

# 4 Spektrum einer Matrix bzw. einer linearen Funktion

## 4.1 Determinanten

Im folgenden betrachten wir quadratische Matrizen. Wir schreiben dabei eine  $n \times n$  Matrix  $A$  (über dem Körper  $K$ ) primär als Zeilenvektor, dessen Elemente die Spalten von  $A$  sind; also

$$A = (a_1, \dots, a_n),$$

wobei  $a_i$  die  $i$ -te Spalte der Matrix  $A$  ist. Für solche quadratische Matrizen  $A$  werden wir eine Abbildung in den Grundkörper  $K$  angeben, die Determinante von  $A$ , welche ein Kriterium für die Singularität (also Nichtinvertierbarkeit) von  $A$  darstellt.

**Definition 4.1.1:** Sei  $D$  eine Abbildung von quadratischen Matrizen der Dimension  $n$  in den Grundkörper, also

$$D : \text{Mat}_{n \times n}(K) \longrightarrow K .$$

Wir fassen  $D$  auf als Funktion der Spalten der Matrix  $A$ , also  $D(A) = D(a_1, \dots, a_n)$ . Im folgenden sind  $i, j$  immer Spaltenindizes, also  $i, j \in \{1, \dots, n\}$ .

$D$  ist **multilinear**, g.d.w. für alle  $i, j$  und alle  $\lambda \in K$  gilt:

$$(D1) \quad D(\dots, b_i + c_i, \dots) = D(\dots, b_i, \dots) + D(\dots, c_i, \dots), \text{ und}$$

$$(D2) \quad D(\dots, \lambda a_i, \dots) = \lambda D(\dots, a_i, \dots).$$

$D$  ist **alternierend** g.d.w. für alle  $i \neq j$  gilt:

$$(D3) \quad D(\dots, a_i, \dots, a_j, \dots) = -D(\dots, a_j, \dots, a_i, \dots).$$

$D$  ist **identitätserkennend** g.d.w.

$$(D3') \quad D(A) = 0 \text{ falls } A \text{ zwei identische Spalten hat.}$$

$D$  ist **1-erhaltend** g.d.w.

$$(D4) \quad D(I_n) = 1.$$

$D$  heisst eine **Determinantenfunktion** g.d.w.  $D$  multilinear, alternierend und 1-erhaltend ist; also wenn  $D$  die Eigenschaften (D1), (D2), (D3) und (D4) besitzt.

**Satz 4.1.2:** Eine Abbildung  $D : \text{Mat}_{n \times n}(K) \rightarrow K$  ist eine Determinantenfunktion g.d.w. sie die Eigenschaften (D1), (D2), (D3') und (D4) besitzt.

*Beweis:* " $\Rightarrow$ ": hat  $A$  zwei identische Spalten, etwa  $a_i$  und  $a_j$ , so gilt wegen (D3) die Beziehung  $D(A) = -D(A)$ , also  $D(A) = 0$ .  $D$  ist also identitätserkennend.

" $\Leftarrow$ ": wir zeigen, dass  $D$  auch die Eigenschaft (D3) hat, also alternierend ist.

$$\begin{aligned} 0 & \stackrel{(D3')}{=} D(\dots, a_i + a_j, \dots, a_i + a_j, \dots) \\ & \stackrel{(D1)}{=} D(\dots, a_i, \dots, a_i, \dots) + D(\dots, a_i, \dots, a_j, \dots) + \\ & \quad D(\dots, a_j, \dots, a_i, \dots) + D(\dots, a_j, \dots, a_j, \dots) \\ & \stackrel{(D3')}{=} D(\dots, a_i, \dots, a_j, \dots) + D(\dots, a_j, \dots, a_i, \dots) . \end{aligned}$$

Daraus folgt (D3). □

**Satz 4.1.3:** Sei  $D : \text{Mat}_{n \times n}(K) \rightarrow K$  eine Determinantenfunktion. Wenn die Spalten von  $A$  linear abhängig sind, so ist  $D(A) = 0$ .

*Beweis:* o.B.d.A. sei  $a_n = \sum_{i=1}^{n-1} \lambda_i a_i$ . Dann ist

$$D(a_1, \dots, a_{n-1}, a_n) \stackrel{(D1),(D2)}{=} \sum_{i=1}^{n-1} \lambda_i D(a_1, \dots, a_{n-1}, a_i) \stackrel{(D3')}{=} 0 .$$

Damit gilt die Behauptung. □

**Satz 4.1.4:** Sei  $D_2 : \text{Mat}_{2 \times 2}(K) \rightarrow K$  die folgende Abbildung von  $2 \times 2$  Matrizen in den Grundkörper:

$$D_2 : \text{Mat}_{2 \times 2}(K) \longrightarrow K \\ \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \mapsto a_{11}a_{22} - a_{12}a_{21} .$$

Dann ist diese Funktion  $D_2$  die einzige Determinantenfunktion auf  $2 \times 2$  Matrizen.

*Beweis:* Man prüft leicht nach, dass  $D_2$  eine Determinantenfunktion ist.

Um die Eindeutigkeit zu zeigen, stellen wir eine beliebige  $2 \times 2$  Matrix  $A$  dar als

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = (a_{11}e_1 + a_{21}e_2, a_{12}e_1 + a_{22}e_2) ,$$

wobei

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} , \quad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} .$$

Sei  $\tilde{D}_2$  ebenfalls eine Determinantenfunktion. Dann gilt wegen (D1)

$$\tilde{D}_2(A) = \tilde{D}_2(a_{11}e_1, a_{12}e_1 + a_{22}e_2) + \tilde{D}_2(a_{21}e_2, a_{12}e_1 + a_{22}e_2) . \quad (*)$$

Der erste Summand in (\*) kann geschrieben werden als

$$\begin{aligned} \tilde{D}_2(a_{11}e_1, a_{12}e_1 + a_{22}e_2) &\stackrel{(D1)}{=} \tilde{D}_2(a_{11}e_1, a_{12}e_1) + \tilde{D}_2(a_{11}e_1, a_{22}e_2) \\ &\stackrel{(D2)}{=} a_{11}a_{12}\tilde{D}_2(e_1, e_1) + a_{11}a_{22}\tilde{D}_2(e_1, e_2) \\ &\stackrel{(D3'),(D4)}{=} a_{11}a_{22} . \end{aligned}$$

Der zweite Summand in (\*) kann geschrieben werden als

$$\begin{aligned} \tilde{D}_2(a_{21}e_2, a_{12}e_1 + a_{22}e_2) &\stackrel{(D1)}{=} \tilde{D}_2(a_{21}e_2, a_{12}e_1) + \tilde{D}_2(a_{21}e_2, a_{22}e_2) \\ &\stackrel{(D2)}{=} a_{21}a_{12}\tilde{D}_2(e_2, e_1) + a_{21}a_{22}\tilde{D}_2(e_2, e_2) \\ &\stackrel{(D3'),(D4)}{=} -a_{21}a_{12} . \end{aligned}$$

Also ist  $\tilde{D}_2(A) = a_{11}a_{22} - a_{12}a_{21} = D_2(A)$ , d.h. die Determinantenfunktion  $D_2$  ist auf  $2 \times 2$  Matrizen eindeutig bestimmt. □

Im folgenden wollen wir nun für beliebige Dimension zeigen, dass es genau eine Determinantenfunktion gibt. Dass dies für  $1 \times 1$  Matrizen ( $a$ ) gilt, ist klar. Die eindeutig bestimmte Determinantenfunktion ist  $D_1(a) = a$ .

**Satz 4.1.5:** Sei  $n \geq 3$  und sei  $D : \text{Mat}_{(n-1) \times (n-1)}(K) \rightarrow K$  eine Determinantenfunktion. Für eine  $n \times n$  Matrix  $A$  bezeichne  $A_{ij}$  die Matrix, welche aus  $A$  entsteht durch Streichung der  $i$ -ten Zeile und  $j$ -ten Spalte. Dann ist für jedes  $i \in \{1, \dots, n\}$  die Funktion

$$f_i : \text{Mat}_{n \times n}(K) \longrightarrow K \\ A \mapsto \sum_{j=1}^n (-1)^{i+j} a_{ij} D(A_{ij})$$

eine Determinantenfunktion. Es gibt also für jede Dimension von Matrizen eine Determinantenfunktion.

*Beweis:* Klarerweise ist  $D(A_{ij})$  unabhängig von der  $j$ -ten Spalte von  $A$ , also hängt  $a_{ij}D(A_{ij})$  linear von der  $j$ -ten Spalte von  $A$  ab. Daraus sieht man, dass  $f_i$  linear abhängig ist von den Spalten von  $A$ , also die Eigenschaften (D1) und (D2) für  $f_i$  erfüllt sind. Dennoch geben wir hier einen ausführlichen Beweis.

Zunächst zu (D1): seien

$$A = (a_1, \dots, a_n), \quad B = (b_1, \dots, b_n), \quad C = (c_1, \dots, c_n), \quad \text{und} \quad k \in \{1, \dots, n\}$$

so, dass  $a_j = b_j = c_j$  für  $j \neq k$  und  $a_k = b_k + c_k$ . Dann gilt

$$\begin{aligned} f_i(A) &= \sum_{j=1}^n (-1)^{i+j} a_{ij} D(A_{ij}) \\ &= \sum_{j=1, j \neq k}^n (-1)^{i+j} a_{ij} (D(B_{ij}) + D(C_{ij})) + (-1)^{i+k} (b_{ik} + c_{ik}) D(A_{ik}) \\ &= \left[ \sum_{j=1, j \neq k}^n (-1)^{i+j} b_{ij} D(B_{ij}) + (-1)^{i+k} b_{ik} D(B_{ik}) \right] + \\ &\quad \left[ \sum_{j=1, j \neq k}^n (-1)^{i+j} c_{ij} D(C_{ij}) + (-1)^{i+k} c_{ik} D(C_{ik}) \right] \\ &= f_i(B) + f_i(C) \end{aligned}$$

Sodann zu (D2): seien

$$A = (a_1, \dots, a_n), \quad B = (b_1, \dots, b_n), \quad \text{und} \quad k \in \{1, \dots, n\}, \quad \lambda \in K$$

so, dass  $a_j = b_j$  für  $j \neq k$  und  $b_k = \lambda a_k$ . Dann gilt

$$\begin{aligned} f_i(B) &= \sum_{j=1, j \neq k}^n (-1)^{i+j} b_{ij} D(B_{ij}) + (-1)^{i+k} b_{ik} D(B_{ik}) \\ &= \sum_{j=1, j \neq k}^n (-1)^{i+j} a_{ij} \lambda D(A_{ij}) + (-1)^{i+k} \lambda a_{ik} D(A_{ik}) \\ &= \lambda f_i(A) \end{aligned}$$

Wir zeigen nun, dass  $f_i$  auch die Eigenschaft (D3') besitzt. Dazu nehmen wir an, dass die Matrix  $A$  zwei identische Spalten, etwa die  $p$ -te und die  $q$ -te mit  $p \neq q$ , hat. Ist nun  $j$  so, dass  $j \notin \{p, q\}$ , dann hat die  $(n-1) \times (n-1)$  Matrix  $A_{ij}$  zwei identische Spalten, und daher gilt

$$D(A_{ij}) = 0 \quad \text{für } j \notin \{p, q\}.$$

Somit haben wir

$$f_i(A) = (-1)^{i+p} a_{ip} D(A_{ip}) + (-1)^{i+q} a_{iq} D(A_{iq}).$$

Ist nun, o.B.d.A.,  $p < q$ , dann können wir offensichtlich mittels  $q-1-p$  Vertauschungen benachbarter Spalten die Matrix  $A_{iq}$  transformieren in die Matrix  $A_{ip}$ . Wegen der Eigenschaft (D3) für  $D$  gilt also

$$D(A_{iq}) = (-1)^{q-1-p} D(A_{ip}).$$

Nun ist aber laut unserer Annahme  $a_{ip} = a_{iq}$ , und daher

$$f_i(A) = [(-1)^{i+p} + (-1)^{i+q} (-1)^{q-1-p}] a_{ip} D(A_{ip}) = 0.$$

Schliesslich sehen wir, dass  $f_i$  auch die Eigenschaft (D4) besitzt. Ist nämlich  $A = I_n$ , dann ist

$$f_i(A) = (-1)^{2i} D(I_{n-1}) = 1.$$

Also ist jedes  $f_i$  eine Determinantenfunktion. Wie wir oben gesehen haben, gibt es für  $n = 1, 2$  Determinantenfunktionen. Somit gibt es für jedes  $n$  eine Determinantenfunktion.  $\square$

**Beispiel 4.1.6:** Wir gehen aus von der Determinantenfunktion  $D_2$  auf  $2 \times 2$  Matrizen. So wie im Satz 4.1.5 ergeben sich daraus die Determinantenfunktionen  $f_i$  auf  $3 \times 3$  Matrizen wie folgt:

$$\begin{aligned}
 f_1 \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} &= \\
 a_{11} \cdot D_2 \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12} \cdot D_2 \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} + a_{13} \cdot D_2 \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} &= \\
 a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} &= \\
 f_2 \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} &= f_3 \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.
 \end{aligned}$$

Wir sehen, dass man den eindeutig bestimmten Wert der Determinantenfunktionen  $f_i$  auf  $3 \times 3$  Matrizen wie folgt nach der sogenannten **Regel von Sarrus** bestimmen kann: man bildet die 3 Produkte der Parallelen zur Diagonale, sowie die negativen Werte der 3 Produkte der Parallelen zur Antidiagonale, und summiert diese 6 Summanden auf.  $\square$

Um die Eindeutigkeit der Determinantenfunktion für beliebiges  $n$  nachzuweisen, gehen wir zunächst einen Umweg. Wir untersuchen gewisse Eigenschaften von Permutationen (also Bijektionen) auf endlichen Mengen. Vergleiche dazu Beispiel 1.5.6(6).

