1 & 1 & 0 \\ -\frac{5}{2} & 1 & 1 & 0 \\ -\frac{1}{2\sqrt{2}} & 0 & 0 & 0 \end{pmatrix}, \quad S_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\tilde{A}_1 = S_1^t \tilde{A} S_1 = \begin{pmatrix} \frac{89}{16} & -\frac{11}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} \\ -\frac{11}{2\sqrt{2}} & 2 & 0 & 0 \\ \frac{1}{2\sqrt{2}} & 0 & 0 & 0 \\ \frac{-1}{2\sqrt{2}} & 0 & 0 & 0 \end{pmatrix}, \quad S_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{11}{4\sqrt{2}} & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\tilde{A}_2 = S_2^t \tilde{A}_1 S_2 = \begin{pmatrix} -2 & 0 & \frac{1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} \\ 0 & 2 & 0 & 0 \\ \frac{1}{2\sqrt{2}} & 0 & 0 & 0 \\ \frac{-1}{2\sqrt{2}} & 0 & 0 & 0 \end{pmatrix}, \quad S_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

$$\tilde{A}_3 = S_3^t \tilde{A}_2 S_3 = \begin{pmatrix} -2 & 0 & 0 & \frac{-1}{2} \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -\frac{1}{2} & 0 & 0 & 0 \end{pmatrix}, \quad S_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}$$

$$\tilde{A}_4 = S_4^t \tilde{A}_3 S_4 = \begin{pmatrix} 0 & 0 & 0 & \frac{-1}{2} \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -\frac{1}{2} & 0 & 0 & 0 \end{pmatrix}, \quad S = S_1 S_2 S_3 S_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{3}{8} & \frac{1}{\sqrt{2}} & \frac{-1}{2} & \frac{1}{2} \\ \frac{19}{8} & \frac{1}{\sqrt{2}} & \frac{1}{2} & \frac{-1}{2} \\ -\sqrt{2} & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

$$\tilde{A}_4 = S^t \tilde{A} S, \quad 2y_1^2 - y_3 = 0$$

## 28.7.5 Normalform

**Theorem 28.11** Zu jeder bzgl. eines kartesischen Koordinatensystems  $\alpha$  gegebenen quadratischen Menge  $\mathcal{Q}$  gibt es ein kartesisches Koordinatensystem  $\delta$ , bzgl. dessen  $\mathcal{Q}$  in einer der folgenden Formen beschrieben wird (bis auf Multiplikation mit einer Konstanten) wobei  $0 < a_i \leq a_j$  für  $i < j \leq p$  resp.  $p < i < j \leq p+q$ . Zusätzlich  $a_{p+q} = 1$  in (2).

$$(1) \mathcal{Q} = \{P \in \mathbb{P} \mid \mathbf{z} = P^\delta, \frac{z_1^2}{a_1^2} + \dots + \frac{z_p^2}{a_p^2} - \frac{z_{p+1}^2}{a_{p+1}^2} - \dots - \frac{z_{p+q}^2}{a_{p+q}^2} = 1\} \text{ mit } p \geq 1$$

$$(2) \mathcal{Q} = \{P \in \mathbb{P} \mid \mathbf{z} = P^\delta, \frac{z_1^2}{a_1^2} + \dots + \frac{z_p^2}{a_p^2} - \frac{z_{p+1}^2}{a_{p+1}^2} - \dots - \frac{z_{p+q}^2}{a_{p+q}^2} = 0\} \text{ mit } p \geq q \geq 1$$

$$(3) \mathcal{Q} = \{P \in \mathbb{P} \mid \mathbf{z} = P^\delta, z_n = \frac{z_1^2}{a_1^2} + \dots + \frac{z_p^2}{a_p^2} - \frac{z_{p+1}^2}{a_{p+1}^2} - \dots - \frac{z_{p+q}^2}{a_{p+q}^2}\} \text{ mit } p \geq q \text{ und } 1 \leq p+q$$

## 28.7.6 Mittelpunkte +

Ein Punkt  $P$  von  $\mathcal{Q}$  ist ein *Mittelpunkt* von  $\mathcal{Q}$  falls

$$\vec{x} + P \in \mathcal{Q} \Rightarrow -\vec{x} + P \in \mathcal{Q} \quad \text{für alle } \vec{x} \in V$$

