Kapitel 3: Lineare Abbildungen, Matrizen und Polynome

§1 Der Rang einer Matrix

Voraussetzung wie üblich: K beliebiger Körper.
Wir wenden nun die Ergebnisse des letzten Kapitels auf Matrizen und lineare Gleichungssysteme an. Gegeben sei eine beliebige Matrix $A \in K^{m\times n}$.

- Sei $\mathsf{SR}(A) \subseteq K^m$ der Spaltenraum von A.
  Dann heißt $\dim \mathsf{SR}(A)$ der ***Spaltenrang*** von A.
- Sei $\mathsf{ZR}(A) \subseteq K^n$ der Zeilenraum von A.
  Dann heißt $\dim \mathsf{ZR}(A)$ der ***Zeilenrang*** von A.

Was wir bereits wissen (siehe §2):
Betrachte die lineare Abbildung $\varphi_A \colon K^n \to K^m$, $v \mapsto Av$. Dann ist

$$\mathsf{SR}(A) = \mathsf{Bild}(\varphi_A) = \{Av \mid v \in K^n\} \text{ und}$$
$$\mathsf{Kern}(\varphi_A) = N(A) := \{v \in K^n \mid Av = 0_m\}.$$

Also folgt mit Kern-Bild-Dimensionsformel: $\boxed{n = \dim N(A) + \dim \mathsf{SR}(A).}$

Bemerkung 1.1. Wie oben sei $A^{m\times n}$ gegeben.

Nach Kapitel 1, §2, ändert sich die Lösungsmenge $N(A) = \{x \in K^n \mid Ax = 0_m\}$ nicht bei elementaren Zeilenumformungen. Sei $A' \in K^{m\times n}$ in Zeilenstufenform und $A \to A'$ (Gauß-Algorithmus). Dann folgt also

$$\dim \mathsf{SR}(A) = n - \dim N(A) = n - \dim N(A') = \dim \mathsf{SR}(A').$$

d.h., der Spaltenrang ändert sich nicht bei elementaren Zeilenumformungen.

Beispiel 1.2. Wir betrachten das Beispiel aus Kapitel 1, §2:

$$A = \begin{bmatrix} 0 & 0 & 2 & 2 & 0 & 2 \\ 2 & 4 & 2 & 6 & 0 & 4 \\ 3 & 6 & 0 & 6 & 1 & 4 \end{bmatrix} \quad \to \quad A' = \begin{bmatrix} 1 & 2 & 0 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}.$$

Also $r = 3$ Stufen; Pivots $j_1 = 1$, $j_2 = 3$, $j_3 = 5$; freie Variablen x_2, x_4, x_6.

Zugehöriges homogenes Gleichungssystem:
$$\begin{aligned} x_1 + 2x_2 + 2x_4 + x_6 &= 0 \\ x_3 + x_4 + x_6 &= 0 \\ x_5 + x_6 &= 0 \end{aligned}.$$

Damit ist $N(A) = N(A') = \{v(x_2, x_4, x_5) \mid x_2, x_4, x_6 \in \mathbb{Q} \text{ beliebig}\}$, wobei

$$v(x_2, x_4, x_5) := \begin{bmatrix} -2x_2 - 2x_4 - x_6 \\ x_2 \\ -x_4 - x_6 \\ x_4 \\ -x_6 \\ x_6 \end{bmatrix} \qquad \left(\text{System} : \begin{aligned} x_1 + 2x_2 + 2x_4 + x_6 &= 0 \\ x_3 + x_4 + x_6 &= 0 \\ x_5 + x_6 &= 0 \end{aligned} \right)$$

Wie erhält man eine Basis für $N(A)$? Dazu bemerken wir, dass
$v(x_2, x_4, v_6) = x_2 v(1,0,0) + x_4 v(0,1,0) + x_6 v(0,0,1)$ gilt, mit

$$v(1,0,0) := \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad v(0,1,0) := \begin{bmatrix} -2 \\ 0 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \qquad v(0,0,1) := \begin{bmatrix} -1 \\ 0 \\ -1 \\ 0 \\ -1 \\ 1 \end{bmatrix}.$$

Behauptung: $v(1,0,0), v(0,1,0), v(0,0,1)$ sind l.u. Denn sind $s_1, s_2, s_3 \in \mathbb{Q}$, so ist $s_1 v(1,0,0) + s_2 v(0,1,0) + s_3 v(0,0,1) = v(s_1, s_2, s_3)$, wobei die 2. Komponente dieses Spaltenvektors gleich s_1 ist, die 4. Komponente gleich s_2 und die 6. Komponente gleich s_3. Gilt also $v(s_1, s_2, s_3) = 0_6$, so folgt $s_1 = s_2 = s_3 = 0$.

Also ist $B = \{v(1,0,0), v(0,1,0), v(0,0,1)\}$ eine Basis von $N(A)$ und damit $\dim \mathsf{SR}(A) = 6 - 3 = 3$. Dies können wir auch direkt anhand von $\dim \mathsf{SR}(A) = \dim \mathsf{SR}(A')$ feststellen. Dazu betrachte noch einmal

$$A' = \begin{bmatrix} 1 & 2 & 0 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \qquad \text{mit Pivots } j_1 = 1,\ j_2 = 3,\ j_3 = 5.$$

Die 1., 3., 5. Spalte von A' sind also gleich den Einheitsvektoren e_1, e_2, e_3. Damit ist $e_1, e_2, e_3 \in \mathsf{SR}(A') \subseteq \mathbb{Q}^3$, und weil $\{e_1, e_2, e_3\}$ eine Basis von $\mathbb{Q}^3$ ist, folgt $\mathsf{SR}(A') = \mathbb{Q}^3$ und $\dim \mathsf{SR}(A') = 3$.

Bemerkung 1.3. Sei $A' \in K^{m\times n}$ in Zeilenstufenform mit $r \geqslant 0$ Stufen.

Seien $1 \leqslant j_1 < \ldots < j_r \leqslant n$ die Pivots. Die zugehörigen Spalten von A' sind genau die Einheitsvektoren $e_1, \ldots, e_r \in K^m$. Jede andere Spalte von A' hat nur Einträge ungleich 0 in den ersten r Zeilen, also ist jede Spalte von A' eine Linearkombination von $e_1, \ldots e_r$. Damit folgt

$$\dim \mathsf{SR}(A') = \dim \langle e_1, \ldots, e_r \rangle_K = r \qquad \text{und} \qquad \dim N(A') = n - r.$$

Satz 1.4 (Rang-Satz). Sei $A \in K^{m\times n}$ gegeben.

Dann gilt $\dim \mathsf{SR}(A) = \dim \mathsf{ZR}(A)$; diese Zahl wird mit $\mathsf{Rang}(A)$ bezeichnet und heißt der ***Rang*** von A. Ist $A \to A'$ (Gauß-Algorithmus), wobei $A' \in K^{m\times n}$ in Zeilenstufenform mit r Stufen ist, so gilt $\mathsf{Rang}(A) = \mathsf{Rang}(A') = r \leqslant \min\{n, m\}$.