**Definition 4.1.7:** Sei  $n \in \mathbb{N}$ . Eine **Permutation** auf der Menge  $\{1, \dots, n\}$  ist eine Bijektion auf dieser Menge. Wir schreiben üblicherweise eine Permutation  $\sigma$  als

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}.$$

Die Menge der Permutationen auf  $\{1, \dots, n\}$  bezeichnen wir mit  $P_n$ .

Die Hintereinanderausführung auf  $P_n$  nennen wir auch **Produkt**.

Eine **Transposition** auf  $\{1, \dots, n\}$  ist eine Permutation  $\tau$ , welche zwei Elemente  $i$  und  $j$  ( $i \neq j$ ) vertauscht, und alle anderen Elemente fix lässt; in Zeichen  $\tau : i \leftrightarrow j$ .

Ist  $\sigma \in P_n$ , so heisst ein Paar  $(i, j)$  **Fehlstelle** von  $\sigma$ , wenn gilt:  $i < j$  und  $\sigma(i) > \sigma(j)$ .

Hat  $\sigma \in P_n$  genau  $f(\sigma)$  Fehlstellen, so heisst  $(-1)^{f(\sigma)}$  die **Signatur** von  $\sigma$ , in Zeichen  $\text{sign}(\sigma)$ .

Wie wir schon in Kapitel 1.5 gesehen haben, bildet  $P_n$  mit der Hintereinanderausführung von Funktionen, also dem Produkt, eine (nicht-abelsche) Gruppe.

**Beispiel 4.1.8:** Für  $n = 0$  gibt es offensichtlich nur eine Permutation; ebenso für  $n = 1$ . Wir betrachten die Permutationen in  $P_6$ :

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 5 & 3 & 4 & 2 & 6 \end{pmatrix} \quad \text{und} \quad \tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 1 & 2 & 6 & 3 & 5 \end{pmatrix}.$$

Als Verknüpfungsergebnisse erhalten wir

$$\tau \circ \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 3 & 2 & 6 & 1 & 5 \end{pmatrix} \quad \text{bzw.} \quad \sigma \circ \tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 1 & 5 & 6 & 3 & 2 \end{pmatrix} .$$

$\sigma$  hat die 5 Fehlstellen  $(2, 3), (2, 4), (2, 5), (3, 5), (4, 5)$ , also  $\text{sign}(\sigma) = (-1)^5 = -1$ .  $\square$

**Satz 4.1.9:** *Jede Permutation  $\sigma \in P_n$  kann ausgedrückt werden als Produkt von Transpositionen.*

*Beweis:* Für  $n = 0, 1$  ist die Behauptung trivial: die einzige Permutation  $\sigma = \text{id}$  in  $P_0$  bzw.  $P_1$  ist das leere Produkt von Transpositionen.

Sei also nun  $n \geq 2$  beliebig aber fix gewählt.

Induktionshypothese: “Jede Permutation  $\sigma \in P_{n-1}$  kann ausgedrückt werden als Produkt von Transpositionen”.

Induktionsschritt: wir zeigen die Behauptung für  $n$ . Sei  $\sigma \in P_n$ , und sei  $k := \sigma(n)$ . Sei

$$\tau : n \leftrightarrow k .$$

Dann gilt  $(\tau \circ \sigma)(n) = n$ .  $\tau \circ \sigma$  besteht also aus einer Permutation  $(\tau \circ \sigma)^*$  in  $P_{(n-1)}$  und einer Zuordnung von  $n$  auf  $n$ . Laut Induktionshypothese gibt es Transpositionen  $\tau_1^*, \dots, \tau_r^* \in P_{(n-1)}$  mit  $(\tau \circ \sigma)^* = \tau_1^* \circ \dots \circ \tau_r^*$ . Diese Transpositionen induzieren Transpositionen  $\tau_1, \dots, \tau_r \in P_n$  mit  $\tau \circ \sigma = \tau_1 \circ \dots \circ \tau_r$ . Somit haben wir

$$\sigma = \tau^{-1} \circ \tau_1 \circ \dots \circ \tau_r ,$$

dargestellt als Produkt von Transpositionen.  $\square$

**Satz 4.1.10:** *Für  $\sigma, \tau \in P_n$  gilt  $\text{sign}(\tau \circ \sigma) = \text{sign}(\tau) \cdot \text{sign}(\sigma)$ . Insbesondere gilt  $\text{sign}(\sigma^{-1}) = \text{sign}(\sigma)$ .*

*Beweis:* Für  $\sigma \in P_n$  sei

$$v(\sigma) := \prod_{i < j} (\sigma(j) - \sigma(i)) .$$

Insbesondere sei

$$v := v(\text{id}) = \prod_{i < j} (j - i) .$$

Da  $\sigma$  eine Bijektion ist, entsprechen die Faktoren in  $v$  und  $v(\sigma)$  einander eindeutig, eventuell bis auf das Vorzeichen. Konkret haben wir

$$v(\sigma) = \text{sign}(\sigma)v .$$

Ebenso haben wir für  $\tau, \sigma \in P_n$  die Beziehung

$$v(\tau\sigma) = \text{sign}(\tau)v(\sigma) .$$

Daraus folgt

$$\text{sign}(\tau\sigma)v = v(\tau\sigma) = \text{sign}(\tau)v(\sigma) = \text{sign}(\tau)\text{sign}(\sigma)v ,$$

also, wegen  $v \neq 0$ , erhalten wir  $\text{sign}(\tau\sigma) = \text{sign}(\tau) \cdot \text{sign}(\sigma)$ .

Insbesondere gilt  $\text{sign}(\sigma)\text{sign}(\sigma^{-1}) = \text{sign}(\sigma\sigma^{-1}) = \text{sign}(\text{id}) = 1$ .  $\square$

**Definition 4.1.11:** Eine Permutation  $\sigma$  heisst **gerade**, wenn  $\text{sign}(\sigma) = 1$ , und **ungerade**, wenn  $\text{sign}(\sigma) = -1$ .

**Satz 4.1.12 (Leibnizsche Determinantenformel):** Für jedes  $n \in \mathbb{N}$  gibt es eine eindeutig bestimmte Determinantenfunktion  $D_n : \text{Mat}_{n \times n}(K) \rightarrow K$ , welche beschrieben werden kann als

$$D_n(A) = \sum_{\sigma \in P_n} \text{sign}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} .$$

*Beweis:* Wir haben schon in Satz 4.1.5 bewiesen, dass es für jedes  $n$  eine Determinantenfunktion gibt. Wir zeigen nun, dass diese Determinantenform  $D$  notwendigerweise von der obigen Form ist.

Mittels der Einheitsvektoren  $e_i$  schreiben wir zunächst die Matrix  $A = (a_{ij})$  als

$$A = (a_{11}e_1 + \cdots + a_{n1}e_n, \dots, a_{1n}e_1 + \cdots + a_{nn}e_n) .$$

Wegen (D1) und (D2) können wir  $D(A)$  schreiben als Summe von Termen der Form

$$D(a_{\sigma(1),1}e_{\sigma(1)}, \dots, a_{\sigma(n),n}e_{\sigma(n)}) = a_{\sigma(1),1} \cdots a_{\sigma(n),n} D(e_{\sigma(1)}, \dots, e_{\sigma(n)})$$

mit  $1 \leq \sigma(i) \leq n$  für jedes  $i$ . Wegen (D3') sind alle diese Terme 0 mit Ausnahme derjenigen, für welche gilt  $\sigma(i) \neq \sigma(j)$  für  $i \neq j$ , also derjenigen, für welche  $\sigma$  eine Permutation auf  $\{1, \dots, n\}$  ist. Es gilt also

$$D(A) = \sum_{\sigma \in P_n} a_{\sigma(1),1} \cdots a_{\sigma(n),n} D(e_{\sigma(1)}, \dots, e_{\sigma(n)}) .$$

Wegen (D3) können wir mittels  $\sigma^{-1}$  die Spalten in  $(e_{\sigma(1)}, \dots, e_{\sigma(n)})$  vertauschen und dadurch die Einheitsmatrix  $I_n$  erzeugen. Wegen (D4) ist  $D(I_n) = 1$  und ausserdem gilt  $\text{sign}(\sigma) = \text{sign}(\sigma^{-1})$ . Wir erhalten also

$$D(e_{\sigma(1)}, \dots, e_{\sigma(n)}) = \text{sign}(\sigma) D(e_1, \dots, e_n) = \text{sign}(\sigma) .$$

Somit haben wir gezeigt, dass eine Determinantenfunktion notwendigerweise von der Form

$$D_n(A) = \sum_{\sigma \in P_n} \text{sign}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n}$$

ist. □

Aus Satz 4.1.12 sehen wir auch, dass die in Satz 4.1.5 eingeführten Determinantenfunktionen  $f_i$  unabhängig von  $i$  sind, also allesamt dasselbe Ergebnis liefern.

**Definition 4.1.13:** Die laut Satz 4.1.12 eindeutig bestimmte Determinantenfunktion auf  $\text{Mat}_{n \times n}(K)$  bezeichnen wir mit  $\det$ . Für eine  $n \times n$  Matrix  $A = (a_{ij})$  heisst

$$\det(A) = \sum_{\sigma \in P_n} \text{sign}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n}$$

die **Determinante** von  $A$ . Wir schreiben die Determinante auch als  $\det(A) = |a_{ij}|$ . Die Darstellung der Determinante  $\det(A)$  wie in Satz 4.1.5, also

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

heisst die (**Laplace-**) **Entwicklung von  $\det(A)$  nach der  $i$ -ten Zeile.** (P.S. Laplace, 1719–1790)

**Satz 4.1.14:** Für eine quadratische Matrix  $A$  gilt:  $\det(A) = \det(A^T)$ .

*Beweis:* Sei  $B = A^T$ , also  $b_{ji} = a_{ij}$  für alle  $i, j$ . Für jede Permutation  $\sigma$  gilt dann

$$a_{\sigma(1),1} \cdots a_{\sigma(n),n} = a_{1,\sigma^{-1}(1)} \cdots a_{n,\sigma^{-1}(n)} = b_{\sigma^{-1}(1),1} \cdots b_{\sigma^{-1}(n),n} .$$

Durchläuft  $\sigma$  ganz  $P_n$ , so durchläuft auch  $\sigma^{-1}$  ganz  $P_n$ , und ausserdem gilt  $\text{sign}(\sigma) = \text{sign}(\sigma^{-1})$ . Daraus erhalten wir

$$\det(A) = \sum_{\sigma \in P_n} \text{sign}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} = \sum_{\sigma^{-1} \in P_n} \text{sign}(\sigma^{-1}) b_{\sigma^{-1}(1),1} \cdots b_{\sigma^{-1}(n),n} = \det(B) . \square$$

**Korollar 1:** Für jeden Spaltenindex  $j$  gilt auch eine Entwicklung von  $\det(A)$  nach der  $j$ -ten Spalte, also

$$\det(A) = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det(A_{ij}) .$$

**Korollar 2:** Die Eigenschaften (D1), (D2), (D3), (D3') und (D4) der Determinantenfunktion gelten nicht nur für Spalten, sondern auch für Zeilen.

**Beispiel 4.1.15:** Ist  $A$  eine quadratische Matrix in oberer Dreiecksform, also  $a_{ij} = 0$  für alle  $i > j$ , so ist  $\det(A) = |A|$  besonders einfach zu bestimmen, nämlich als das Produkt der Diagonalelemente

$$|A| = \prod_{i=1}^n a_{ii} .$$

Das sieht man aus der Spaltenentwicklung von  $|A|$ .

Dasselbe gilt für untere Dreiecksmatrizen wegen der Zeilenentwicklung von  $|A|$ . □

Eine Möglichkeit, die Determinante einer Matrix  $A$  zu berechnen, ist die Transformation mittels Zeilen- (bzw. Spalten-) umformungen in eine obere Dreiecksmatrix, unter Ausnutzung der Beziehungen (D1 – 4). Dann braucht man nur noch das Produkt der Diagonalelemente zu bilden.

**Beispiel 4.1.16:** Für  $a_1, \dots, a_n \in K$  betrachten wir die Matrix

$$A_n(a_1, \dots, a_n) = \begin{pmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{pmatrix}$$

bzw. ihre Determinante

$$V_n(a_1, \dots, a_n) := \det(A_n) .$$

Man nennt  $V_n(a_1, \dots, a_n)$  die **Vandermonde Determinante** für  $a_1, \dots, a_n$ .

Sind etwa  $a_i = a_j$  für  $i \neq j$ , so hat  $A_n(a_1, \dots, a_n)$  zwei identische Zeilen und ist also wegen

(D3') gleich 0.