**Korollar 28.12**  $P$  ist Mittelpunkt von  $\mathcal{Q}$  genau dann, wenn

$$A\mathbf{p} = -\frac{1}{2}\mathbf{b} \quad \text{wobei } \mathbf{p} = \vec{p}^\alpha, P = \vec{p} + O_\alpha$$

*Beweis.* Transformiere in den Ursprung  $P$

$$\begin{pmatrix} 1 & \mathbf{x}^t \\ \mathbf{0} & E \end{pmatrix} \begin{pmatrix} c & \frac{1}{2}\mathbf{b}^t \\ \frac{1}{2}\mathbf{b} & A \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0}^t \\ \mathbf{x} & E \end{pmatrix} = \begin{pmatrix} c' & \frac{1}{2}\mathbf{b}^t + \mathbf{x}^t A \\ \frac{1}{2}\mathbf{b} + A\mathbf{x} & A \end{pmatrix}, \quad c' = c + \mathbf{b}^t \mathbf{x} + \mathbf{x}^t A \mathbf{x}$$

Also  $P$  Mittelpunkt genau dann, wenn  $\frac{1}{2}\mathbf{b} + A\mathbf{p} = \mathbf{0}$ .  $\square$

## 28.8 Zur Numerik der symmetrischer Matrizen +

## 28.8.1 Cholesky-Zerlegung

Die Matrix  $W$  aus obigem Satz kann man sogar als untere Dreiecksmatrix wählen, da der Gaußalgorithmus hier ohne Vertauschungen auskommt. Daher kann man aus dem Ansatz

$$\boxed{R^t R = A \quad R \text{ obere Dreiecksmatrix}}$$

die Zeilen von  $R$  der Reihe nach berechnen - genau dann, wenn  $A$  positiv definit

$$\begin{aligned} \blacktriangleright \quad r_{kk} &= \sqrt{a_{kk} - (|r_{1k}|^2 + \dots + |r_{k-1,k}|^2)} \\ \blacktriangleright \quad r_{kj} &= \frac{1}{r_{kk}} [a_{kj} - (r_{k1}r_{1j} + \dots + r_{k,k-1}r_{k-1,j})] \end{aligned}$$

Da Dreiecksmatrizen leicht zu invertieren sind, hat man so ein effektives Verfahren zur Lösung linearer Gleichungssysteme mit positiv definiten Matrix.

$$\begin{pmatrix} 4 & -2 & 2 \\ -2 & 10 & 2 \\ 2 & 2 & 4 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ & 0 & 0 \\ & & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}$$

$$\begin{aligned}
&= \begin{pmatrix} 2 & 0 & 0 \\ -1 & & 0 \\ 1 & & \end{pmatrix} \begin{pmatrix} 2 & -1 & 1 \\ 0 & & \\ 0 & 0 & \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ -1 & 3 & 0 \\ 1 & & \end{pmatrix} \begin{pmatrix} 2 & -1 & 1 \\ 0 & 3 & \\ 0 & 0 & \end{pmatrix} \\
&= \begin{pmatrix} 2 & 0 & 0 \\ -1 & 3 & 0 \\ 1 & 1 & \end{pmatrix} \begin{pmatrix} 2 & -1 & 1 \\ 0 & 3 & 1 \\ 0 & 0 & \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ -1 & 3 & 0 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & -1 & 1 \\ 0 & 3 & 1 \\ 0 & 0 & 2 \end{pmatrix}
\end{aligned}$$

### 28.8.2 Jacobiverfahren

Ursprung vieler Verfahren zur Bestimmung der EW einer symmetrischen Matrix  $A$  ist das folgende Verfahren von Jacobi. Sei  $\sigma A$  die Quadratsumme der  $n(n-1)/2$  Einträge oberhalb der Diagonalen:  $\sigma A = \sum_{i < j} a_{ij}^2$ . Dass  $A$  Diagonalmatrix ist, bedeutet gerade  $\sigma A = 0$ , und dem wollen wir nun durch ein Iterationsverfahren immer näher kommen.