Beweis. Nach Bemerkungen 1.1 und 1.3 gilt $\dim \mathsf{SR}(A) = \dim \mathsf{SR}(A') = r \geqslant 0$. Nach Aufgabe 10.3 ändert sich der Zeilenraum nicht bei elementaren Zeilenumformungen. Also gilt $\dim \mathsf{ZR}(A) = \dim \mathsf{ZR}(A')$.
Seien $1 \leqslant j_1 < \ldots < j_r \leqslant n$ die Pivots für A'; für $1 \leqslant i \leqslant m$ sei $z_i \in K^{1\times n}$ die i-te Zeile von A'. Dann ist $\mathsf{ZR}(A') = \langle z_1, \ldots, z_r\rangle_K$ (weil $z_i = 0_n$ für $i > r$). Die Definition der Zeilenstufenform zeigt: Für $1 \leqslant i \leqslant r$ ist die j_i-te Komponente von z_i gleich 1, und gleich 0 in allen anderen Zeilen. Also ist in allen $v \in \langle z_1, \ldots, z_{i-1}\rangle_K$ die j_i-te Komponente gleich 0, und damit $z_i \notin \langle z_1, \ldots, z_{i-1}\rangle_K$ für $1 \leqslant i \leqslant r$. Nach Lemma 4.8 sind $(z_1, \ldots, z_r)$ linear unabhängig, also $\dim \mathsf{ZR}(A') = r$ und damit schließlich

$$\dim \mathsf{SR}(A) = \dim \mathsf{SR}(A') = r = \dim \mathsf{ZR}(A') = \dim \mathsf{ZR}(A). \qquad \square$$

Folgerung 1.5. Sei $A \in K^{m\times n}$ gegeben.

(a) Ist $m = n$, so gilt: $\quad \mathsf{Rang}(A) = n \;\Leftrightarrow\; A$ invertierbar.
(b) Ist $B \in K^{n\times p}$, so gilt $\mathsf{Rang}(AB) \leqslant \min\{\mathsf{Rang}(A), \mathsf{Rang}(B)\}$.

Beweis. (a) Wegen $n = \dim N(A) + \mathsf{Rang}(A)$ und mit Beispiel 4.10 folgt:
$n = \mathsf{Rang}(A) \Leftrightarrow \dim N(A) = 0 \Leftrightarrow \{x \in K^n \mid Ax = 0_n\} = \{0_n\} \Leftrightarrow A$ invertierbar.
(b) Sei $\varphi\colon K^n \to K^m$, $v \mapsto Av$. Dann ist $\mathsf{SR}(A) = \mathsf{Bild}(\varphi) = \{Av \mid v \in K^n\}$.
Analog gilt $\mathsf{SR}(AB) = \{ABw \mid w \in K^p\}$. Nun ist $Bw \in K^n$ für alle $w \in K^p$, also

$$\mathsf{SR}(AB) = \{ABw \mid w \in K^p\} \subseteq \{Av \mid v \in K^n\} = \mathsf{SR}(A),$$

d.h., $\mathsf{Rang}(AB) \leqslant \mathsf{Rang}(A)$. Andererseits gilt $N(B) \subseteq N(AB)$, denn ist $w \in N(B)$, d.h., $Bw = 0$, so folgt auch $ABw = 0_m$. Damit also $\dim N(B) \leqslant \dim N(AB)$. Mit

$$p = \dim N(B) + \mathsf{Rang}(B) = \dim N(AB) + \mathsf{Rang}(AB)$$

erhalten wir $\mathsf{Rang}(AB) = p - \dim N(AB) \leqslant p - \dim N(B) = \mathsf{Rang}(B)$. $\square$

Zum Schluss wenden wir dies auf lineare Gleichungssysteme an:

Satz 1.6. Sei $A \in K^{m\times n}$, $b \in K^m$ und $L := \{x \in K^n \mid Ax = b\}$.

(a) Es gilt: $\quad L \neq \varnothing \quad\Leftrightarrow\quad b \in \mathsf{SR}(A) \quad\Leftrightarrow\quad \mathsf{Rang}(A) = \mathsf{Rang}([A \mid b])$.

(b) Ist $L \neq \varnothing$, so sei $x_0 \in L$ eine feste Lösung. Dann gilt

$$L = \{x_0 + v \mid v \in N(A)\} \quad \text{wobei} \quad N(A) := \{v \in K^n \mid Av = 0_m\}$$

der Nullraum von A ist; es gilt $\dim N(A) = n - \mathsf{Rang}(A)$.

Beweis. (a) Seien $v_1, \ldots, v_n \in K^m$ die Spalten von A. Dann gilt $v_i = Ae_i$ wobei $\{e_1, \ldots, e_n\}$ die Einheitsvektoren in K^n sind. Sei $x \in K^n$ beliebig; dann ist $x = x_1e_1 + \ldots + x_ne_n$ mit $x_1, \ldots, x_n \in K$. Wie in Kapitel 2, Beispiel 2.8, sieht man sofort: $Ax = Ax_1e_1 + \ldots + Ax_ne_n = x_1v_1 + \ldots + x_nv_n$. Also gilt $Ax = b$ genau dann, wenn b eine Linearkombination von $v_1, \ldots, v_n$ ist, d.h., genau dann, wenn $b \in \mathsf{SR}(A)$ gilt. Damit gilt die erste Äquivalenz in (a).
Zur zweiten Äquivalenz: Nach Definition ist zunächst klar, dass $\mathsf{SR}(A) \subseteq \mathsf{SR}([A \mid b])$ gilt, also auch $\mathsf{Rang}(A) \leqslant \mathsf{Rang}([A \mid b])$.

Ist nun $\mathsf{Rang}(A) = \mathsf{Rang}([A \mid b])$, so folgt $\mathsf{SR}([A \mid b]) = \mathsf{SR}(A)$ (siehe Satz 4.11), also auch $b \in \mathsf{SR}([A \mid b]) = \mathsf{SR}(A)$.
Umgekehrt: Ist $b \in \mathsf{SR}(A) = \langle v_1, \ldots, v_n\rangle_K$, so folgt $\{v_1, \ldots, v_n, b\} \subseteq \mathsf{SR}(A)$, also auch $\mathsf{SR}([A \mid b]) = \langle v_1, \ldots, v_n, b\rangle_K \subseteq \mathsf{SR}(A)$, d.h., $\mathsf{SR}(A) = \mathsf{SR}([A \mid b])$.

(b) Sei nun $L \neq \varnothing$ und $x_0 \in L$ fest. Zu zeigen:

$$L = \{x_0 + v \mid v \in N(A)\} \quad \text{wobei} \quad N(A) := \{v \in K^n \mid Av = 0_m\}.$$

"$\subseteq$" Sei $x \in L$, also $Ax = b$. Dann ist $A(x - x_0) = Ax - Ax_0 = b - b = 0_m$, also $v := x - x_0 \in N(A)$, d.h., $x = x_0 + v$ mit $v \in N(A)$.
"$\supseteq$" Sei $x = x_0 + v$ mit $v \in N(A)$. Dann ist $Av - 0_m$ und damit $Ax - A(x_0 + v) = Ax_0 + Av = b + 0_m = b$, also $x \in L$. $\square$

Betrachten wir noch einmal das Beispiel aus Kapitel 1, §2:

$$[A \mid b] = \left[\begin{array}{ccccc|c} 0 & 0 & 2 & 2 & 0 & 2 \\ 2 & 4 & 2 & 6 & 0 & 4 \\ 3 & 6 & 0 & 6 & 1 & 4 \end{array}\right] \quad\to\quad [A' \mid b'] = \left[\begin{array}{ccccc|c} 1 & 2 & 0 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{array}\right].$$

Das neue Gleichungssystem ist also:
$$\begin{aligned} x_1 + 2x_2 + 2x_4 &= 1 \\ x_3 + x_4 &= 1 \\ x_5 &= 1 \end{aligned}$$

Es gilt $r = 3$; die Pivots sind $j_1 = 1$, $j_2 = 3$, $j_3 = 5$; die freien Variablen sind x_2, x_4.