Allgemein zeigen wir:

$$V_n(a_1, \dots, a_n) = \prod_{1 \leq i < j \leq n} (a_j - a_i). \quad (*)$$

Dazu gehen wir induktiv nach  $n$  vor. Für  $n = 1$  gilt (\*) offensichtlich. Sei also nun  $n > 1$ . Wir multiplizieren für  $j = n - 1, \dots, 1$  die  $j$ -te Spalte mit  $a_1$  und subtrahieren sie von der  $(j + 1)$ -ten Spalte. Dabei verändert sich wegen (D1) und (D2) die Determinante nicht. Anschliessend ziehen wir jeweils aus der  $i$ -ten Zeile (für  $i > 1$ ) den Faktor  $a_i - a_1$  heraus. Schliesslich entwickeln wir die Determinante nach der ersten Zeile. Wir erhalten also

$$\det(V_n(a_1, \dots, a_n)) = \begin{vmatrix} 1 & 0 & \cdots & 0 \\ 1 & a_2 - a_1 & \cdots & a_2^{n-1} - a_1 a_2^{n-2} \\ \vdots & \vdots & & \vdots \\ 1 & a_n - a_1 & \cdots & a_n^{n-1} - a_1 a_n^{n-2} \end{vmatrix} = \prod_{i=2}^n (a_i - a_1) V_{n-1}(a_2, \dots, a_n).$$

Mit der Induktionshypothese ergibt sich daraus die Behauptung (\*).  $\square$

Die Determinantenfunktion ist multiplikativ; das sehen wir im folgenden Satz.

**Satz 4.1.17:** Für  $A, B \in \text{Mat}_{n \times n}(K)$  gilt:

$$|A \cdot B| = |A| \cdot |B|.$$

*Beweis:* Sei  $C = AB$ . Sei  $A = (a_{ij})$ ,  $B = (b_{ij})$ , und sei  $a_i$  die  $i$ -te Spalte von  $A$ , analog für  $B$  und  $C$ . Dann lässt sich die  $k$ -te Spalte von  $C$  schreiben als

$$c_k = b_{1k}a_1 + \cdots + b_{nk}a_n.$$

Daraus sehen wir

$$\begin{aligned} |AB| &= |b_{11}a_1 + \cdots + b_{n1}a_n, \dots, b_{1n}a_1 + \cdots + b_{nn}a_n| \\ &= \sum_{\sigma \in P_n} |b_{\sigma(1),1}a_{\sigma(1)}, \dots, b_{\sigma(n),n}a_{\sigma(n)}| \\ &= \sum_{\sigma \in P_n} b_{\sigma(1),1} \cdots b_{\sigma(n),n} |a_{\sigma(1)}, \dots, a_{\sigma(n)}| \\ &= \sum_{\sigma \in P_n} b_{\sigma(1),1} \cdots b_{\sigma(n),n} \text{sign}(\sigma) |a_1, \dots, a_n| \\ &= |A| \cdot \sum_{\sigma \in P_n} \text{sign}(\sigma) b_{\sigma(1),1} \cdots b_{\sigma(n),n} \\ &= |A| \cdot |B| \end{aligned} \quad \square$$

**Korollar 1:** Ist  $A$  invertierbar, dann gilt  $|A| \neq 0$  und  $|A^{-1}| = 1/|A|$ .

*Beweis:*  $|A| \cdot |A^{-1}| = |I| = 1$ .  $\square$

**Korollar 2:** Ist  $A$  eine  $n \times n$  Matrix, und  $P$  eine reguläre (invertierbare)  $n \times n$  Matrix, so gilt

$$|A| = |PAP^{-1}|.$$



*Beweis:*

$$|PAP^{-1}| = |P| \cdot |A| \cdot |P^{-1}| = |P| \cdot |A| \cdot \frac{1}{|P|} = |A|. \quad \square$$

**Definition 4.1.18:** Sei  $\mathcal{A}$  die Darstellungsmatrix einer linearen Abbildung  $h$  eines  $n$ -dimensionalen Vektorraums  $V$  in sich. Dann ist  $|\mathcal{A}|$  die **Determinante** von  $h$ ; geschrieben als  $|h|$ .

*Beachte:* Darstellungsmatrizen bzgl. verschiedener Basen sind ähnlich (vgl. Def.3.3.39), wegen Korollar 2 hängt also  $|h|$  nicht von der Basisdarstellung ab.  $\square$

Die Determinante charakterisiert also die Invertierbarkeit einer quadratischen Matrix  $A$  (vgl. Satz 4.1.3 und Korollar zu Satz 4.1.17). Wir wollen nun aus der Determinante eine Formel für die Inverse herleiten. Dazu führen wir den Begriff der adjungierten Matrix ein. Wir erinnern daran, dass  $A_{ij}$  die Matrix bezeichnet, welche wir aus  $A$  erhalten durch Streichung der  $i$ -ten Zeile und  $j$ -ten Spalte.

**Definition 4.1.19:** Sei  $A \in \text{Mat}_{n \times n}(K)$ . Dann ist die **adjungierte Matrix** bzw. **Adjungierte** von  $A$ , in Zeichen  $\text{adj}(A)$ , diejenige  $n \times n$  Matrix, deren Element in der  $i$ -ten Zeile und  $j$ -ten Spalte

$$(-1)^{i+j} |A_{ji}|$$

ist.

Man beachte die Umkehrung der Indizes in Definition 4.1.19.

**Satz 4.1.20:** Für jede quadratische Matrix gilt  $(\text{adj}(A))^T = \text{adj}(A^T)$ .

*Beweis:*  $(\text{adj}(A))^T = ((-1)^{i+j} |A_{ij}|) = ((-1)^{i+j} |(A^T)_{ji}|) = \text{adj}(A^T)$ .  $\square$

**Satz 4.1.21:** Für jede  $n \times n$  Matrix  $A$  gilt:

$$A \cdot \text{adj}(A) = |A| \cdot I_n = \text{adj}(A) \cdot A.$$

*Beweis:* Sei  $B = \text{adj}(A)$  und  $C = AB$ . Dann gilt

$$c_{ij} = \sum_{k=1}^n a_{ik} b_{kj} = \sum_{k=1}^n a_{ik} (-1)^{k+j} |A_{jk}| = \begin{cases} |A| & \text{falls } i = j, \\ 0 & \text{falls } i \neq j. \end{cases}$$

Die letzte Gleichheit gilt deshalb, weil im Fall  $i \neq j$  der Ausdruck die Determinante einer Matrix darstellt, deren  $j$ -te Zeile gleich ihrer  $i$ -ten Zeile ist. Somit haben wir gezeigt  $A \cdot \text{adj}(A) = |A| I_n$ .

Den zweiten Teil der Beziehung sieht man mittels des ersten Teils und Satz 7.19:

$$\text{adj}(A) \cdot A = (A^T \cdot \text{adj}(A)^T)^T = (A^T \cdot \text{adj}(A^T))^T = (|A^T| \cdot I_n)^T = |A| \cdot I_n. \quad \square$$

**Satz 4.1.22:** Eine quadratische Matrix  $A$  ist invertierbar g.d.w.  $|A| \neq 0$ ; in diesem Fall ist

$$A^{-1} = \frac{1}{|A|} \text{adj}(A).$$

*Beweis:* Laut Satz 4.1.3 folgt aus der Singularität von  $A$ , dass  $|A| = 0$ . Andererseits folgt wegen des Korollars 1 zu Satz 4.1.17 aus der Invertierbarkeit von  $A$ , dass  $|A| \neq 0$ . Ist nun  $|A| \neq 0$ , so gilt wegen Satz 4.1.21

$$A \cdot \frac{1}{|A|} \cdot \text{adj}(A) = I_n .$$

Somit ist  $A$  invertierbar und

$$A^{-1} = \frac{1}{|A|} \cdot \text{adj}(A)$$

ist die Inverse. □

Determinanten von Diagonalmatrizen bzw. Blockdiagonalmatrizen sind besonders einfach auszurechnen. Sind  $B = (b_{ij})_{p \times p}$  und  $C = (c_{ij})_{q \times q}$  quadratische Matrizen, so ist

$$A = (a_{ij})_{n \times n}, \text{ mit } n = p + q \text{ und } a_{ij} = \begin{cases} b_{ij} & \text{falls } i, j \leq p \\ c_{i-p, j-p} & \text{falls } p + 1 \leq i, j \\ 0 & \text{sonst} \end{cases}$$

die Blockdiagonalmatrix mit Blöcken  $B$  und  $C$ . Die Matrix  $A$  hat also die Gestalt

$$A = \begin{pmatrix} B & \vdots & 0 \\ \dots & \dots & \dots \\ 0 & \vdots & C \end{pmatrix} \quad (*)$$

**Satz 4.1.23:**

- (a) Ist  $A = (a_{ij})$  eine  $n \times n$  Diagonalmatrix, so ist  $|A| = \prod_{i=1}^n a_{ii}$ .
- (b) Ist  $A = (a_{ij})_{n \times n}$  eine Blockdiagonalmatrix mit Blöcken  $B = (b_{ij})_{p \times p}$  und  $C = (c_{ij})_{q \times q}$  wie in (\*), so ist  $|A| = |B| \cdot |C|$ .

*Beweis:* (a) offensichtlich aus der Definition der Determinante.

(b) Induktion nach  $p = \dim B$ .

“ $p = 1$ ”:  $B = (b)$ , also durch Entwicklung nach der 1. Zeile erhalten wir

$$|A| = (-1)^{1+1} \cdot b \cdot |C| = |B| \cdot |C|$$

“ $p - 1 \rightarrow p$ ”: wieder entwickeln wir nach der 1. Zeile und verwenden die Induktionshypothese:

$$\begin{aligned} |A| &= \sum_{j=1}^p (-1)^{1+j} b_{1j} |A_{1j}| = \sum_{j=1}^p (-1)^{1+j} b_{1j} \cdot |B_{1j}| \cdot |C| = \\ &|C| \cdot \sum_{j=1}^p (-1)^{1+j} b_{1j} \cdot |B_{1j}| = |C| \cdot |B| \quad \square \end{aligned}$$

**Definition 4.1.24:** Die Untergruppe von  $GL_n(K)$ , welche aus den Matrizen mit Determinante 1 besteht, heisst die **spezielle lineare Gruppe** (special linear group) vom Grad  $n$  über  $K$ ,  $SL_n(K)$ .

## Determinanten und lineare Gleichungssysteme

Mittels Determinanten können wir für ein SLG  $Ax = b$  mit regulärer quadratischer Matrix  $A$  eine Lösungsformel angeben. Diese Formel heisst die **Cramersche Regel** (Gabriel Cramer, 1704–1752).

**Satz 4.1.25:** Sei  $A \in \text{Mat}_{n \times n}(K)$  mit  $|A| \neq 0$ , und sei  $b$  ein Spaltenvektor über  $K$  der Länge  $n$ . Seien  $a_1, \dots, a_n$  die Spalten von  $A$ , also  $A = (a_1, \dots, a_n)$ . Dann können die Komponenten  $x_i$  der eindeutigen Lösung von  $Ax = b$  ausgedrückt werden als

$$x_i = \frac{|a_1, \dots, a_{i-1}, b, a_{i+1}, \dots, a_n|}{|A|}.$$

*Beweis:* Für die eindeutig bestimmte Lösung  $x = (x_1, \dots, x_n)^T$  gilt

$$A \cdot x = x_1 a_1 + \dots + x_n a_n = b.$$

Daraus ergibt sich

$$\begin{aligned} |a_1, \dots, a_{i-1}, b, a_{i+1}, \dots, a_n| &= |a_1, \dots, a_{i-1}, x_1 a_1 + \dots + x_n a_n, a_{i+1}, \dots, a_n| \\ &= \sum_{j=1}^n |a_1, \dots, a_{i-1}, x_j a_j, a_{i+1}, \dots, a_n| \\ &= \sum_{j=1}^n x_j \cdot |a_1, \dots, a_{i-1}, a_j, a_{i+1}, \dots, a_n| \\ &= x_i \cdot |a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_n|. \end{aligned}$$

Dabei gilt die letzte Gleichung deswegen, weil alle anderen Summanden eine Determinante mit zwei gleichen Spalten beinhalten. Daraus ergibt sich unmittelbar die Darstellung

$$x_i = \frac{|a_1, \dots, a_{i-1}, b, a_{i+1}, \dots, a_n|}{|A|}. \quad \square$$

**Beispiel 4.1.26:** Wir betrachten das SLG aus Beispiel 2.2.1, also

$$\begin{pmatrix} 0 & 1 & 2 \\ 1 & -2 & 1 \\ 0 & 3 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 23 \end{pmatrix},$$

welches die eindeutig bestimmte Lösung  $x = (12, 5, -2)^T$  besitzt.

Als Determinante der Koeffizientenmatrix erhalten wir  $|A| = 10$ .

Ersetzen wir die erste Spalte von  $A$  durch die rechte Seite  $b$ , so ist die Determinante dieser Matrix 120.

Somit ergibt sich nach der Cramerschen Regel  $x_1 = 120/10 = 12$ . □

## 4.2 Eigenwerte und Eigenvektoren

In Kapitel 4.1 haben wir gesehen, dass für die Darstellungsmatrizen einer linearen Abbildung  $h$  bzgl. der Basen  $B, B'$  folgende Beziehung gilt:

$$\mathcal{A}(h, B', B') = \mathcal{A}_{B'}^B \cdot \mathcal{A}(h, B, B) \cdot (\mathcal{A}_{B'}^B)^{-1}.$$

In Def. 3.3.39 haben wir zwei  $n \times n$  Matrizen  $A, A'$  ähnlich genannt, wenn es eine reguläre Matrix  $P$  gibt, mit

$$A' = P^{-1}AP.$$

In diesem Kapitel wollen wir untersuchen, unter welchen Bedingungen eine Matrix zu einer Diagonalmatrix ähnlich ist; also unter welchen Bedingungen es eine Basis  $B$  gibt, bezüglich derer ein linearer Endomorphismus  $h$  eine Diagonaldarstellung besitzt.