**Lemma 28.13** *Zu jeder symmetrischen  $n \times n$ -Matrix  $A_k$  gibt es eine orthogonale Matrix  $S_{k+1}$ , sodass*

$$\sigma A_{k+1} \leq q \sigma A_k \text{ für } A_{k+1} = S_{k+1}^t A_k S_{k+1} \text{ und } q = 1 - \frac{2}{n(n-1)}.$$

Sei nun  $A_0 = A$  und die  $A_k, S_k$  induktiv nach dem Lemma bestimmt. Dann gilt  $\lim_{k \rightarrow \infty} \sigma A_k = 0$ . Die Produktmatrizen  $U_k = S_1 \cdot \dots \cdot S_k$  sind ebenfalls orthogonal. Fasst man  $n \times n$ -Matrizen als  $n^2$ -Tupel auf, so bilden die orthogonalen Matrizen eine beschränkte und abgeschlossene Teilmenge von  $\mathbb{R}^{n^2}$  bestehend aus den  $(u_{11}, u_{12}, \dots, u_{nn})$ , die den Gleichungen  $\sum_i u_{ij}^2 = 1$  und  $\sum_i u_{ij} u_{ik} = 0$  für  $j \neq k$  genügen - insbesondere  $|u_{ij}| \leq 1$ . Also hat die Folge  $U_k$  eine konvergente Teilfolge  $U_{k_l}$  mit Limes  $U$  - und  $U$  ist auch orthogonal.

$$\sigma(U^t A U) = \lim_{l \rightarrow \infty} \sigma(S_{k_l}^t \cdot \dots \cdot S_1^t A S_1 \cdot \dots \cdot S_{k_l}) \leq \lim_{l \rightarrow \infty} q^{k_l} \sigma A = 0$$

und somit ist  $U^t A U$  Diagonalmatrix. Das Dumme ist: Nur der oberste Chef von Herrn Bolzano weiss, welche  $U_k$  zu der konvergenten Teilfolge gehören. Nur wenn man  $n$  verschiedene EW hat, konvergieren die  $U_k$  gegen 'das' Hauptachsensystem.

Beweis des Lemmas. Zur Vereinfachung der Notation betrachten wir  $A$  und nehmen an, dass  $a_{12}^2$  das grösste der  $a_{ij}^2$ ,  $i < j$ , ist (andernfalls muss man die Basis umordnen, d.h. mit einer Permutationsmatrix transformieren). Über den Fall  $n=2$  finden wir eine Drehmatrix  $D_\phi$ , sodass

$$D_\phi^t \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} D_\phi$$

Diagonalmatrix ist und setzen

$$S = \begin{pmatrix} D_\phi & 0 \\ 0 & E_{n-2} \end{pmatrix}$$

Für  $A' = S^t A S$  haben wir  $a'_{12} = 0$ ,  $a'_{ij} = a_{ij}$  für  $2 < i \leq j$  und

$$a'_{1j} = a_{1j} \cos \phi + a_{2j} \sin \phi, \quad a'_{2j} = -a_{1j} \sin \phi + a_{2j} \cos \phi$$

und damit  $\sigma A' = \sigma A - a_{12}^2 \leq q \sigma A$ .

Man kann zeigen, dass die Diagonalelemente  $a_{ii}^{(k)}$  von  $A_k$  um weniger als  $\sqrt{q^k \sigma A}$  von den in passender Reihenfolge genommenen Hauptmomenten  $\lambda_1, \dots, \lambda_n$  abweichen. Dieselbe Schranke gilt für  $|\mathbf{A}\mathbf{u}_j^k - \lambda_j \mathbf{u}_j^k|$ , d.h. die Spalten von  $U_k$  verhalten sich im Rahmen der Rechengenauigkeit wie Eigenvektoren, obwohl die Abweichung (gemessen an den Winkeln) von einer tatsächlichen Basis von Eigenvektoren absolut gross sein kann. Beispiel.