$$L = \left\{ \begin{bmatrix} 1 - 2x_2 - 2x_4 \\ x_2 \\ 1 - x_4 \\ x_4 \\ 1 \end{bmatrix} \,\middle|\, x_2, x_4 \in \mathbb{Q} \text{ beliebig} \right\} \quad \text{mit} \quad x_0 := \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \in L,$$

siehe Kapitel 1, Beispiel 2.9. (Wir erhalten immer eine spezielle Lösung x_0, indem man die freien Variablen, also hier x_2, x_4, gleich 0 setzt.)

$$N(A) = N(A') = \{v(x_2, x_4) \mid x_2, x_4 \in \mathbb{Q}\} \quad \text{mit} \quad v(x_2, x_4) := \begin{bmatrix} -2x_2 - 2x_4 \\ x_2 \\ -x_4 \\ x_4 \\ 0 \end{bmatrix}.$$

Wie in Beispiel 1.2 ist eine Basis von $N(A)$ gegeben durch $v(1,0), v(0,1)$. Damit
$$L = \{x = x_0 + sv(1,0) + tv(0,1) \mid s, t \in \mathbb{Q}\}.$$

§2 Darstellende Matrix einer linearen Abbildung

Seien V und W Vektorräume über einem Körper K. Wir setzen
$$\mathrm{Hom}(V, W) := \{\varphi\colon V \to W \mid \varphi \text{ linear}\} \subseteq \mathrm{Abb}(V, W);$$
hier steht “Hom” für ***Homomorphismus***.

Ist $V = W$, so schreiben wir auch $\mathrm{End}(V) := \mathrm{Hom}(V, V)$;
hier steht “End” für ***Endomorphismus***.

Seien $\varphi, \psi \in \mathrm{Hom}(V, W)$ und $s \in K$. Dann definieren wir Abbildungen
$$\varphi + \psi\colon V \to W \quad \text{und} \quad s\varphi\colon V \to W$$
durch $(\varphi + \psi)(v) := \varphi(v) + \psi(v)$ und $(s\varphi)(v) := s\varphi(v)$ für alle $v \in V$.

Bemerkung 2.1. Mit obigen Bezeichnungen gilt.

(a) Die Abbildungen $\varphi + \psi\colon V \to W$ und $s\varphi\colon V \to W$ sind wieder linear.

(b) Mit diesen Verknüpfungen ist $\mathrm{Hom}(V, W)$ ein K-Vektorraum.

Beweis. Unmittelbares Nachrechnen (selbst!). □

Wollen den Zusammenhang von $\mathrm{Hom}(V, W)$ und Matrizen studieren.
Zuerst betrachten wir den Fall $V = K^n$ und $W = K^m$.

Bemerkung 2.2

Seien $n, m \geqslant 1$ und $\varphi \in \mathrm{Hom}(K^n, K^m)$. Dann gibt es genau eine Matrix $A \in K^{m \times n}$, so dass $\varphi(v) = Av$ für alle $v \in K^n$ gilt.

Beweis. Sei dazu $B = \{e_1, \ldots, e_n\}$ die Basis der Einheitsvektoren von K^n (siehe Kapitel 2, §2). Für $1 \leqslant j \leqslant n$ sei $w_j := \varphi(e_j) \in K^m$. Sei $A \in K^{m \times n}$ die Matrix mit Spalten $w_1, \ldots, w_n$, d.h., es gilt $\varphi(e_j) = w_j = Ae_j$ für $1 \leqslant j \leqslant n$. Sei nun $v \in K^n$ beliebig, mit Komponenten $v_1, \ldots, v_n$. Dann folgt:
$$\varphi(v) = \varphi\Big(\sum_{j=1}^n v_j e_j\Big) = \sum_{j=1}^n v_j \varphi(e_j) = \sum_{j=1}^n v_j A e_j = \sum_{j=1}^n A v_j e_j = A\Big(\sum_{j=1}^n v_j e_j\Big) = Av.$$
Ist auch $A' \in K^{m \times n}$ so, dass $\varphi(v) = A'v$ gilt für alle $v \in K^n$, so folgt $Ae_j = \varphi(e_j) = A'e_j$ für $1 \leqslant j \leqslant n$, also sind alle Spalten von A und A' gleich. □

Lemma 2.3. Sei V und W Vektorräume über K.

Sei V e.e., $n = \dim V \geqslant 1$ und $B = \{v_1, \ldots, v_n\}$ eine Basis von V.

(a) Sind $\varphi, \psi \in \mathrm{Hom}(V, W)$ gegeben mit $\varphi(v_i) = \psi(v_i)$ für $1 \leqslant i \leqslant n$, so gilt $\varphi = \psi$.

(b) Seien $w_1, \ldots, w_n \in W$ beliebig. Dann gibt es genau ein $\varphi \in \mathrm{Hom}(V, W)$ mit $\varphi(v_i) = w_i$ für $1 \leqslant i \leqslant n$.

Beweis. (a) Sei $v \in V$ beliebig. Dann ist $v = s_1 v_1 + \ldots + s_n v_n$ mit $s_i \in K$, also
$$\varphi(v) = \varphi\Big(\sum_{i=1}^n s_i v_i\Big) = \sum_{i=1}^n s_i \varphi(v_i) = \sum_{i=1}^n s_i \psi(v_i) = \psi(v).$$
(b) Wir definieren $\varphi\colon V \to W$ wie folgt. Sei $v \in V$ beliebig. Dann gibt es eine eindeutige Darstellung $v = s_1 v_1 + \ldots + s_n v_n$ mit $s_i \in K$. Wir setzen
$$\varphi(v) := \sum_{i=1}^n s_i w_i \in W.$$
Seien $v, w \in V$, mit eindeutigen Darstellungen $v = s_1 v_1 + \ldots + s_n v_n$ und $w = t_1 v_1 + \ldots + t_n v_n$, wobei $s_i, t_i \in K$. Dann gilt also $\varphi(v) = \sum_{i=1}^n s_i w_i$ und $\varphi(w) = \sum_{i=1}^n t_i w_i$. Hieraus folgt sofort $\varphi(v + w) = \sum_{i=1}^n (s_i + t_i) w_i = \varphi(v) + \varphi(w)$

und $\varphi(sv) = \sum_{i=1}^n s s_i w_i = s\varphi(v)$ für $s \in K$. Also gibt es φ mit der gewünschten Eigenschaft. Die Eindeutigkeit ist dann klar nach (a). □

Beispiel 2.4

Sei K Körper und $U \leqslant K^n$ ein beliebiger Teilraum. Dann gibt es eine Matrix $A \in M_n(K)$, so dass $U = \{x \in K^n \mid Ax = 0_n\}$ gilt.

Denn: Ist $U = \{0_V\}$, so setze $A := I_n$; ist $U = K^n$, so setze $A = 0_{n\times n}$. Sei nun $1 \leqslant d := \dim U < n$; sei $\{v_1, \ldots, v_d\}$ eine Basis von U; nach Kapitel 2, Satz 4.12, können wir diese zu einer Basis $B := \{v_1, \ldots, v_d, v_{d+1}, \ldots, v_n\}$ von K^n ergänzen. Nach Lemma 2.3 erhalten wir eine lineare Abbildung $\varphi\colon K^n \to K^n$ mit $\varphi(v_i) = 0_n$ für $1 \leqslant i \leqslant d$ und $\varphi(v_i) = v_i$ für $d+1 \leqslant i \leqslant n$. Dann hat Bild$(\varphi) = \langle v_{d+1}, \ldots v_n\rangle_K$ Dimension $n-d$, also ist $\dim$ Kern$(\varphi) = n-(n-d) = d$. Wegen $U \subseteq$ Kern(φ) folgt Kern$(\varphi) = U$. Nach Bemerkung 2.2 gibt es eine Matrix $A \in M_n(K)$ mit $\varphi(v) = Av$ für alle $v \in K^n$. Dann ist $U =$ Kern$(\varphi) = \{x \in K^n \mid Ax = \varphi(x) = 0_n\}$. □

Beispiele dazu siehe Übung.