Wir werden in diesem Kapitel immer quadratische  $n \times n$  Matrizen über einem Körper  $K$  betrachten, wenn nichts anderes explizit vereinbart wird.

**Definition 4.2.1:** Sei  $A$  eine  $n \times n$  Matrix und sei  $\lambda$  ein Skalar ( $\in K$ ). Wenn es einen nicht-trivialen (also vom Nullvektor verschiedenen) Spaltenvektor  $x$  gibt, sodass

$$Ax = \lambda x,$$

dann ist  $\lambda$  ein **Eigenwert** der Matrix  $A$  und  $x$  ein zum Eigenwert  $\lambda$  gehöriger **Eigenvektor**.

Das **Spektrum** von  $A$ , geschrieben  $\sigma(A)$ , ist die Menge der Eigenwerte von  $A$ .  $\square$

**Bemerkung:** Der Spektralradius  $\rho$  einer  $n \times n$  Matrix über dem komplexen Zahlkörper  $\mathbb{C}$  ist der Betrag des betragsmässig grössten Eigenwerts von  $A$ ; also

$$\rho(A) = \max\{|\lambda| \mid \lambda \in \sigma(A)\}.$$

Der Spektralradius ist etwa bei der numerischen Approximation von Lösungen linearer Gleichungssysteme  $A \cdot x = b$  ( $A$  regulär) von Bedeutung. Falls  $\rho(I - B^{-1}A) < 1$ , dann konvergiert das Iterationsverfahren

$$x_{k+1} = (I - B^{-1}A)x_k + B^{-1}b$$

für jeden Startvektor  $x_0$  gegen die exakte Lösung  $x^*$  des linearen Gleichungssystems.  $\square$

**Satz 4.2.2:** Ein Skalar  $\lambda$  ist ein Eigenwert der Matrix  $A$  genau dann, wenn

$$|A - \lambda I_n| = 0.$$

*Beweis:* Natürlich können wir  $Ax = \lambda x$  auch schreiben als  $(A - \lambda I_n)x = 0$ . Also ist  $\lambda$  ein Eigenwert von  $A$  genau dann, wenn das homogene SLG

$$(A - \lambda I_n)x = 0$$

eine nichttriviale Lösung besitzt. Wegen der Sätze 2.2.32 und 2.3.5 ist das genau dann der Fall, wenn die Matrix  $A - \lambda I_n$  nicht invertierbar ist. Das wiederum ist wegen Satz 4.1.22 genau dann der Fall, wenn gilt  $|A - \lambda I_n| = 0$ .  $\square$

**Korollar:** Ähnliche Matrizen haben dieselben Eigenwerte.

*Beweis:* Wegen Satz 4.1.17 und zugehörigem Korollar gilt für eine invertierbare Matrix  $P$

$$|P^{-1}AP - \lambda I_n| = |P^{-1}(A - \lambda I_n)P| = |P^{-1}| \cdot |A - \lambda I_n| \cdot |P| = |A - \lambda I_n|.$$

Damit folgt die Behauptung unmittelbar aus dem Satz. □

Schreiben wir  $A = (a_{ij})$ , so sehen wir, dass offensichtlich

$$|A - \lambda I_n| = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix}$$

ein Polynom vom Grad  $n$  in  $\lambda$  ist.

**Definition 4.2.3:** Das Polynom  $|A - \lambda I_n|$  vom Grad  $n$  in der Variablen  $\lambda$  heisst das **charakteristische Polynom** der Matrix  $A$ , geschrieben als  $c_A(\lambda)$ . Unter der **charakteristischen Gleichung** von  $A$  verstehen wir die Gleichung

$$c_A(\lambda) = |A - \lambda I_n| = 0. \quad \square$$

Somit könnten wir Satz 4.2.2 auch folgendermassen ausdrücken: Die Eigenwerte von  $A$  sind die Nullstellen des charakteristischen Polynoms von  $A$ . Ist der Grundkörper etwa  $K = \mathbb{C}$  (oder ein anderer algebraisch abgeschlossener Körper), so hat jedes univariate Polynom mit Koeffizienten in  $K$  auch (mindestens) eine Lösung in  $K$ . Insbesondere sieht man daraus, dass jede Matrix einen Eigenwert hat. Weiters lässt sich das charakteristische Polynom über einem algebraisch abgeschlossenen Körper in  $n$  Linearfaktoren zerlegen, also

$$c_A(\lambda) = (-1)^n (\lambda - \lambda_1)^{r_1} (\lambda - \lambda_2)^{r_2} \cdots (\lambda - \lambda_k)^{r_k}, \quad (*)$$

wobei  $\lambda_1, \dots, \lambda_k$  die Eigenwerte von  $A$  sind und  $r_1 + \cdots + r_k = n$ .

**Definition 4.2.4:** Die Exponenten  $r_1, \dots, r_k$  in der Gleichung  $(*)$  heissen die zu den Eigenwerten  $\lambda_1, \dots, \lambda_k$  von  $A$  gehörigen **algebraischen Vielfachheiten**.

**Beispiel 4.2.5:** (a) Wir betrachten die Matrix

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

in  $\text{Mat}_{2 \times 2}(\mathbb{R})$ . Das charakteristische Polynom von  $A$  ist

$$c_A(\lambda) = |A - \lambda I_2| = \begin{vmatrix} -\lambda & 1 \\ -1 & -\lambda \end{vmatrix} = \lambda^2 + 1 = (\lambda + i)(\lambda - i).$$

Über dem Körper  $\mathbb{R}$  hat also  $A$  keinen Eigenwert. Aber natürlich können wir die Matrix  $A$  auch über dem Grundkörper  $\mathbb{C}$  betrachten, und dann hat  $A$  die Eigenwerte  $\pm i$ , jeden mit algebraischer Vielfachheit 1.

(b) Wir betrachten die Matrix

$$B = \begin{pmatrix} -3 & 1 & -1 \\ -7 & 5 & -1 \\ -6 & 6 & -2 \end{pmatrix}$$

in  $\text{Mat}_{3 \times 3}(\mathbb{R})$ . Das charakteristische Polynom von  $B$  ist

$$c_B(\lambda) = |B - \lambda I_3| = \begin{vmatrix} -3 - \lambda & 1 & -1 \\ -7 & 5 - \lambda & -1 \\ -6 & 6 & -2 - \lambda \end{vmatrix} = -(\lambda + 2)^2(\lambda - 4) .$$

$B$  hat also die reellen Eigenwerte  $-2$  und  $4$ , und zwar mit den algebraischen Vielfachheiten  $2$  bzw.  $1$ .  $\square$

**Lemma 4.2.6:** Sei  $A \in \text{Mat}_{n \times n}(K)$  und sei  $\lambda$  ein Eigenwert von  $A$ . Dann ist die Menge (von Spaltenvektoren)

$$E_{A,\lambda} = \{x \in \text{Mat}_{n \times 1}(K) \mid Ax = \lambda x\}$$

(also die Menge der zu  $A$  und  $\lambda$  gehörigen Eigenvektoren zusammen mit dem Nullvektor) ein linearer Unterraum von  $\text{Mat}_{n \times 1}(K)$ .

*Beweis:* für  $x, y \in E_{A,\lambda}$  und  $\mu, \nu \in K$  ist offensichtlich auch  $\mu x + \nu y \in E_{A,\lambda}$ .  $\square$

**Definition 4.2.7:** Seien  $A$  und  $\lambda$  wie in Lemma 4.2.6. Der lineare Raum  $E_{A,\lambda}$  heisst der zu  $A$  und  $\lambda$  gehörige **Eigenraum**.  $\dim(E_{A,\lambda})$  heisst die **geometrische Vielfachheit** des Eigenwertes  $\lambda$ .

Ist  $A$  aus dem Kontext klar, so schreiben wir statt  $E_{A,\lambda}$  auch einfach  $E_\lambda$ .

**Beispiel 4.2.8:** Sei  $B$  die  $3 \times 3$  Matrix aus Beispiel 4.2.5(b). Der Eigenraum zum Eigenvektor  $-2$ , also  $E_{B,-2} = E_{-2}$  besteht aus den Lösungen des SLG

$$\begin{pmatrix} -1 & 1 & -1 \\ -7 & 7 & -1 \\ -6 & 6 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} .$$

Man prüft leicht nach, dass

$$E_{-2} = \text{span}((1, 1, 0)^T) .$$

Die geometrische Vielfachheit des Eigenwertes  $-2$  ist also  $1$  und somit kleiner als die algebraische Vielfachheit.

Der Eigenraum zum Eigenwert  $4$ , also  $E_4$ , besteht aus den Lösungen des SLG

$$\begin{pmatrix} -7 & 1 & -1 \\ -7 & 1 & -1 \\ -6 & 6 & -6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} .$$

Man prüft leicht nach, dass

$$E_4 = \text{span}((0, 1, 1)^T) .$$

Die geometrische Vielfachheit des Eigenwertes 4 ist also 1 und somit gleich wie die algebraische Vielfachheit.  $\square$

**Satz 4.2.9:** Sind  $\lambda_1, \lambda_2$  zwei verschiedene Eigenwerte von  $A$ , dann gilt  $E_{\lambda_1} \cap E_{\lambda_2} = \{0\}$ .

*Beweis:*  $x \in E_{\lambda_1} \cap E_{\lambda_2}$  g.d.w.  $\lambda_1 x = Ax = \lambda_2 x$ . Somit ist  $x$  im Durchschnitt der Eigenräume g.d.w.  $(\lambda_1 - \lambda_2)x = 0$ , also  $x = 0$ .  $\square$

Wir wissen, dass sich jede lineare Abbildung zwischen endlichdimensionalen Vektorräumen darstellen lässt als Multiplikation mit einer Matrix (Darstellungsmatrix). Andererseits ist natürlich jede Abbildung der Form

$$x \longrightarrow Ax$$

linear. Dieser Zusammenhang führt auf natürliche Weise zum Begriff des Eigenwertes einer linearen Abbildung.

**Definition 4.2.10:** Sei  $V$  ein endlichdimensionaler Vektorraum über  $K$ . Sei  $f : V \rightarrow V$  eine lineare Abbildung und sei  $\lambda$  ein Skalar im Grundkörper  $K$ . Wenn es einen Vektor  $x \in V \setminus \{0\}$  gibt mit

$$f(x) = \lambda x ,$$

dann ist  $\lambda$  ein **Eigenwert** von  $f$  und  $x$  ein zum Eigenwert  $\lambda$  gehöriger **Eigenvektor**.

Das **Spektrum** von  $f$ , geschrieben  $\sigma(f)$ , ist die Menge der Eigenwerte von  $f$ .  $\square$

Das Spektrum eines Endomorphismus  $f$  ist das Spektrum einer beliebigen Darstellungsmatrix von  $f$ .

**Beispiel 4.2.11:** (a) Sei  $f$  die lineare Abbildung

$$f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3 \\ (x, y, z) \mapsto (y + z, x + z, x + y) .$$

Bezüglich der natürlichen geordneten Basis von  $\mathbb{R}^3$  hat  $f$  die Darstellungsmatrix

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

Die Matrix  $A$  hat das charakteristische Polynom

$$c_A(\lambda) = -(\lambda + 1)^2(\lambda - 2) .$$

Somit hat  $A$  und daher auch  $f$  die Eigenwerte -1 und 2, und zwar mit algebraischer Vielfachheit 2 bzw. 1.

Bezüglich der geordneten Basis

$$B' = ((1, 1, 1), (1, 1, 0), (1, 0, 0))$$

hat  $f$  die Darstellungsmatrix

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

Die Matrix  $A'$  hat das charakteristische Polynom

$$c_{A'}(\lambda) = -(\lambda + 1)^2(\lambda - 2) .$$

Die Eigenwerte von  $A$  und  $A'$  stimmen also überein.

(b) Sei  $\text{Diff}(\mathbb{R}, \mathbb{R})$  der Vektorraum aller unendlich oft differenzierbaren reellen Funktionen (über  $\mathbb{R}$ ). Die Differentiationsabbildung

$$D : \text{Diff}(\mathbb{R}, \mathbb{R}) \longrightarrow \text{Diff}(\mathbb{R}, \mathbb{R})$$

ist linear. Ein Eigenvektor von  $D$  ist eine von der Nullfunktion verschiedene differentierbare Funktion  $f(x)$ , sodass für ein  $\lambda \in \mathbb{R}$  gilt:  $D(f) = \lambda f$ . Aus der Analysis und der Theorie der Differentialgleichung ergibt sich, dass die Eigenvektoren von  $D$  die Funktionen der Form

$$f(x) = ke^{\lambda x}$$

sind. Die lineare Abbildung  $\text{Diff}$  hat also unendlich viele Eigenwerte  $\lambda \in \mathbb{R} \setminus \{0\}$ .  $\square$

**Satz 4.2.12:** Sei  $f : V \rightarrow V$  eine lineare Abbildung, seien  $\lambda_1, \dots, \lambda_k$  verschiedene Eigenwerte von  $f$  und sei für jedes  $i \in \{1, \dots, k\}$  der Vektor  $x_i$  ein Eigenvektor zum Eigenwert  $\lambda_i$ . Dann sind die Vektoren  $x_1, \dots, x_k$  linear unabhängig.

*Beweis:* Wir führen den Beweis mittels Induktion über  $k$ . Für  $k = 1$  ist  $x_1$  natürlich linear unabhängig, weil  $x_1 \neq 0$ .