$$A = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & 0 \end{pmatrix}.$$

$$\begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix} \text{ hat EW/Hauptachsen } -\frac{1}{2} : \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \text{ und } \frac{1}{2} : \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

$$S_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad A_1 = S_1^t A S_1 = \begin{pmatrix} -\frac{1}{2} & 0 & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{2} & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 \end{pmatrix}.$$

$$\begin{pmatrix} -\frac{1}{2} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 \end{pmatrix} \text{ EW/Hauptachsen } -1 : \frac{1}{\sqrt{3}} \begin{pmatrix} -\sqrt{2} \\ 1 \end{pmatrix} \text{ und } \frac{1}{2} : \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ \sqrt{2} \end{pmatrix}.$$

$$S_2 = \begin{pmatrix} \frac{\sqrt{2}}{\sqrt{3}} & 0 & \frac{1}{\sqrt{3}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{3}} & 0 & \frac{\sqrt{2}}{\sqrt{3}} \end{pmatrix}, \quad A_2 = S_2^t A_1 S_2 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}, \quad U = S_1 S_2 = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix}.$$

Dass man hier nach endlich vielen Schritten fertig ist, war ein Glücksfall. Es gibt jedoch eine wichtige Abänderung des Verfahrens (nach Givens), in der man  $A$  durch endlich viele Drehungen auf Tridiagonalgestalt transformiert. Von da kann man auf das charakteristische Polynom kommen, und die Nullstellen mit Näherungsverfahren berechnen.

### 28.8.3 Matrixphilosophie

Der Gaussalgorithmus und seine gesternte Variante haben die erfreuliche Eigenschaft, innerhalb jeden Körpers, z.B. dem der rationalen Zahlen, ausführbar zu sein. Demgegenüber beruhen Hauptachsentransformation für symmetrische reelle Matrizen ebenso wie die Jordansche Normalform wesentlich auf den Abschlusseigenschaften von  $\mathbb{R}$  bzw.  $\mathbb{C}$ . Ein direkter Algorithmus wie Gauss (auch wenn man noch solche Schritte wie das Normieren erlaubt) ist ohne vorherige Bestimmung der Eigenwerte nicht möglich, da man sonst ein allgemeines Verfahren zur Nullstellenbestimmung rationaler Polynome in einer Variablen mithilfe der arithmetischen Operationen und Wurzeln hätte, was aber nach Galois und Abel nicht sein kann. Aber auch bei den meisten praktizierten Näherungsverfahren für die Transformationen beruht der Nachweis der Konvergenz auf dem abstrakten Beweis (bzw. der Voraussetzung) der Existenz der Transformation. Daher ist es sinnvoll, begrifflich einfache Existenzbeweise zu studieren und sie durch Übungsalgorithmen besser zu verstehen.

## 28.9 Schneller Beweis der Trägheit

Angenommen

$$U^t A U = A' = \begin{pmatrix} E_p & O & O \\ O & -E_q & O \\ O & O & O \end{pmatrix}, \quad V^t A V = A'' = \begin{pmatrix} E_k & O & O \\ O & -E_l & O \\ O & O & O \end{pmatrix}$$

mit invertierbaren  $U, V$ . Dann  $A' = S^t A'' S$  mit  $S = V^{-1}U$ . Also o.B.d.A.  $A = A''$  und  $p \leq k$ . Multiplikation mit invertierbaren Matrizen ändert den Rang nicht. Somit  $r = p+q = k+l$ . Seien  $B, B', T$  die  $r$ -Minoren von  $A$  bzw.  $A'$  bzw.  $S$ . Dann  $B' = T^t B T$ . Daher o.B.d.A.  $r = n$ . Sei  $Q$  die durch  $A$  bzgl. der kanonischen Basis  $\alpha$  definierte quadratische Form und  $\beta : \mathbf{v}_1, \dots, \mathbf{v}_n$  die Basis mit  $S = {}^\alpha T_\beta$ . Setze

$$U_1 = \text{Spann}\{\mathbf{e}_1, \dots, \mathbf{e}_k\}, \quad V_2 = \text{Spann}\{\mathbf{v}_{p+1}, \dots, \mathbf{v}_n\}$$

Dann ist  $Q$  auf  $U_1$  positiv definit, auf  $V_2$  negativ definit. Es folgt

$$U_1 \cap V_2 = \{\mathbf{0}\}$$

und es ist  $\mathbf{e}_1, \dots, \mathbf{e}_k, \mathbf{v}_{p+1}, \dots, \mathbf{v}_n$  linear unabhängig. Daher  $k \leq p$  und somit  $p = k$ .  $\square$