Von nun an seien V und W e.e., mit $n = \dim V \geqslant 1$ und $m = \dim W \geqslant 1$.
Sei $B = \{v_1, \ldots, v_n\}$ eine Basis von V und $C = \{w_1, \ldots, w_m\}$ eine Basis von W.
Sind $v \in V$ und $w \in W$, so lassen sich v, w eindeutig schreiben als

$$v = x_1v_1 + \ldots + x_nv_n \qquad \text{und} \qquad w = y_1w_1 + \ldots + y_mw_m \qquad (\text{mit } x_i, y_i \in K).$$

Wie in Kapitel 2, Satz 4.6, haben wir die zugehörigen Koordinatenvektoren

$$M_B(v) := \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in K^n \qquad \text{und} \qquad M_C(w) := \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} \in K^m.$$

Definition 2.5. Mit obigen Bezeichnungen sei $\varphi \in \mathrm{Hom}(V,W)$.

Für $1 \leqslant j \leqslant n$ ist $\varphi(v_j) \in W$, also lässt sich $\varphi(v_j)$ eindeutig schreiben als

$$\varphi(v_j) = \sum_{i=1}^m a_{ij}w_i \qquad \text{mit } a_{ij} \in K.$$

Dann heißt $M_C^B(\varphi) := [a_{ij}] \in K^{m\times n}$ die (darstellende) Matrix von φ bezüglich der Basen B und C. Die j-te Spalte von A ist also der Koordinatenvektor $M_C(\varphi(v_j))$.

Beachte: Die Reihenfolge der Vektoren in den Basen B und C ist wichtig.
Ändert man die Reihenfolge, so ändert sich auch $M_C^B(\varphi)$ entsprechend.
(Siehe Übung: Permutationsmatrizen.)

Beispiel 2.6

(a) Sei $V = W$ und $\varphi = \mathrm{id}_V$ die identische Abbildung. Dann ist $\mathrm{id}_V(v_j) = v_j$ für $1 \leqslant j \leqslant n$, also $M_B^B(\mathrm{id}_V) = I_n$ (Einheitsmatrix). Beachte: Benutzen wir zwei Basen B, C von $V = W$ wie oben, so ist $T := M_C^B(\mathrm{id}_V)$ nicht unbedingt gleich I_n. Die Einträge von T beschreiben, wie sich die $v_j \in B$ als Linearkomination der $w_i \in C$ ausdrücken lassen. Daher heißt T auch ***Basiswechselmatrix***.

(b) Sei $A \in K^{m\times n}$; betrachte die lineare Abbildung $\varphi_A\colon K^n \to K^m$, $v \mapsto Av$.
Sei B die Basis der Einheitsvektoren in $V = K^n$, also $M_B(v) = v$ für $v \in K^n$.
Sei C die Basis der Einheitsvektoren in $W = K^m$, also $M_C(w) = w$ für $w \in K^m$.
Für $1 \leqslant j \leqslant n$ ist $M_C(\varphi(v_j)) = M_C(Av_j) = Av_j$ die j-te Spalte von A, also folgt $A = M_C^B(\varphi_A)$. (Jede Matrix ist die darstellende Matrix einer linearen Abbildung.)

Lemma 2.7. Sei $\varphi \in \mathrm{Hom}(V,W)$.

Für alle $v \in V$ gilt $M_C(\varphi(v)) = M_C^B(\varphi) \cdot M_B(v)$ (Produkt Matrix mal Spaltenvektor).

Beweis. Wie oben sei $M_C^B(\varphi) = [a_{ij}]$ und $v = x_1v_1 + \ldots + x_nv_n$ mit $x_i \in K$. Dann

$$\varphi(v) = \sum_{j=1}^n x_j\varphi(v_j) = \sum_{j=1}^n x_j\Big(\sum_{i=1}^m a_{ij}w_i\Big) = \sum_{i=1}^m\Big(\sum_{j=1}^n a_{ij}x_j\Big)w_i.$$

Nun ist $y_i := \sum_{j=1}^n a_{ij}x_j$ genau die i-te Komponente des Produkts $[a_{ij}] \cdot M_B(v)$.
Aber nach obiger Formel ist y_i auch die i-te Komponente von $M_C(\varphi(v))$. □

Beispiel 2.8

Sei $V =$ Abb(K,K) der Vektorraum aller Funktionen $f\colon K \to K$. Sei $c \in K$ fest.
Für $f \in V$ definiere eine neue Funktion $f_c \in V$ durch $f_c(x) = f(x+c)$ für alle $x \in K$. Dann ist die Abbildung $\varphi\colon V \to V$, $f \mapsto f_c$, linear.
Denn für $f, g \in V$ und $s \in K$ gilt

$$(f+g)_c(x) = (f+g)(x+c) = f(x+c) + g(x+c) = f_c(x) + g_c(x) = (f_c + g_c)(x)$$

und $(sf)_c(x) = (sf)(x+c) = sf(x+c) = sf_c(x) = (sf_c)(x)$ für alle $x \in K$.

Sei nun K ein unendlicher Körper, $n \geqslant 1$ und $P_n(K) \subseteq V$ der Teilraum der Polynomfunktionen vom Grad $\leqslant n$, mit Basis $B = \{p_0, p_1, p_2, \dots, p_n\}$ wobei $p_i(x) = x^i$ für alle $x \in K$ (siehe Kapitel 2, §3). Ist $f \in P_n(K)$, so gibt es eindeutige $a_i \in K$ mit $f(x) = \sum_{i=0}^n a_i x^i$ für alle $x \in K$. Dann ist

$$M_B(f) = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} \in K^n.$$

Mit dem Binomischen Lehrsatz folgt für alle $x \in K$:

$$f(x+c) = \sum_{i=0}^n a_n (x+c)^i = \sum_{i=0}^n a_i \sum_{j=0}^i \binom{i}{j} x^j c^{i-j} = \sum_{j=0}^n \Big(\underbrace{\sum_{i=j}^n \binom{i}{j} a_i c^{i-j}}_{=:b_j \in K} \Big) x^j$$

Also ist auch $f_c = \sum_{j=0}^n b_j p_j \in P_n(K)$. Durch Einschränkung erhalten wir damit auch eine lineare Abbildung $\varphi\colon P_n(K) \to P_n(K),\ f \mapsto f_c$.

Sei nun $n = 3$. Dann ist

$$\varphi(p_0)(x) = p_0(x+c) = 1, \qquad \varphi(p_1)(x) = p_1(x+c) = x + c,$$
$$\varphi(p_2)(x) = p_2(x+c) = (x+c)^2 = x^2 + 2cx + c^2,$$
$$\varphi(p_3)(x) = p_3(x+c) = (x+c)^3 = x^3 + 3cx^2 + 3c^2x + c^3$$

für alle $x \in K$. Also ist

$$M_B^B(\varphi) = \begin{bmatrix} 1 & c & c^2 & c^3 \\ 0 & 1 & 2c & 3c^2 \\ 0 & 0 & 1 & 3c \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Um f_c für ein beliebiges $f \in P_3(K)$ zu bestimmen, brauchen wir nur das Produkt $M_B^B(\varphi) \cdot M_B(f)$ zu bilden. Ist also z.B. $f(x) = x^3 - x$ für alle $x \in K$, so ist $f = -p_1 + p_3$. Wir berechnen

$$M_B(f_c) = M_B^B(\varphi) \cdot M_B(f) = \begin{bmatrix} 1 & c & c^2 & c^3 \\ 0 & 1 & 2c & 3c^2 \\ 0 & 0 & 1 & 3c \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -c + c^3 \\ -1 + 3c^2 \\ 3c \\ 1 \end{bmatrix}.$$

Also ist $f_c(x) = x^3 + 3cx^2 + (3c^2 - 1)x + (c^3 - c)$ für alle $x \in K$.