Wir nehmen nun in der Induktionshypothese an, dass für ein beliebiges aber fixes  $k$  schon gelte, dass Eigenvektoren  $x_1, \dots, x_k$  zu verschiedenen Eigenwerten  $\lambda_1, \dots, \lambda_k$  linear unabhängig sind.

Seien also nun  $x_1, \dots, x_{k+1}$  Eigenvektoren zu verschiedenen Eigenwerten  $\lambda_1, \dots, \lambda_{k+1}$ . Wenn wir auf die Gleichung

$$a_1x_1 + \dots + a_kx_k + a_{k+1}x_{k+1} = 0 \quad (*)$$

die Abbildung  $f$  anwenden und die Beziehung  $f(x_i) = \lambda_i x_i$  verwenden, so erhalten wir

$$a_1\lambda_1x_1 + \dots + a_k\lambda_kx_k + a_{k+1}\lambda_{k+1}x_{k+1} = 0 . \quad (**)$$

Kombinieren wir nun diese Gleichungen in der Form  $(**) - \lambda_{k+1}(*),$  so erhalten wir

$$a_1(\lambda_1 - \lambda_{k+1})x_1 + \dots + a_k(\lambda_k - \lambda_{k+1})x_k = 0 .$$

Aufgrund der Induktionshypothese und der Verschiedenheit der Eigenwerte gilt also

$$a_1 = \dots = a_k = 0 .$$

Aus  $(*)$  und  $x_{k+1} \neq 0$  schliessen wir weiters  $a_{k+1} = 0$ . Somit sind die Vektoren  $x_1, \dots, x_{k+1}$  linear unabhängig.  $\square$

**Definition 4.2.13:** Die lineare Abbildung  $f : V \rightarrow V$  auf einem endlichdimensionalen Vektorraum  $V$  heisst **diagonalisierbar**, wenn es eine geordnete Basis  $(v_1, \dots, v_n)$  von  $V$  gibt, bezüglich derer die Darstellungsmatrix von  $f$  eine Diagonalmatrix ist.

Ebenso nennen wir eine quadratische Matrix  $A$  **diagonalisierbar**, wenn sie ähnlich ist zu einer Diagonalmatrix.



**Satz 4.2.14:** (a) Eine lineare Abbildung  $f : V \rightarrow V$  auf einem endlichdimensionalen Vektorraum  $V$  ist diagonalisierbar g.d.w.  $V$  eine Basis besitzt, welche aus Eigenvektoren von  $f$  besteht.

(b) Eine quadratische  $n \times n$  Matrix  $A$  ist diagonalisierbar g.d.w. sie  $n$  linear unabhängige Eigenvektoren besitzt.

*Beweis:* (a) Laut Definition ist  $f$  diagonalisierbar g.d.w. es eine geordnete Basis  $(v_1, \dots, v_n)$  gibt, sodass

$$\begin{aligned} f(v_1) &= \lambda_1 v_1 \\ f(v_2) &= \lambda_2 v_2 \\ &\vdots \\ f(v_n) &= \lambda_n v_n . \end{aligned}$$

Das ist offensichtlich genau dann der Fall, wenn die  $\lambda_i$  Eigenwerte mit zugehörigen Eigenvektoren  $v_i$  sind.

(b) Wir interpretieren die Matrix  $A$  als Darstellungsmatrix der linearen Abbildung

$$\begin{aligned} f_A : K_n &\longrightarrow K_n \\ x &\longmapsto Ax . \end{aligned}$$

Die Matrix  $A$  ist diagonalisierbar g.d.w.

es eine reguläre Matrix  $P$  und eine Diagonalmatrix  $D$  gibt, sodass  $D = P^{-1}AP$ . Diese Matrix  $P$  ist natürlich eine Basistransformationsmatrix. Das ist also der Fall g.d.w.

die lineare Abbildung  $f_A$  diagonalisierbar ist, also (wegen (a)) g.d.w.

es eine geordnete Basis  $(v_1, \dots, v_n)$  von  $K_n$  und  $\lambda_1, \dots, \lambda_n$  gibt, sodass  $Av_i = f_A(v_i) = \lambda_i v_i$  für alle  $i$ . Das aber heisst nichts anderes als dass  $A$   $n$  linear unabhängige Eigenvektoren besitzt.  $\square$

**Bemerkung:** Laut Definition ist  $A$  diagonalisierbar g.d.w. es eine invertierbare Matrix  $P$  und eine Diagonalmatrix  $D$  gibt, sodass  $D = P^{-1}AP$ .  $A$  und  $D$  haben dieselben Eigenwerte. Sei

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} .$$

Die kanonische Basis von  $K_n$  ist eine Basis aus Eigenvektoren von  $D$ :  $De_i = \lambda_i e_i$ . Die Vektoren  $Pe_1, \dots, Pe_n$  sind Eigenvektoren von  $A$ , denn

$$A(Pe_i) = PDP^{-1}(Pe_i) = PDe_i = P\lambda_i e_i = \lambda_i(Pe_i) .$$

Ausserdem sind die Vektoren  $Pe_1, \dots, Pe_n$  linear unabhängig, denn

$$0 = \sum \mu_i(Pe_i) = P\left(\sum \mu_i e_i\right) \iff \sum \mu_i e_i = 0 \iff \mu_i = 0 \forall i .$$

Somit gibt es  $n$  linear unabhängige Eigenvektoren von  $A$ .

Umgekehrt seien  $x_1, \dots, x_n$  linear unabhängig und  $\lambda_1, \dots, \lambda_n$  so, dass  $Ax_i = \lambda_i x_i$  für alle  $i$  gilt. Sei  $P$  die (Basistransformations-) Matrix, für welche gilt

$$Pe_i = x_i \quad \forall i ,$$

also  $x_i$  ist die  $i$ -te Spalte von  $P$ . Dann gilt für  $D := P^{-1}AP$ :

$$De_i = P^{-1}APe_i = \lambda_i e_i \quad \forall i.$$

Also ist  $A$  ähnlich zur Diagonalmatrix

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

und somit diagonalisierbar. □

**Beispiel:** Sei  $A$  die Darstellungsmatrix einer linearen Abbildung von  $\mathbb{R}^2$  nach  $\mathbb{R}^2$ :

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

$A$  ist diagonalisierbar, denn für

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

gilt  $D = P^{-1}AP$ .

Die Diagonalmatrix  $D$  hat offensichtlich die Eigenwerte 1 und 2, und das sind wegen der Ähnlichkeit auch die Eigenwerte von  $A$ .

$(1, 0)^T$  ist ein Eigenvektor von  $D$  zum Eigenwert 1, und

$$P \cdot (1, 0)^T = (0, 1)^T$$

ist ein Eigenvektor von  $A$  zum Eigenwert 1.

$(0, 1)^T$  ist ein Eigenvektor von  $D$  zum Eigenwert 2, und

$$P \cdot (0, 1)^T = (1, -1)^T$$

ist ein Eigenvektor von  $A$  zum Eigenwert 2. □

Nun wollen wir zeigen, dass eine lineare Abbildung  $f : V \rightarrow V$  bzw. die zugehörige Matrix diagonalisierbar ist genau dann, wenn für jeden Eigenwert die algebraische und die geometrische Vielfachheit übereinstimmen. Dazu brauchen wir zunächst einige Vorbereitung.

**Satz 4.2.15:** Sei  $f : V \rightarrow V$  eine lineare Abbildung auf dem  $n$ -dimensionalen Vektorraum  $V$ . Seien  $\lambda_1, \dots, \lambda_k$  die verschiedenen Eigenwerte von  $f$ , und  $d_1, \dots, d_k$  die zugehörigen geometrischen Vielfachheiten. Dann gilt

$$d_1 + \dots + d_k \leq n. \quad (*)$$

In  $(*)$  gilt die Gleichheit g.d.w.  $f$  diagonalisierbar ist.

*Beweis:* Für jedes  $i \in \{1, \dots, k\}$  sei  $B_i = \{b_{i1}, \dots, b_{id_i}\}$  eine Basis von  $E_{\lambda_i}$ . Aus Satz 4.2.12 folgt, dass

$$v_1 + \dots + v_k = 0, \quad \text{mit } v_i \in E_{\lambda_i}$$

nur möglich ist für  $v_i = 0$  für alle  $i$ .

Weiters ist  $B := \bigcup_{i=1}^k B_i$  linear unabhängig. Denn für

$$v_i = \mu_{i1}b_{i1} + \dots + \mu_{id_i}b_{id_i} \in B_i$$

und

$$v_1 + \dots + v_k = 0$$

folgt wegen der obigen Überlegung  $v_i = 0$  für alle  $i$ , und daraus weiter  $\mu_{ij} = 0$  für alle  $i, j$ . Somit haben wir

$$d_1 + \dots + d_k = \left| \bigcup_{i=1}^k B_i \right| \leq n.$$

Schliesslich gilt die Gleichheit genau dann, wenn  $V$   $n$  linear unabhängige Eigenvektoren hat. Laut Satz 4.2.14 ist das genau dann der Fall, wenn  $f$  diagonalisierbar ist.  $\square$

**Satz 4.2.16:** Für jeden Eigenwert  $\lambda$  einer linearen Abbildung  $f : V \rightarrow V$  auf einem endlichdimensionalen Vektorraum  $V$  ist die algebraische Vielfachheit mindestens so gross wie die geometrische Vielfachheit.

*Beweis:* Sei  $d$  die geometrische Vielfachheit von  $\lambda$ , und sei  $(v_1, \dots, v_d)$  eine geordnete Basis von  $E_\lambda$ . Wir erweitern diese Folge linear unabhängiger Vektoren zu einer Basis  $B = (v_1, \dots, v_d, \dots, v_n)$  von  $V$ . Die Darstellungsmatrix  $A$  der linearen Abbildung  $f$  bzgl. der Basis  $B$  hat dann die Gestalt

$$A = \begin{pmatrix} \lambda I_d & C \\ 0 & D \end{pmatrix}.$$

Das charakteristische Polynom von  $A$  hat also die Form

$$c_A(x) = (\lambda - x)^d p(x),$$

wobei  $p(x)$  ein Polynom vom Grad  $n - d$  ist. Die algebraische Vielfachheit von  $\lambda$  ist also mindestens  $d$ .  $\square$

**Satz 4.2.17:** Sei  $f : V \rightarrow V$  eine lineare Abbildung auf dem endlichdimensionalen Vektorraum  $V$ . Die folgenden Aussagen sind äquivalent:

- (i)  $f$  ist diagonalisierbar;
- (ii) für jeden Eigenwert  $\lambda$  von  $f$  ist die geometrische Vielfachheit gleich der algebraischen Vielfachheit.

*Beweis:* Die Summe der algebraischen Vielfachheiten ist der Grad des charakteristischen Polynoms, nämlich  $n = \dim(V)$ . Mit den Sätzen 4.2.15 und 4.2.16 folgt die Behauptung.  $\square$

Wir haben also nun ein Kriterium für die Diagonalisierbarkeit einer linearen Abbildung bzw. einer Matrix  $A$  (welche ja einer linearen Abbildung entspricht). Wie aber bestimmt man eine invertierbare Matrix  $P$ , sodass  $P^{-1}AP$  eine Diagonalmatrix ist?

**Satz 4.2.18:** Sei  $A$  eine diagonalisierbare  $n \times n$  Matrix. Seien  $\lambda_1, \dots, \lambda_n$  die Eigenwerte von  $A$  (dabei komme in dieser Liste jeder Eigenwert so oft vor, wie seine algebraische Vielfachheit beträgt), und seien  $p_1, \dots, p_n \in K_n$  so, dass für jeden Index  $i$  der Vektor

$p_i$  ein Eigenvektor zum Eigenwert  $\lambda_i$  ist, und dass die Vektoren  $p_1, \dots, p_n$  überdies eine Basis von  $K_n$  bilden. Sei  $P$  diejenige  $n \times n$  Matrix, deren  $i$ -te Spalte  $p_i$  ist. Dann gilt

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ & & \ddots \\ 0 & & & \lambda_n \end{pmatrix}.$$

*Beweis:* Sei  $D$  die Diagonalmatrix, welche so wie oben in der Diagonale die Eigenwerte enthält. Die Beziehung  $P^{-1}AP = D$  können wir auch auf folgende Weise schreiben:  $AP = PD$ . Für die Matrix  $P$  bestehend aus den Eigenvektoren so wie im Satz, gilt für jede Spalte:

$$Ap_i = \lambda_i p_i.$$

Insgesamt gilt somit  $AP = PD$ , und da die gewählten Eigenvektoren linear unabhängig sind, ist die Matrix  $P$  auch invertierbar.  $\square$

**Beispiel 4.2.19:** Wir betrachten die Matrix

$$A = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}.$$

Das charakteristische Polynom von  $A$  ist

$$c_A(\lambda) = -(\lambda - 4)(\lambda + 2)^2.$$

Die Matrix  $A$  hat also den Eigenwert 4 mit algebraischer Vielfachheit 1 und den Eigenwert -2 mit algebraischer Vielfachheit 2. So wie im Satz 4.2.18 seien also

$$\lambda_1 = 4, \quad \lambda_2 = -2, \quad \lambda_3 = -2.$$

Als Eigenvektoren wählen wir (so wie im Satz 4.2.18 angegeben)

$$p_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \quad p_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad p_3 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

Dann gilt für die Matrix  $P = (p_1, p_2, p_3)$

$$P^{-1}AP = \begin{pmatrix} 4 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

Wir erreichen also die Diagonalisierung mittels einer Matrix, deren Spalten eine Basis von  $\mathbb{C}^3$  sind bestehend aus Eigenvektoren von  $A$ .  $\square$

Einige Anwendungsbeispiele:

**Beispiel 4.2.20 (Äquilibrium in einem System):** Wir betrachten die Entwicklung der Arbeitslosigkeit in einer Volkswirtschaft. Dazu nehmen wir an, dass 75% der Personen, welche am Beginn eines Jahres arbeitslos sind, im Laufe dieses Jahres Arbeit finden, und dass 5% der Personen, welche am Beginn eines Jahres Arbeit haben, im Laufe dieses Jahres arbeitslos werden. Wir betrachten die Entwicklung der Arbeitslosenrate über die Jahre hinweg. Am Beginn der Studie, also am Ende des 0-ten Jahres, sei  $L_0$  der Anteil der Bevölkerung, welche keine Arbeit hat, und  $M_0 = 1 - L_0$  der Anteil der Bevölkerung mit Arbeit. In analoger Weise für die Folgejahre. Also

$L_i$  ..... Anteil der Arbeitslosen am Ende des  $i$ -ten Jahres  
 $M_i$  ..... Anteil der Arbeitenden am Ende des  $i$ -ten Jahres

Der Übergang von einem Jahr zum nächsten drückt sich also aus als

$$\begin{pmatrix} L_{i+1} \\ M_{i+1} \end{pmatrix} = \underbrace{\begin{pmatrix} \frac{1}{4} & \frac{1}{20} \\ \frac{3}{4} & \frac{19}{20} \end{pmatrix}}_U \begin{pmatrix} L_i \\ M_i \end{pmatrix} .$$

Ausgehend von den Startwerten gilt dann für ein Jahr  $k$ :

$$\begin{pmatrix} L_k \\ M_k \end{pmatrix} = \begin{pmatrix} \frac{1}{4} & \frac{1}{20} \\ \frac{3}{4} & \frac{19}{20} \end{pmatrix}^k \begin{pmatrix} L_0 \\ M_0 \end{pmatrix} .$$

Wegen

$$\lim_{k \rightarrow \infty} \begin{pmatrix} \frac{1}{4} & \frac{1}{20} \\ \frac{3}{4} & \frac{19}{20} \end{pmatrix}^k = \lim_{k \rightarrow \infty} \frac{1}{16} \begin{pmatrix} 1 + \frac{15}{5^k} & 1 - \frac{1}{5^k} \\ 15 - \frac{15}{5^k} & 15 + \frac{1}{5^k} \end{pmatrix} = \frac{1}{16} \begin{pmatrix} 1 & 1 \\ 15 & 15 \end{pmatrix}$$

gilt also

$$\begin{pmatrix} L_k \\ M_k \end{pmatrix} \xrightarrow{k \rightarrow \infty} \frac{1}{16} \begin{pmatrix} 1 & 1 \\ 15 & 15 \end{pmatrix} \begin{pmatrix} L_0 \\ M_0 \end{pmatrix} = \begin{pmatrix} \frac{1}{16} \\ \frac{15}{16} \end{pmatrix} .$$

Dieses System konvergiert also unabhängig von den Ausgangswerten  $L_0, M_0$  zu einem Äquilibrium.

Wie aber sind haben wir oben die  $k$ -te Potenz der Übergangsmatrix berechnen können? Diese Überlegung wollen wir nun nachholen. Hat man einmal die Vermutung für die Formel von  $U^k$ , so lässt sich die Korrektheit mittels Induktion leicht nachweisen. Wie aber kommt man zu dieser Vermutung?

Die Eigenwerte von  $U$  sind die Nullstellen des charakteristischen Polynoms

$$c_U(\lambda) = \left(\frac{1}{4} - \lambda\right)\left(\frac{19}{20} - \lambda\right) - \frac{3}{80} = (5\lambda - 1)(\lambda - 1) = 0 .$$

Die Eigenwerte sind also  $\lambda_1 = \frac{1}{5}$  und  $\lambda_2 = 1$ . Zugehörige Eigenvektoren sind etwa  $v_1 = (1, -1)^T$  und  $v_2 = (1, 15)^T$ . Wegen Satz 4.2.18 bewirkt also die Matrix  $P = (v_1, v_2)$  eine Ähnlichkeitstranformation in eine Diagonalmatrix, nämlich

$$\underbrace{\frac{1}{16} \begin{pmatrix} 15 & -1 \\ 1 & 1 \end{pmatrix}}_{P^{-1}} \cdot \underbrace{\frac{1}{20} \begin{pmatrix} 5 & 1 \\ 15 & 19 \end{pmatrix}}_U \cdot \underbrace{\begin{pmatrix} 1 & 1 \\ -1 & 15 \end{pmatrix}}_P = \begin{pmatrix} \frac{1}{5} & 0 \\ 0 & 1 \end{pmatrix} .$$

Somit können wir für beliebiges  $k \in \mathbb{N}$  die  $k$ -te Potenz der Übergangsmatrix  $U$  einfach berechnen:

$$\begin{aligned} U^k &= P \cdot \begin{pmatrix} \frac{1}{5} & 0 \\ 0 & 1 \end{pmatrix}^k \cdot P^{-1} \\ &= \frac{1}{16} \begin{pmatrix} 1 & 1 \\ -1 & 15 \end{pmatrix} \begin{pmatrix} \frac{1}{5^k} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 15 & -1 \\ 1 & 1 \end{pmatrix} \\ &= \frac{1}{16} \begin{pmatrix} 1 + \frac{15}{5^k} & 1 - \frac{1}{5^k} \\ 15 - \frac{15}{5^k} & 15 + \frac{1}{5^k} \end{pmatrix}. \end{aligned}$$

**Beispiel 4.2.21 (Fibonacci-Zahlen):** Leonardo von Pisa (1175–1250), besser bekannt als Fibonacci (Sohn des Bonaccio, filius Bonacii) betrachtete die rekursive Folge

$$a_0 = 0, \quad a_1 = 1, \quad \text{und} \quad a_{n+2} = a_{n+1} + a_n \quad \text{für} \quad n \geq 0.$$

Mittels der Hilfsfolge  $(b_n)_{n \geq 1}$ , definiert als  $b_n := a_{n-1}$ , können wir das Bildungsgesetz der Fibonacci-Folge auch schreiben als ein System von Differenzgleichungen der Form

$$\begin{pmatrix} a_{n+2} \\ b_{n+2} \end{pmatrix} = \begin{pmatrix} a_{n+1} + b_{n+1} \\ a_{n+1} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}}_A \begin{pmatrix} a_{n+1} \\ b_{n+1} \end{pmatrix}.$$

Wir bestimmen die Eigenwerte von  $A$ :

$$c_A(\lambda) = \lambda^2 - \lambda - 1,$$

also

$$\lambda_1 = \frac{1}{2}(1 + \sqrt{5}), \quad \lambda_2 = \frac{1}{2}(1 - \sqrt{5}).$$

Zugehörige Eivenektoren sind

$$\text{zu } \lambda_1 : \begin{pmatrix} 1 \\ -\lambda_2 \end{pmatrix}, \quad \text{zu } \lambda_2 : \begin{pmatrix} 1 \\ -\lambda_1 \end{pmatrix}.$$

Mittels der Matrix  $P$  bestehend aus diesen Eigenvektoren, also

$$P = \begin{pmatrix} 1 & 1 \\ -\lambda_2 & -\lambda_1 \end{pmatrix}, \quad P^{-1} = \frac{1}{\lambda_2 - \lambda_1} \begin{pmatrix} -\lambda_1 & -1 \\ \lambda_2 & 1 \end{pmatrix},$$

können wir also  $A$  diagonalisieren:

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Die Potenzen  $A^n$  der Matrix  $A$  sind also leicht berechenbar. Wir benutzen dazu noch den Zusammenhang  $\lambda_1 \lambda_2 = -1$ . Für beliebiges  $n \in \mathbb{N}$  gilt also:

$$\begin{aligned} A^n &= \frac{1}{\lambda_2 - \lambda_1} \begin{pmatrix} 1 & 1 \\ -\lambda_2 & -\lambda_1 \end{pmatrix} \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix} \begin{pmatrix} -\lambda_1 & -1 \\ \lambda_2 & 1 \end{pmatrix} \\ &= \frac{1}{\lambda_2 - \lambda_1} \begin{pmatrix} \lambda_2^{n+1} - \lambda_1^{n+1} & \lambda_2^n - \lambda_1^n \\ \lambda_2^n - \lambda_1^n & \lambda_2^{n-1} - \lambda_1^{n-1} \end{pmatrix}. \end{aligned}$$

Aus den Anfangswerten  $b_1 = a_0 = 0$  und  $a_1 = 1$  ergibt sich somit

$$\begin{pmatrix} a_{n+1} \\ b_{n+1} \end{pmatrix} = A^n \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} = A^n \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\lambda_2 - \lambda_1} \begin{pmatrix} \lambda_2^{n+1} - \lambda_1^{n+1} \\ \lambda_2^n - \lambda_1^n \end{pmatrix},$$

und schliesslich

$$a_n = \frac{1}{\lambda_2 - \lambda_1} (\lambda_2^n - \lambda_1^n) = \frac{1}{2^n \sqrt{5}} \left( (1 + \sqrt{5})^n - (1 - \sqrt{5})^n \right).$$

Somit haben wir eine explizite Formel für die Fibonacci-Folge ermittelt.  $\square$

**Beispiel 4.2.22 (Systeme gewöhnlicher Differentialgleichungen mit konstanten Koeffizienten):** Seien  $y_1(t), \dots, y_n(t)$  differenzierbare Funktionen von  $\mathbb{R}$  nach  $\mathbb{R}$ , bzw. von  $\mathbb{C}$  nach  $\mathbb{C}$ . Wir schreiben diese Funktionen als Vektor

$$y(t) = \begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix},$$

$y(t)$  ist also eine Funktion von  $\mathbb{R}$  nach  $\mathbb{R}^n$ , bzw. von  $\mathbb{C}$  nach  $\mathbb{C}^n$ . Die Ableitung  $y'(t) = \frac{dy}{dt}(t)$  von  $y$  nach  $t$  ist komponentenweise definiert als

$$y'(t) = \begin{pmatrix} y'_1(t) \\ \vdots \\ y'_n(t) \end{pmatrix}.$$

Für eine gegebene Matrix  $A$  in  $\text{Mat}_{n \times n}(\mathbb{R})$  bzw. in  $\text{Mat}_{n \times n}(\mathbb{C})$  ist

$$y'(t) = Ay(t) \quad (*)$$

ein **homogenes System linearer gewöhnlicher Differentialgleichungen (hSLGDG) mit konstanten Koeffizienten**. Eine Lösung dieses Systems ist eine differenzierbare Funktion, welche  $(*)$  erfüllt. Die Menge aller Lösungen eines solchen hSLGDG ist ein linearer Teilraum des Vektorraums aller differenzierbaren Funktionen von  $\mathbb{R}$  nach  $\mathbb{R}^n$ , bzw. von  $\mathbb{C}$  nach  $\mathbb{C}^n$ . Man kann zeigen, dass dieser Lösungsraum die Dimension  $n$  hat. Eine Basis

$$(\varphi_1(t), \dots, \varphi_n(t))$$

des Lösungsraumes nennt man auch ein **Fundamentalsystem** des hSLGDG  $(*)$ . Es gilt die folgende Beziehung:

$n$  Lösungen  $(\varphi_1, \dots, \varphi_n)$  von  $(*)$  bilden ein Fundamentalsystem

$$\iff$$

$$\det(\varphi_1(t_0), \dots, \varphi_n(t_0)) \neq 0 \text{ für ein } t_0.$$

Hat nun die Matrix  $A$  den Eigenwert  $\lambda$  mit zugehörigem Eigenvektor  $x$ , dann ist

$$\varphi(t) = xe^{\lambda t}$$

eine Lösung des hSLGDG  $(*)$ , denn

$$\varphi'(t) = \lambda xe^{\lambda t} = Axe^{\lambda t} = A\varphi(t).$$

Tatsächlich gilt folgender Zusammenhang (der typischerweise in Vorlesungen zu Differentialgleichungen bewiesen wird):

**Satz 4.2.23:** *Besitzt die Matrix  $A$  eine Basis von Eigenvektoren  $x_1, \dots, x_n$  zu den Eigenwerten  $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ , so bilden die Funktionen  $\varphi_k(t) = x_k e^{\lambda_k t}$  für  $1 \leq k \leq n$  ein Fundamentalsystem des hSLGDG (\*).*

Konkret betrachten wir das Differentialgleichungssystem

$$y'(t) = \underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}}_A y(t).$$

Die Matrix  $A$  hat das charakteristische Polynom

$$c_A(\lambda) = \lambda^2(\lambda - 3).$$

Der Eigenwert  $\lambda_1 = \lambda_2 = 0$  hat also die algebraische Vielfachheit 2, der Eigenwert  $\lambda_3 = 3$  hat die algebraische Vielfachheit 1. Eigenvektoren  $x_i$  zu den Eigenwerten  $\lambda_i$  sind

$$x_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad x_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad x_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Diese Eigenvektoren  $(x_1, x_2, x_3)$  bilden offensichtlich eine Basis von  $\mathbb{R}^3$ . Wegen des obigen Satzes erhalten wir daraus das folgende Fundamentalsystem für den Lösungsraum:

$$\varphi_1(t) = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} e^{0t} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \varphi_2(t) = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} e^{0t} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad \varphi_3(t) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} e^{3t}.$$

Die allgemeine Lösung des Differentialgleichungssystems lautet also

$$y(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{pmatrix} = a \cdot \varphi_1(t) + b \cdot \varphi_2(t) + c \cdot \varphi_3(t) = \begin{pmatrix} a + b + ce^{3t} \\ -b + ce^{3t} \\ -a + ce^{3t} \end{pmatrix},$$

wobei  $a, b, c$  beliebige Konstante sind.