Satz 2.9. Seien weiterhin wie oben B eine Basis von V und C eine Basis von W.

Die Abbildung $\Phi_C^B\colon \mathrm{Hom}(V, W) \to K^{m \times n}$, $\varphi \mapsto M_C^B(\varphi)$, ist linear und bijektiv. Insbesondere folgt $\dim \mathrm{Hom}(V, W) = (\dim V) \cdot (\dim W) < \infty$.

Beweis. Man rechnet sofort nach (selbst!), dass Φ_C^B linear ist. Ist $\Phi_C^B(\varphi) = 0_{m \times n}$, so folgt mit Lemma 2.7, dass $M_C(\varphi(v)) = 0_m$ gilt für alle $v \in V$. Nach Kapitel 2, Satz 4.6, ist $M_C\colon W \to K^m$, $w \mapsto M_C(w)$, linear und bijektiv, also folgt $\varphi(v) = 0_W$ für alle $v \in V$, d.h., $\varphi = \underline{0}$. Also ist Kern$(M_C^B) = \{\underline{0}\}$ und damit M_C^B injektiv.

Umgekehrt sei $A \in K^{m \times n}$ beliebig. Für $1 \leqslant j \leqslant n$ sei $z_j \in W$ der eindeutige Vektor, so dass $M_C(z_j) \in K^m$ die j-te Spalte von A ist. Nach Lemma 2.3 gibt es genau ein $\varphi \in \mathrm{Hom}(V, W)$ mit $\varphi(v_j) := z_j$ für $1 \leqslant j \leqslant n$. Nach Konstruktion ist $M_C^B(\varphi) = A$. Also ist M_C^B auch surjektiv, und damit bijektiv.

Da $K^{m \times n}$ e.e. und $\dim(K^{m \times n}) = nm$ (Kapitel 2, Beispiel 4.4), ist auch $\mathrm{Hom}(V, W)$ e.e. und $\dim \mathrm{Hom}(V, W) = \dim(K^{m \times n}) = mn$, siehe Kapitel 2, Beispiel 4.14. □

Sei nun U weiterer e.e. K-Vektorraum mit $p = \dim U \geqslant 1$. Dann können wir die Hintereinanderausführung von linearen Abbildungen $U \to V \to W$ betrachten. Sei $D = \{u_1, \dots, u_p\}$ eine Basis von U.

Satz 2.10. Seien $\varphi \in \mathrm{Hom}(V, W)$ und $\psi \in \mathrm{Hom}(U, V)$.

Dann ist auch $\varphi \circ \psi\colon U \to W$ linear, d.h., $\varphi \circ \psi \in \mathrm{Hom}(U, W)$, und es gilt

$$M_C^D(\varphi \circ \psi) = M_C^B(\varphi) \cdot M_B^D(\psi) \qquad \text{(Matrixprodukt).}$$

Beweis. Nach Aufgabe 13.1 ist $\varphi \circ \psi\colon U \to W$ linear. Sei $M_C^B(\varphi) = [a_{ij}] \in K^{m \times n}$ und $M_B^D(\psi) = [a'_{jk}] \in K^{n \times p}$. Außerdem sei $M_C^D(\varphi \circ \psi) = [g_{ik}] \in K^{m \times n}$, also $(\varphi \circ \psi)(u_k) = \sum_{i=1}^m g_{ik} w_i$ für $1 \leqslant k \leqslant p$. Andererseits gilt aber auch

$$(\varphi \circ \psi)(u_k) = \varphi(\psi(u_k))$$
$$= \varphi\Big(\sum_{j=1}^n a'_{jk} v_j\Big) = \sum_{j=1}^n a'_{jk} \varphi(v_j) = \sum_{j=1}^n a'_{jk} \Big(\sum_{i=1}^m a_{ij} w_i\Big) = \sum_{i=1}^m \Big(\sum_{j=1}^n a_{ij} a'_{jk}\Big) w_i$$

und damit $g_{ik} = \sum_{j=1}^n a_{ij} a'_{jk} = (i, k)$-Eintrag von $A \cdot A'$. □

Der obige Satz ist letztlich die Begründung für die Definition des Matrixprodukts.

Beispiel 2.11

Sei $V = W$ und $\varphi = \mathrm{id}_V$. Sind B, C zwei Basen von V, so haben wir die Basiswechselmatrizen $T = M_C^B(\mathrm{id}_V) \in M_n(K)$ und $T' = M_B^C(\mathrm{id}_V) \in M_n(K)$ (siehe Beispiel 2.6). Nach Satz 2.10 gilt

$$T \cdot T' = M_C^B(\mathrm{id}_V) \cdot M_B^C(\mathrm{id}_V) = M_C^C(\mathrm{id}_V \circ \mathrm{id}_V) = M_C^C(\mathrm{id}_V) = I_n,$$

also sind T, T' invertierbar und invers zueinander, $T' = T^{-1}$.

Folgerung 2.12. Sei $\varphi \in \mathrm{Hom}(V, W)$, sowie B Basis von V und C Basis von W.

Sei $B' = \{v_1', \ldots, v_n'\}$ eine weitere Basis von V und $C' = \{w_1', \ldots, w_m'\}$ eine weitere Basis von W. Seien $T = M_B^{B'}(\mathrm{id}_V) \in M_n(K)$ und $P = M_C^{C'}(\mathrm{id}_W)$ die zugehörigen Basiswechselmatrizen. Nach Beispiel 2.11 sind T und P invertierbar. Dann gilt $\quad M_{C'}^{B'}(\varphi) = P^{-1} \cdot M_C^B(\varphi) \cdot T.$

Beweis. Es gilt offenbar $\varphi = \varphi \circ \mathrm{id}_V$, also folgt mit Satz 2.10:

$$M_C^{B'}(\varphi) = M_C^{B'}(\varphi \circ \mathrm{id}_V) = M_C^B(\varphi) \cdot M_B^{B'}(\mathrm{id}_V) = M_C^B(\varphi) \cdot T.$$

Andererseits ist auch $\varphi = \mathrm{id}_W \circ \varphi$, also

$$M_C^{B'}(\varphi) = M_C^{B'}(\mathrm{id}_W \circ \varphi) = M_C^{C'}(\mathrm{id}_W) \cdot M_{C'}^{B'}(\varphi) = P \cdot M_{C'}^{B'}(\varphi). \qquad \square$$

Beispiel 2.13. Wichtiger Spezialfall.

Sei $\varphi \in \mathrm{End}(V)$. Wir schreiben dann einfach $M_B(\varphi) := M_B^B(\varphi)$. Ist B' weitere Basis von V wie oben, so sei $T = M_B^{B'}(\mathrm{id}_V)$ die zugehörige Basiswechselmatrix. Nach Beispiel 2.11 ist $T^{-1} = M_{B'}^B(\mathrm{id}_V)$. Also besagt Folgerung 2.12 (mit $B = C$, $B' = C'$) nun einfach: $\quad M_{B'}(\varphi) = T^{-1} \cdot M_B(\varphi) \cdot T.$

Allgemein heißen zwei Matrizen $A, A' \in M_n(K)$ ***ähnlich***, wenn es eine invertierbare Matrix $T \in M_n(K)$ gibt mit $A' = T^{-1} \cdot A \cdot T$.