Solche homogenen Systeme linearer gewöhnlicher Differentialgleichungen mit konstanten Koeffizienten lassen sich also vollständig lösen mittels Methoden der linearen Algebra.

Wir betrachten noch ein zweites konkretes Differentialgleichungssystem, nämlich

$$\begin{pmatrix} y_1'(t) \\ y_2'(t) \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix}.$$

Die Koeffizientenmatrix dieses Systems hat die Eigenwerte

$$\lambda_1 = i, \quad \lambda_2 = -i$$

mit den zugehörigen Eigenvektoren

$$x_1 = \begin{pmatrix} 1 \\ \frac{1}{2} - \frac{1}{2}i \end{pmatrix}, \quad x_2 = \begin{pmatrix} 1 \\ \frac{1}{2} + \frac{1}{2}i \end{pmatrix}.$$



Als Fundamentalsystem für den Lösungsraum können wir also etwa wählen

$$\varphi_1(t) = \begin{pmatrix} 1 \\ \frac{1}{2} - \frac{1}{2}i \end{pmatrix} e^{it}, \quad \varphi_2(t) = \begin{pmatrix} 1 \\ \frac{1}{2} + \frac{1}{2}i \end{pmatrix} e^{-it}.$$

Natürlich ist dann auch

$$(\psi_1(t), \psi_2(t)) = \left(-\frac{i}{2}\varphi_1(t) + \frac{i}{2}\varphi_2(t), \frac{1}{2}\varphi_1(t) + \frac{1}{2}\varphi_2(t)\right)$$

ein Fundamentalsystem, welches wir unter Verwendung der Zusammenhänge

$$\begin{aligned} e^{it} &= \cos t + i \sin t & \text{bzw.} & \quad \cos t = \frac{1}{2}(e^{it} + e^{-it}) \\ e^{-it} &= \cos t - i \sin t & & \quad \sin t = \frac{-i}{2}(e^{it} - e^{-it}) \end{aligned}$$

auch schreiben können als

$$\psi_1(t) = \begin{pmatrix} \sin t \\ -\frac{1}{2} \cos t + \frac{1}{2} \sin t \end{pmatrix}, \quad \psi_2(t) = \begin{pmatrix} \cos t \\ \frac{1}{2} \cos t + \frac{1}{2} \sin t \end{pmatrix}.$$

Somit hat die allgemeine Lösung des Differentialgleichungssystems die Gestalt

$$\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = \begin{pmatrix} a \sin t + b \cos t \\ \frac{a}{2}(\sin t - \cos t) + \frac{b}{2}(\sin t + \cos t) \end{pmatrix},$$

wobei  $a, b$  beliebige Konstante sind. □

### Das Minimalpolynom einer Matrix

Im Vektorraum  $\text{Mat}_{n \times n}(K)$  sei für alle  $1 \leq i, j \leq n$  die Matrix  $E_{ij}$  so, dass sie überall 0 enthält, ausser in der  $i$ -ten Zeile und  $j$ -ten Spalte, wo der Eintrag 1 ist. Die Menge

$$\{E_{ij} \mid 1 \leq i, j \leq n\}$$

ist offensichtlich linear unabhängig und spannt ausserdem den ganzen Raum  $\text{Mat}_{n \times n}(K)$  auf. Der Vektorraum der  $n \times n$  Matrizen hat also die Dimension  $n^2$ . Das bedeutet auch, dass jede Menge von  $n^2 + 1$  Matrizen oder mehr linear abhängig ist. Insbesondere ist für eine beliebige  $n \times n$  Matrix  $A$  die Menge der Potenzen

$$\{A^0 = I_n, A, A^2, \dots, A^{n^2}\} = \{A^k \mid 0 \leq k \leq n^2\}$$

linear abhängig, d.h. es gibt ein Polynom vom Grad  $n^2$

$$p(x) = a_0 + a_1x + \dots + a_{n^2}x^{n^2} \in K[x],$$

sodass  $p(A) = 0$ . Analog gilt das natürlich auch für jede lineare Abbildung  $f \in \text{Hom}_K(V, V)$ , wobei  $V$  ein Vektorraum der Dimension  $n$  ist; denn nach Satz 3.3.31 ist  $\text{Hom}_K(V, V) \cong \text{Mat}_{n \times n}(K)$ .

Diese Schranke für die lineare Abhängigkeit der Potenzen einer quadratischen Matrix kann aber deutlich verbessert werden. Es gibt nämlich ein Polynom  $p(x)$  vom Grad

höchstens  $n$ , sodass  $p(A) = 0$ . Dieser Satz geht zurück auf Cayley (Arthur Cayley, 1821–1895) und Hamilton (William R. Hamilton, 1805–1865).

**Satz 4.2.24 (Cayley-Hamilton):** *Jede quadratische Matrix  $A$  ist Nullstelle ihres charakteristischen Polynoms, also  $c_A(A) = 0$ .*

*Beweis:* Sei  $B := A - \lambda I_n$ , und sei

$$c_A(\lambda) = |A - \lambda I| = |B| = c_0 + c_1\lambda + \cdots + c_n\lambda^n$$

das charakteristische Polynom der quadratischen Matrix  $A$ . Die Adjungierte zur Matrix  $B$ ,  $\text{adj}(B)$ , ist per Definition eine  $n \times n$  Matrix, deren Elemente Polynome in  $\lambda$  sind vom Grad höchstens  $n - 1$  (vgl. Def.4.1.19). Es gibt also Matrizen  $B_0, B_1, \dots, B_{n-1} \in \text{Mat}_{n \times n}(K)$ , sodass wir  $B$  ausdrücken können als

$$\text{adj}(B) = B_0 + B_1\lambda + \cdots + B_{n-1}\lambda^{n-1} .$$

In Satz 4.1.21 haben wir gezeigt, dass  $B \cdot \text{adj}(B) = |B| \cdot I_n$  gilt. Daraus ergibt sich

$$|B| \cdot I_n = B \cdot \text{adj}(B) = (A - \lambda I_n) \cdot \text{adj}(B) = A \cdot \text{adj}(B) - \lambda \cdot \text{adj}(B),$$

und weiters die polynomiale Beziehung

$$c_0 I + c_1 I \lambda + \cdots + c_n I \lambda^n = AB_0 + \cdots + AB_{n-1} \lambda^{n-1} - B_0 \lambda - \cdots - B_{n-1} \lambda^n .$$

Polynome sind nur dann gleich, wenn ihre jeweiligen Koeffizienten gleich sind (Koeffizientenvergleich), also muss gelten

$$\begin{aligned} c_0 I &= AB_0 \\ c_1 I &= AB_1 - B_0 \\ &\vdots \\ c_{n-1} I &= AB_{n-1} - B_{n-2} \\ c_n I &= -B_{n-1} . \end{aligned}$$

Nun multiplizieren wir die erste Gleichung mit  $A^0 = I$ , die zweite mit  $A^1$ , die dritte mit  $A^2$  und so fort. Wir erhalten

$$\begin{aligned} c_0 A^0 &= AB_0 \\ c_1 A^1 &= A^2 B_1 - AB_0 \\ &\vdots \\ c_{n-1} A^{n-1} &= A^n B_{n-1} - A^{n-1} B_{n-2} \\ c_n A^n &= -A^n B_{n-1} . \end{aligned}$$

Durch Addition dieser Gleichung ergibt sich die gewünschte Gleichung  $c_A(A) = 0$ . □

Ein vom Nullpolynom verschiedenes Polynom  $p(x)$  heisst **monisch**, wenn sein führender Koeffizient, also der Koeffizient des höchsten Terms  $x^k$  in  $p$ , 1 ist.

**Satz 4.2.25:** *Sei  $k$  der kleinste Grad eines Polynoms, welches  $A$  als Nullstelle hat. Dann gibt es genau ein monisches Polynom  $m(x)$  vom Grad  $k$ , also von der Form  $m(x) = x^k + m_{k-1}x^{k-1} + \cdots + m_0$ , welches  $A$  als Nullstelle hat.*

*Beweis:* Wir nehmen an,  $r(x)$  wäre ebenfalls ein monisches Polynom vom Grad  $k$  mit  $r(A) = 0$ . Dann ist  $m - r$  ein Polynom vom Grad höchstens  $k - 1$  mit  $A$  als Nullstelle:  $(m - r)(A) = m(A) - r(A) = 0$ . Die Polynome  $m$  und  $r$  müssen also gleich sein.  $\square$

**Definition 4.2.26:** Sei  $A \in \text{Mat}_{n \times n}(K)$ . Das eindeutig bestimmte monische Polynom kleinsten Grades  $m_A(x)$  mit  $m_A(A) = 0$  heisst das **Minimalpolynom** der Matrix  $A$ .

**Satz 4.2.27:** Ist  $p(x)$  ein Polynom mit  $p(A) = 0$ , dann teilt das Minimalpolynom  $m_A(x)$  das Polynom  $p(x)$ .

*Beweis:* Wir dividieren  $p(x)$  durch  $m_A(x)$  (etwa mittels des Euklidischen Algorithmus) und erhalten die folgende eindeutige Darstellung mittels Quotienten  $q(x)$  und Rest  $r(x)$ :

$$p(x) = q(x)m_A(x) + r(x), \quad \text{mit } r(x) = 0 \text{ oder } \text{grad}(r) < \text{grad}(m_A).$$

Nun gilt sowohl  $p(A) = 0$  als auch  $m_A(A) = 0$ . Also muss auch gelten  $r(A) = 0$ . Das ist aber nur möglich für  $r = 0$ , also falls  $m_A(x)$  das Polynom  $p(x)$  teilt.  $\square$

**Korollar:** Das Minimalpolynom  $m_A(x)$  ist ein Teiler des charakteristischen Polynoms  $c_A(x)$ .

*Beweis:* das folgt unmittelbar aus Satz 4.2.24 und Satz 4.2.27.  $\square$

**Satz 4.2.28:** Das Minimalpolynom  $m_A(x)$  und das charakteristische Polynom  $c_A(x)$  haben dieselben Nullstellen in  $K$ .

*Beweis:* Sei  $\lambda \in K$  eine Nullstelle von  $c_A(x)$ , also ein Eigenwert der Matrix  $A$ . Es gibt also einen vom Nullvektor verschiedenen Vektor  $v$ , sodass  $Av = \lambda v$ .

Für ein beliebiges Polynom  $a(x) \in \text{Pol}(K)$ ,

$$a(x) = a_0 + a_1x + \dots + a_kx^k,$$

gilt

$$a(A)v = a_0v + a_1Av + \dots + a_kA^k v = a_0v + a_1\lambda v + \dots + a_k\lambda^k v = a(\lambda)v.$$

Also ist  $a(\lambda)$  ein Eigenwert von  $a(A)$ , d.h.  $a(\lambda)$  ist eine Nullstelle von  $c_{a(A)}(x)$ .

All das gilt also auch für den Fall des Minimalpolynoms, also  $a(x) = m_A(x)$ . Für jede Nullstelle  $\lambda$  von  $c_A(x)$  ist also  $m_A(\lambda)$  eine Nullstelle von

$$c_{m_A(A)}(x) = c_0(x) = |-xI_n| = (-1)^n x^n.$$

(Dabei ist der zweite Ausdruck das charakteristische Polynom der Nullmatrix.) Das Polynom auf der rechten Seite hat nur 0 als Nullstelle. Daraus sehen wir  $m_A(\lambda) = 0$ . Somit ist  $\lambda$  eine Nullstelle von  $m_A(x)$ .  $\square$

**Beispiel 4.2.29:** (a) Das charakteristische Polynom der Matrix

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

ist  $c_A(x) = (x-2)^3$ . Sowohl  $A-2I_3$  als auch  $(A-2I_3)^2$  sind verschieden von der Nullmatrix. Somit ist das charakteristische Polynom gleich dem Minimalpolynom,  $m_A(x) = c_A(x)$ .

(b) Das charakteristische Polynom der Matrix

$$B = \begin{pmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{pmatrix}$$

ist  $c_B(x) = (x - 1)(x - 2)^2$ . Wegen Satz 4.2.28 ist das Minimalpolynom von  $B$  entweder

$$(x - 1)(x - 2) \quad \text{oder} \quad (x - 1)(x - 2)^2 .$$

Es gilt  $(A - I)(A - 2I) = 0$ , also ist  $m_B(x) = (x - 1)(x - 2)$ . □

**Satz 4.2.30:** Eine quadratische Matrix ist invertierbar genau dann, wenn der konstante Term in ihrem charakteristischen Polynom verschieden von 0 ist.

*Beweis:* Ist  $A$  invertierbar, dann ist wegen Satz 4.2.2 0 kein Eigenwert von  $A$ , also  $c_A(0) \neq 0$ . Das bedeutet, dass der konstante Term in  $c_A(x)$  verschieden von 0 ist.

Andererseits nehmen wir an, dass der konstante Term in  $c_A(x)$  verschieden von 0 ist. Aus dem Satz von Cayley-Hamilton wissen wir  $c_A(A) = 0$ , was also mit obiger Hypothese geschrieben werden kann als

$$A \cdot p(A) = I_n ,$$

für ein gewisses Polynom  $p(x)$ .