Grundlegende Fragestellungen:

Für $\varphi \in \mathrm{End}(V)$ finde Basis B von V, so dass $A = M_B(\varphi)$ eine "möglichst einfache" Gestalt hat. Oder Matrixversion: Für $A \in M_n(K)$ finde invertierbare Matrix $T \in M_n(K)$ so dass $A' = T^{-1} \cdot A \cdot T$ eine "möglichst einfache" Gestalt hat.

Erster Spezialfall im folgenden Abschnitt; allgemeine Lösung in LAAG2.

§3 Eigenräume und Diagonalisierbarkeit

In Kapitel 1, §4 haben wir Eigenwerte von Matrizen eingeführt. Erinnerung: $\lambda \in K$ heißt Eigenwert von $A \in M_n(K)$, wenn es ein $0_n \neq v \in K^n$ gibt mit $Av = \lambda v$.

Nun sei $V \neq \{0_V\}$ ein Vektorraum über K und $\varphi\colon V \to V$ eine lineare Abbildung.

Definition 3.1

(a) Sei $\lambda \in K$. Dann heißt λ ein ***Eigenwert*** von φ, wenn es ein $v \in V$ gibt mit

$$v \neq 0_V \qquad \text{und} \qquad \varphi(v) = \lambda v.$$

(b) Sei $v \in V$, $v \neq 0_V$. Dann heißt v ein ***Eigenvektor*** von φ, wenn es ein $\lambda \in K$ gibt mit $\varphi(v) = \lambda v$. (Dann ist also λ der zu v gehörige Eigenwert.)

Bemerkung 3.2. Sei $\lambda \in K$ und $E_\varphi(\lambda) := \{v \in V \mid \varphi(v) = \lambda v\}$.

Es gilt $\quad v \in E_\varphi(\lambda) \quad \Leftrightarrow \quad \varphi(v) - \lambda v = 0_V \quad \Leftrightarrow \quad (\varphi - \lambda \mathrm{id}_V)v = 0_V.$

Also: $\quad E_\varphi(\lambda) = \mathrm{Kern}(\varphi - \lambda \mathrm{id}_V)$, und dies ist ein Teilraum von V.

Insbesondere: $\quad \lambda$ Eigenwert $\quad \Leftrightarrow \quad E_\varphi(\lambda) \neq \{0_V\}.$

Bemerkung 3.3. Sei nun V e.e. und $n = \dim V \geqslant 1$.

Sei $B = \{v_1, \ldots, v_n\}$ eine Basis von V und $A = M_B(\varphi) \in M_n(K)$ die Matrix von φ bezüglich B. Sei $\lambda \in K$. Dann gilt:

λ Eigenwert von $\varphi \quad \Leftrightarrow \quad \lambda$ Eigenwert von A (im Sinne von Kapitel 1, §4).

Beweis. Sei $0_V \neq v \in V$ und $0_n \neq x = M_B(v) \in K^n$. Ist λ ein Eigenwert von φ mit Eigenvektor v, so gilt $\varphi(v) = \lambda v$ und $\lambda M_B(v) = M_B(\lambda v) = M_B(\varphi(v)) = AM_B(v)$. D.h. λ ist ein Eigenwert von A mit Eigenvektor $x = M_B(v)$. Umgekehrt: Ist λ ein Eigenwert zu A mit Eigenvektor v, so gilt $Ax = \lambda x$ und wir erhalten

$$M_B(\varphi(v)) = AM_B(v) = Ax = \lambda x = \lambda M_B(v) = M_B(\lambda v).$$

Weil die Abbildung $V \to K^n$, $w \mapsto M_B(w)$, bijektiv ist, folgt $\varphi(v) = \lambda v$. $\square$

Also können wir die Eigenwerte von φ mit Hilfe der Minimalgleichung $\mu_A\colon K \to K$ wie in Kapitel 1, §4, bestimmen. Erinnerung: Sei $d = d(A) \geqslant 1$ minimal, so dass es $a_0, a_1, \ldots, a_{d-1} \in K$ gibt mit $A^d + a_{d-1}A^{d-1} + \ldots + a_1 A + a_0 I_n = 0_{n\times n}$. Dann sind die a_i eindeutig und $\mu_A(x) = a_0 + a_1 x + \ldots + a_{d-1}x^{d-1} + x^d$ für alle $x \in K$.

Bemerkung 3.4. Sei V e.e. und $n = \dim V \geqslant 1$.

Sei $\lambda \in K$ Eigenwert von φ. Dann heißt $E_\varphi(\lambda)$ der zugehörige ***Eigenraum***.
Berechnung von $E_\varphi(\lambda)$? Dazu sei wieder $B = \{v_1, \ldots, v_n\}$ eine Basis von V und $A = M_B(\varphi) \in M_n(K)$ die Matrix von φ bezüglich B. Sei

$$N(A - \lambda I_n) := \{x \in K^n \mid (A - \lambda I_n)x = 0_n\}$$

Lösungsmenge des homogenen linearen Gleichungssystems mit Matrix $A - \lambda I_n$.
Behauptung: $E_\varphi(\lambda) = \{v \in V \mid M_B(v) \in N(A - \lambda I_n)\}$.
(Denn sei $v \in V$ und $x = M_B(v) \in K^n$ zugehöriger Koordinatenvektor. Im Beweis von Bemerkung 3.3 haben wir gesehen: $v \in E_\varphi(\lambda) \Leftrightarrow Ax = \lambda x$. Nun gilt $Ax = \lambda x \Leftrightarrow (A - \lambda I_n)x = 0_n \Leftrightarrow x \in N(A - \lambda I_n)$, also die Behauptung.)

Beispiel 3.5.

(a) Ist Matrix $A \in M_n(K)$ gegeben, so haben wir die zugehörige lineare Abbildung $\varphi_A \colon K^n \to K^n$, $v \mapsto Av$. Ist $B = \{e_1, \ldots, e_n\}$ die Basis der Einheitsvektoren in K^n, so gilt $A = M_B(\varphi_A)$. Also folgt $E_A(\lambda) := E_{\varphi_A}(\lambda) = N(A - \lambda I_n)$ für $\lambda \in K$.

(b) Set $K = \mathbb{Q}$ und $A = \begin{bmatrix} 5 & 1 & 0 \\ 1 & 6 & 1 \\ 4 & 3 & 9 \end{bmatrix} \in M_3(\mathbb{Q})$ die Waschmittelmatrix in Kapitel 1, §4.

Dann ist $\mu_A(x) = (x-10)(x-5)^2$ für alle $x \in \mathbb{Q}$, also hat φ_A die 2 Eigenwerte $\lambda_1 = 10$ und $\lambda_2 = 5$. Nach Aufgabe 13.4 sind die Eigenräume:

$$E_A(10) = N(A - 10I_3) = \left\langle \begin{bmatrix} 1/19 \\ 5/19 \\ 1 \end{bmatrix} \right\rangle_{\mathbb{Q}} \quad \text{und} \quad E_A(5) = N(A - 5I_3) = \left\langle \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\rangle_{\mathbb{Q}}.$$

Definition 3.6. Sei V e.e. und $\varphi \colon V \to V$ linear; sei $n := \dim V \geqslant 1$.

Dann heißt φ ***diagonalisierbar***, wenn es eine Basis B von V gibt, so dass $A := M_B(\varphi)$ eine Diagonalmatrix ist, also eine Gestalt wie folgt hat:

$$A = [a_{ij}] = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \in M_n(K) \qquad \text{(d.h., } a_{ij} = 0 \text{ für alle } i \neq j\text{).}$$

Dies ist sicherlich die optimal einfachste Form einer darstellenden Matrix!
Beachte: Die $\lambda_i \in K$ müssen hier nicht paarweise verschieden sein.

Bemerkung. Sei $n = \dim V \geqslant 1$ und $B = \{v_1, \ldots, v_n\}$ eine Basis von V. Genau dann ist $M_B(\varphi)$ eine Diagonalmatrix wie oben, wenn $\varphi(v_i) = \lambda_i v_i$ für $1 \leqslant i \leqslant n$ gilt, d.h., wenn jedes v_i ein Eigenvektor von V ist (mit Eigenwert λ_i). Also:

φ diagonalisierbar $\Leftrightarrow$ es gibt eine Basis B aus Eigenvektoren von V.

Bemerkung 3.7. Sei $A \in M_n(K)$.

Wir sagen, dass A diagonalisierbar ist, wenn die zugehörige lineare Abbildung $\varphi_A \colon K^n \to K^n$, $v \mapsto Av$, diagonalisierbar ist. Dies ist äquivalent dazu, dass es eine invertierbare Matrix $T \in M_n(K)$ gibt, so dass $T^{-1}AT$ eine Diagonalmatrix ist.
(Denn sei $B = \{e_1, \ldots, e_n\}$ die Basis der Einheitsvektoren von $V = K^n$. Dann ist $A = M_B(\varphi_A)$. Sei B' weitere Basis von K^n. Dann ist $M_{B'}(\mathrm{id}_V) = T^{-1}AT$, wobei $T := M_B^{B'}(\varphi_A)$ Basiswechselmatrix wie in Beispiel 2.13.)

Einfachstes Beispiel einer nicht-diagonalisierbaren Matrix ist $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$.
(Denn 0 ist einziger Eigenwert, aber Eigenraum hat nur Dimension 1.)

Beispiel 3.8. Sei $K = \mathbb{Q}$.

Sei $A = \begin{bmatrix} -1 & 0 & 1 \\ 3 & 0 & -3 \\ 1 & 0 & -1 \end{bmatrix} \in M_3(\mathbb{Q})$.

Man rechnet nach, dass $A^2 = -2A$ gilt, also ist $\mu_A(x) = x(x+2)$ für alle $x \in \mathbb{Q}$.
Damit haben wir 2 Eigenwerte $\lambda_1 = 0$ und $\lambda_2 = -2$. Die Eigenräume sind:

$$E_A(0) = \left\langle \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\rangle_{\mathbb{Q}} \quad \text{und} \quad E_A(-2) = \left\langle \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix} \right\rangle_{\mathbb{Q}}.$$

Prüfe nach, dass obige 3 Eigenvektoren eine Basis von $\mathbb{Q}^3$ bilden. Also A diagonalisierbar. Sei $B = \{e_1, e_2, e_3\}$ Basis der Einheitsvektoren in $\mathbb{Q}^3$ und B' neue Basis, die aus den obigen 3 Eigenvektoren besteht. Dann ist $T^{-1}AT =$ Diagonalmatrix mit Einträgen $0, 0, -2$ auf der Diagonalen, wobei

$$T := M_B^{B'}(\mathrm{id}_{\mathbb{Q}^3}) = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 3 \\ 0 & 1 & 1 \end{bmatrix} \in M_3(\mathbb{Q}) \qquad \text{(Spalten von } T \text{ sind obige } 3 \text{ Eigenvektoren)}$$

Satz 3.9 (Charakterisierung diagonalisierbarer Matrizen). Sei $A \in M_n(K)$.

Sei $d = d(A) \geqslant 1$ der Minimalgrad und $\mu_A \colon K \to K$ die Minimalgleichung von A (siehe Kapitel 1, §4). Genau dann ist A diagonalisierbar, wenn es d paarweise verschiedene Nullstellen für μ_A in K gibt.

Für den Beweis brauchen wir ein neues Hilfsmittel: “abstrakte” Polynome $\rightsquigarrow$ §4.

- Für die Waschmittelmatrix $A = \begin{bmatrix} 5 & 1 & 0 \\ 1 & 6 & 1 \\ 4 & 3 & 9 \end{bmatrix} \in M_3(\mathbb{Q})$ gilt $d = d(A) = 3$ und $\mu_A(x) = (x-10)(x-5)^2$ für alle $x \in \mathbb{Q}$. Also ist die Bedingung in Satz 3.9 nicht erfüllt, und damit folglich A nicht diagonalisierbar.
- Für $A = \begin{bmatrix} -1 & 0 & 1 \\ 3 & 0 & -3 \\ 1 & 0 & -1 \end{bmatrix} \in M_3(\mathbb{Q})$ gilt $d = d(A) = 2$ und $\mu_A(x) = x(x+2)$ für alle $x \in \mathbb{Q}$. Also ist hier die Bedingung erfüllt, und damit A diagonalisierbar.

§4 Polynome

Was versteht man unter einem (“abstrakten”) Polynom ? — Informelle Antwort: Ein “formaler” Ausdruck der Form $x^3 + x$, wobei man für x irgendwelche Werte einsetzen kann. Man kann solche Ausdrücke addieren und multiplizieren (wobei man einfach die üblichen Rechenregeln benutzt), zum Beispiel:

$(2x^2 - 1) + (x^3 + x) = x^3 + 2x^2 + x - 1$ und

$(2x^2 - 1) \cdot (x^3 + x) = 2x^5 + 2x^3 - x^3 - x = 2x^5 + x^3 - x$.

Aber was genau ist ein “Ausdruck” (eine Funktion?), und woraus soll man Werte einsetzen können ? Beachte: Im Beispiel $K = \mathbb{F}_2$ gilt $x^3 + x = 0$ für alle $x \in \mathbb{F}_2$.

Vielleicht hilft es, sich die allgemeine Form eines solchen Ausdrucks vorzustellen; diese sollte so aussehen: $a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$, wobei $a_0, a_1, \ldots, a_n \in K$.

Letztlich relevant scheinen also eigentlich nur die Koeffizienten $a_0, a_1, a_2, \ldots, a_n$ zu sein. Werden nun sehen, dass diese Beobachtung tatsächlich als Grundlage einer ordentlichen Definition dienen kann.

Sei K beliebiger Körper; betrachte den K-Vektorraum $V = \mathrm{Abb}(\mathbb{N}_0, K)$. Jedes $f \in V$ schreiben wir als Folge $f = (a_n)_{n \geqslant 0}$ (oder einfach (a_n)), wobei $a_n = f(n)$ für alle $n \geqslant 0$. Addition und Skalarmultiplikation sind gegeben durch:

$(a_n) + (b_n) := (a_n + b_n)$ und $s(a_n) := (sa_n)$ (für $s \in K$).

Das neutrale Element bezüglich der Addition ist $\underline{0} = (0, 0, 0, \ldots)$.

Für $f, g \in V$ definiere $f * g \in V$ wie folgt: Sei $f = (a_n)$ und $g = (b_n)$. Für $n \geqslant 0$ sei

$$c_n := \sum_{i=0}^{n} a_i b_{n-i} = a_0 b_n + a_1 b_{n-1} + a_2 b_{n-2} + \ldots + a_{n-1} b_1 + a_n b_0.$$

Dann setze $f * g := (c_n)$. Diese Verknüpfung heißt ***Konvolution***.

Beispiel: $(-1, 0, 2, 0, 0, \ldots) * (0, 1, 0, 1, 0, 0, \ldots) = (0, -1, 0, 1, 0, 2, 0, 0, \ldots)$, denn $c_0 = a_0 b_0 = 0,\ c_1 = a_0 b_1 + a_1 b_0 = -1,\ c_2 = a_0 b_2 + a_1 b_1 + a_2 b_0 = 0$, usw.