Denn  $0 = c_A(A) = c_0I + c_1A + \dots + c_nA^n$  und  $c_0 \neq 0$ , also  $I = A \cdot (-\frac{1}{c_0}q(A)) = A \cdot p(A)$ . Daraus sehen wir, dass  $A$  invertierbar ist. □

**Beispiel 4.2.31:** Die Matrix

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

hat das charakteristische Polynom  $c_A(x) = (x - 1)^3$ . Aus dem Satz von Cayley-Hamilton sehen wir also

$$0 = (A - I_3)^3 = A^3 - 3A^2 + 3A - I_3 ,$$

und weiter

$$A(A^2 - 3A + 3I_3) = I_3 .$$

Also gilt

$$A^{-1} = A^2 - 3A + 3I_3 = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

für die inverse Matrix. □

### 4.3 Resultanten

Mittels Resultanten kann man bestimmen, ob zwei Polynome in  $K[x]$  einen gemeinsamen Faktor haben, also in  $\bar{K}$  (dem algebraischen Abschluss von  $K$ ) eine gemeinsame Nullstelle haben.

**Lemma 4.3.1:** Seien  $a, b \in K[x]$  mit  $\text{grad}(a) = l > 0$  und  $\text{grad}(b) = m > 0$ . Dann haben  $a$  und  $b$  einen nicht-trivialen gemeinsamen Faktor (also  $\text{ggT}(a, b)$  ist nicht konstant) g.d.w. es Polynome  $c, d \in K[x]$  gibt, sodass:

- (i)  $c$  und  $d$  sind nicht beide 0,
- (ii)  $\text{grad}(c) \leq m - 1$  und  $\text{grad}(d) \leq l - 1$ , und
- (iii)  $c \cdot a + d \cdot b = 0$ .

*Beweis:* Zunächst nehmen wir an, dass  $a$  und  $b$  einen gemeinsamen Faktor  $f \in K[x]$  haben. Dann gibt es  $\tilde{a}, \tilde{b} \in K[x]$ , sodass

$$a = f \cdot \tilde{a}, \quad \text{und} \quad b = f \cdot \tilde{b}.$$

Offensichtlich gilt

$$\text{grad}(\tilde{a}) \leq l - 1 \quad \text{und} \quad \text{grad}(\tilde{b}) \leq m - 1.$$

Weiters

$$\tilde{b} \cdot a + (-\tilde{a}) \cdot b = \tilde{b} \cdot f\tilde{a} - \tilde{a} \cdot f\tilde{b} = 0,$$

also haben  $c = \tilde{b}$  und  $d = -\tilde{a}$  die gewünschten Eigenschaften.

Andererseits nehmen wir an, dass  $c$  und  $d$  die Eigenschaften (i)–(iii) besitzen. O.B.d.A. sei  $d \neq 0$ . Haben  $a$  und  $b$  keinen nicht-trivialen gemeinsamen Faktor, dann ist  $\text{ggT}(a, b) = 1$ . Mittels des Erweiterten Euklidischen Algorithmus können wir also Polynome  $\tilde{c}, \tilde{d}$  finden, sodass  $\tilde{c}a + \tilde{d}b = 1$ . Multiplizieren wir diese Gleichung mit  $d$  und benutzen  $db = -ca$ , so erhalten wir

$$d = (\tilde{c}a + \tilde{d}b)d = \tilde{c}da + \tilde{d}db = \tilde{c}da - \tilde{d}ca = (\tilde{c}d - \tilde{d}c)a.$$

Da  $d \neq 0$ , sehen wir daraus  $\text{grad}(d) \geq l$ , im Widerspruch zu (ii). Es muss also einen nicht-trivialen gemeinsamen Faktor von  $a$  und  $b$  geben.  $\square$

Um die Existenz solcher Polynome  $c$  und  $d$  in Lemma 4.3.1 zu entscheiden, verwenden wir Methoden der Linearen Algebra. Die Bedingung “ $c \cdot a + d \cdot b = 0$ ” lässt sich nämlich schreiben als System linearer Gleichungen in den Koeffizienten. Dazu sei

$$\begin{aligned} a &= a_0x^l + \dots + a_l, & a_0 &\neq 0, \\ b &= b_0x^m + \dots + b_m, & b_0 &\neq 0, \\ c &= c_0x^{m-1} + \dots + c_{m-1}, \\ d &= d_0x^{l-1} + \dots + c_{l-1}. \end{aligned}$$

Die Koeffizienten von  $c, d$  sind gesucht. Die Bedingung “ $c \cdot a + d \cdot b = 0$ ” führt dann zu folgendem Gleichungssystem in den Koeffizienten:

$$(*) \quad \begin{array}{rccccccc} a_0c_0 & & + & b_0d_0 & & = & 0 & \text{Koeff. von } x^{l+m-1} \\ a_1c_0 + a_0c_1 & & + & b_1d_0 + b_0d_1 & & = & 0 & \text{Koeff. von } x^{l+m-2} \\ \vdots & & & \vdots & & & \vdots & \\ & & a_lc_{m-1} & + & b_md_{l-1} & = & 0 & \text{Koeff. von } x^0 \end{array}$$

Dieses homogene lineare Gleichungssystem besteht aus  $l + m$  Gleichungen in  $l + m$  Variablen. Es besitzt also eine nichttriviale (nicht beide  $c, d$  gleich 0) Lösung g.d.w. die Determinante der Koeffizientenmatrix 0 ist. Diese Überlegungen motivieren die folgende Definition.

**Definition 4.3.2:** Seien  $a, b \in K[x]$  mit  $\text{grad}(a) = l > 0$  und  $\text{grad}(b) = m > 0$ , also

$$\begin{aligned} a &= a_0 x^l + \cdots + a_l, & a_0 &\neq 0, \\ b &= b_0 x^m + \cdots + b_m, & b_0 &\neq 0. \end{aligned}$$

Dann ist die **Sylvester-Matrix** von  $a$  und  $b$ , in Zeichen  $\text{Syl}(a, b)$ , die Koeffizientenmatrix des homogenen linearen Gleichungssystems (\*). Also  $\text{Syl}(a, b)$  ist die folgende  $(l + m) \times (l + m)$  Matrix, bestehend aus  $m$  Spalten der Koeffizienten von  $a$  und  $l$  Spalten der Koeffizienten von  $b$ :

$$\text{Syl}(a, b) = \begin{pmatrix} a_0 & & & & b_0 & & & & \\ a_1 & a_0 & & & b_1 & b_0 & & & \\ a_2 & a_1 & \cdots & & b_2 & b_1 & \cdots & & \\ \vdots & & \cdots & a_0 & \vdots & & \cdots & b_0 & \\ & \vdots & & a_1 & & \vdots & & b_1 & \\ a_l & & & & b_m & & & & \\ & a_l & & \vdots & b_m & & & \vdots & \\ & & \cdots & & & & \cdots & & \\ & & & a_l & & & & & b_m \end{pmatrix}$$

(alle anderen Eintragungen in  $\text{Syl}(a, b)$  sind 0).

Die **Resultante** von  $a$  und  $b$ , in Zeichen  $\text{Res}(a, b)$ , ist die Determinante der Sylvestermatrix, also

$$\text{Res}(a, b) = \det(\text{Syl}(a, b)).$$

Auf analoge Weise erhalten wir die Resultante bzgl  $x_1$  zweier multivariater Polynome  $a, b \in K[x_1, \dots, x_n]$  mit  $\text{grad}_1(a) = l > 0$  und  $\text{grad}_1(b) = m > 0$ , also

$$\begin{aligned} a &= a_0(x_2, \dots, x_n)x_1^l + \cdots + a_l(x_2, \dots, x_n), & a_0 &\neq 0, \\ b &= b_0(x_2, \dots, x_n)x_1^m + \cdots + b_m(x_2, \dots, x_n), & b_0 &\neq 0. \end{aligned}$$

Diese Resultante  $\text{Res}_{x_1}(a, b)$  hängt also nur von den Variablen  $x_2, \dots, x_n$  ab.  $\square$

**Satz 4.3.3:** Die Polynome  $a, b \in K[x]$  haben einen nicht-trivialen gemeinsamen Faktor, also  $\text{ggT}(a, b) \neq 1$ , g.d.w.  $\text{Res}(a, b) = 0$ .

*Beweis:* Die Resultante ist 0 g.d.w. das Gleichungssystem (\*) eine nicht-triviale Lösung hat. Das ist aber gleichbedeutend damit, dass es Polynome  $c, d$  wie in Lemma 4.3.1 gibt, und somit haben  $a$  und  $b$  einen nicht-trivialen gemeinsamen Faktor.  $\square$

**Satz 4.3.4:** Zu gegebenen Polynomen  $a, b \in K[x]$  mit positivem Grad gibt es  $c, d \in K[x]$ , sodass

$$\text{Res}(a, b) = c \cdot a + d \cdot b.$$

Weiters hängen sowohl  $\text{Res}(a, b)$  als auch die Koeffizienten von  $c$  und  $d$  ganzzahlig polynomial von den Koeffizienten von  $a$  und  $b$  ab.

*Beweis:* Offensichtlich hängt  $\text{Res}(a, b) = \det(\text{Syl}(a, b))$  ganzzahlig polynomial von den Koeffizienten von  $a, b$  ab.

Für die Aussage bzgl. der Koeffizienten von  $c$  und  $d$  siehe Proposition 9 auf Seite 152 in Cox, Little, O'Shea, *Ideals, Varieties, and Algorithms*, 2nd ed., Springer-Verlag (1997).  $\square$

**Satz 4.3.5:** Sei  $K$  ein algebraisch abgeschlossener Körper. Seien

$$a(x_1, \dots, x_n) = \sum_{i=1}^k a_i(x_1, \dots, x_{n-1})x_n^i, \quad b(x_1, \dots, x_n) = \sum_{i=1}^l b_i(x_1, \dots, x_{n-1})x_n^i$$

$n$ -variate Polynome in  $K[x_1, \dots, x_n]$  vom Grad  $k$  bzw.  $l$ , für  $k, l \geq 0$ . Sei

$$r(x_1, \dots, x_{n-1}) = \text{Res}_{x_n}(a, b).$$

(Die Resultante wird also bzgl. der Variablen  $x_n$  berechnet.)

Ist  $(\alpha_1, \dots, \alpha_n) \in K^n$  eine gemeinsame Nullstelle von  $a$  und  $b$ , dann ist  $(\alpha_1, \dots, \alpha_{n-1})$  eine Nullstelle von  $r$ .

Ist andererseits  $(\alpha_1, \dots, \alpha_{n-1})$  eine Nullstelle von  $r$ , dann ist entweder

(i)  $a_k(\alpha_1, \dots, \alpha_{n-1}) = b_l(\alpha_1, \dots, \alpha_{n-1}) = 0$ , oder

(ii) es gibt  $a_n \in K$ , sodass  $(\alpha_1, \dots, \alpha_n)$  eine gemeinsame Nullstelle von  $a$  und  $b$  ist.

*Beweis:* Siehe Theorem 4.3.3. auf Seite 98 in F.Winkler, *Polynomial Algorithms in Computer Algebra*, Springer-Verlag (1996).  $\square$

**Beispiel 4.3.6:** Wir betrachten das folgende System polynomialer (algebraischer) Gleichungen:

$$a_1(x, y, z) = a_2(x, y, z) = a_3(x, y, z) = 0,$$

wobei

$$\begin{aligned} a_1 &= 2xy + yz - 3z^2, \\ a_2 &= x^2 - xy + y^2 - 1, \\ a_3 &= yz + x^2 - 2z^2. \end{aligned}$$

Die Resultante bzgl.  $y$  von  $a_1$  und  $a_3$  können wir etwa berechnen als

$$\text{Res}_y(a_1, a_3) = \begin{vmatrix} 2x+z & z \\ -3z^2 & x^2 - 2z^2 \end{vmatrix} = 2x^3 + x^2z - 4xz^2 + z^3.$$

Wie im Satz 4.3.5 beschrieben, eliminieren wir einige Variable aus den Gleichungen:

$$\begin{aligned} b(x) &= \text{Res}_z(\text{Res}_y(a_1, a_3), \text{Res}_y(a_2, a_3)) \\ &= x^6(x-1)(x+1)(127x^4 - 167x^2 + 4), \\ c(y) &= \text{Res}_z(\text{Res}_x(a_1, a_3), \text{Res}_x(a_2, a_3)) \\ &= (y-1)^3(y+1)^3(3y^2-1)(127y^4 - 216y^2 + 81)(457y^4 - 486y^2 + 81), \\ d(z) &= \text{Res}_y(\text{Res}_x(a_1, a_2), \text{Res}_x(a_1, a_3)) \\ &= 5184 z^{10}(z-1)(z+1)(127z^4 - 91z^2 + 16). \end{aligned}$$

Alle gemeinsamen Nullstellen von  $a_1, a_2, a_3$  haben Koordinaten, welche Nullstellen von  $b, c, d$  sind, etwa  $(1, 1, 1)$ . Aber nicht jede Nullstelle von  $c$ , etwa  $1/\sqrt{3}$ , lässt sich erweitern zu einer gemeinsamen Nullstelle von  $a_1, a_2, a_3$ .  $\square$

Die Methode der Resultanten ist eine von mehreren konstruktiven Zugängen zur Eliminationstheorie von multivariaten Polynomen, also der Lösung von polynomialen (algebraischen) Gleichungssystemen in mehreren Variablen. Eine andere sehr potente Methode beruht auf Gröbner-Basen. Davon lernt man in der Computeralgebra.