Definition 4.1

Wir sagen, dass $f = (a_n) \in V$ endlich ist, wenn es ein $n_0 \geqslant 0$ gibt mit $a_n = 0$ für alle $n > n_0$. In diesem Fall schreiben wir f einfach als $f = (a_0, a_1, \ldots, a_{n_0}, 0, \ldots)$. Ist $f \neq \underline{0}$, so heißt $\mathrm{Grad}(f) := \max\{n \in \mathbb{N}_0 \mid a_n \neq 0\}$ der ***Grad*** von f.

Lemma 4.2. Sei $\mathcal{P}(K) := \{f \in V \mid f \text{ ist endlich}\}$ wie oben.

Dann ist $\mathcal{P}(K)$ ein Teilraum von V. Seien nun $f, g \in \mathcal{P}(K)$. Dann gilt:

(a) Es ist auch $f * g \in \mathcal{P}(K)$.

(b) Sind $f, g \neq \underline{0}$ und $f + g \neq \underline{0}$, so gilt $\mathrm{Grad}(f+g) \leqslant \max\{\mathrm{Grad}(f), \mathrm{Grad}(g)\}$.

(c) Sind $f, g \neq \underline{0}$, so gilt $f * g \neq 0$ und $\mathrm{Grad}(f * g) = \mathrm{Grad}(f) + \mathrm{Grad}(g)$.

Beweis. Seien $f = (a_n)$ und $g = (b_n)$ in $\mathcal{P}(K)$. Ist $f = \underline{0} = (0, 0, \ldots)$, so ist $f + g = g \in \mathcal{P}(K)$ und $sf = \underline{0} \in \mathcal{P}(K)$; ebenso $f * g = \underline{0} \in \mathcal{P}(K)$. Analog: Ist $g = \underline{0}$, so sind $f + g = f \in \mathcal{P}(K)$ und $f * g = \underline{0} \in \mathcal{P}(K)$. Seien nun also $f \neq \underline{0}$ und $g \neq \underline{0}$.

Seien $n_0 = \mathrm{Grad}(f)$ und $m_0 = \mathrm{Grad}(g)$; also $f = (a_0, a_1, \ldots, a_{n_0}, 0, \ldots)$ mit $a_{n_0} \neq 0$, und $g = (b_0, b_1, \ldots, b_{m_0}, 0, \ldots)$ mit $b_{m_0} \neq 0$. Dann ist offensichtlich $sf = (sa_0, sa_1, \ldots, sa_{n_0}, 0, \ldots) \in \mathcal{P}(K)$ für $s \in K$. Ist $d_0 = \max\{n_0, m_0\}$, so gilt auch $f + g = (a_0 + b_0, a_1 + b_1, \ldots, a_{d_0} + b_{d_0}, 0, \ldots) \in \mathcal{P}(K)$.

Ist $f + g \neq \underline{0}$, so ist damit $\mathrm{Grad}(f+g) \leqslant d_0$, also gilt (b).

Außerdem ist damit gezeigt, dass $\mathcal{P}(K)$ ein Teilraum von $V = \mathrm{Abb}(\mathbb{N}_0, K)$ ist.
Nun betrachte $f * g = (c_n)$. Ist $c_n = \sum_{i=0}^{n} a_i b_{n-1} \neq 0$, so gibt es ein i mit $a_i \neq 0$ und $b_{n-i} \neq 0$, d.h., $i \leqslant n_0$ und $n - i \leqslant m_0$, also $n \leqslant m_0 + i \leqslant m_0 + n_0$.
Also folgt $f * g = (c_0, c_1, \ldots, c_{n_0+m_0}, 0, \ldots) \in \mathcal{P}(K)$. Betrachten wir nun $c_{n_0+m_0} = \sum_{i=0}^{n_0+m_0} a_i b_{n_0+m_0-i}$. Ist in dieser Summe $i > n_0$, so folgt $a_i = 0$. Ist $i < n_0$, so ist $n_0 + m_0 - i > m_0$, also $b_{n_0+m_0-i} = 0$. Also bleibt nur $i = n_0$ übrig und es folgt $c_{n_0+m_0} = a_{n_0} b_{m_0} \neq 0$. Also $f * g \neq \underline{0}$ und $\mathrm{Grad}(f * g) = n_0 + m_0$. □

Bemerkung 4.3

(a) Sei $e := (1, 0, 0, \ldots) \in \mathcal{P}(K)$. Dann ist e neutrales Element bezüglich $*$.
(b) Ist $X := (0, 1, 0, 0, \ldots) \in \mathcal{P}(K)$, so gilt $X * (a_n) = (0, a_0, a_1, a_2, \ldots)$.

Beweis. (a) Sei $e = (e_n) = (1, 0, 0, \ldots)$. Dann ist $e * (a_n) = (c_n)$ mit $c_n = \sum_{i=0}^{n} e_i a_{n-i} = \sum_{i=0}^{n} \delta_{i0} a_{n-i} = a_n$.
(b) Seit $X = (X_n) = (0, 1, 0, 0, \ldots)$. Dann ist $X * (a_n) = (c_n)$ mit $c_n = \sum_{i=0}^{n} X_i a_{n-i}$ $= \sum_{i=0}^{n} \delta_{i1} a_{n-i}$. Dies ist gleich 0 falls $n = 0$, und gleich a_{n-1} falls $n \geqslant 1$. □

Definition 4.4 (Formale Definition von "abstrakten" Polynomen über K).

Bezeichnen wir wieder $X := (0, 1, 0, 0, \ldots) \in \mathcal{P}(K)$ wie oben, so schreiben wir auch $K[X] := \mathcal{P}(K)$. Dann definieren wir $X^0 := e = (1, 0, 0, \ldots)$, $X^1 := X$,

$$X^2 := X * X = (0, 0, 1, 0, 0, \ldots), \quad X^3 := X * X^2 = (0, 0, 0, 1, 0, 0, \ldots), \quad \ldots$$

also $X^n := X * X^{n-1}$ für alle $n \geqslant 1$. Ist $f = (a_n) \in \mathcal{P}(K)$ und $n_0 \geqslant 0$ mit $f = (a_0, a_1, \ldots, a_{n_0}, 0, \ldots)$, so erhalten wir eine eindeutige Darstellung

$$(*) \qquad f = \sum_{i=0}^{n_0} a_i X^i = a_0 e + a_1 X + a_2 X^2 + \ldots + a_{n_0} X^{n_0}.$$

Daher heißen die Elemente von $\mathcal{P}(K)$ auch ***Polynome*** in der ***Unbestimmten*** X.
(Wir könnten auch irgendein anderes Symbol anstelle von X nehmen; dies ist immer nur ein anderer Name für die Folge $(0, 1, 0, 0, \ldots) \in \mathcal{P}(K)$.) Da jedes f eine eindeutige Darstellung wie in $(*)$ besitzt, ist $\{X^i \mid i \in \mathbb{N}_0\}$ eine Basis von $\mathcal{P}(K)$. Sei

$$\mathcal{P}_n(K) := \langle X^0, X^1, X^2, \ldots, X^n \rangle_K \subseteq \mathcal{P}(K) \qquad \text{für } n \geqslant 0.$$

Schließlich: Wir fassen K als Teilmenge von $K[X]$ auf, indem wir ein Element $a \in K$ mit $ae = aX^0 = (a, 0, 0, \ldots) \in \mathcal{P}(K)$ identifizieren.

Bemerkung 4.5

Sei $f = (a_n) \in \mathcal{P}(K)$; sei $n_0 \geqslant 0$ mit $f = (a_0, a_1, \ldots, a_{n_0}, 0, \ldots)$. Dann erhalten wir eine Polynomfunktion $\dot{f} \in P(K)$ (siehe Kapitel 2, §3) durch die Definition

$$\dot{f}(x) := a_0 + a_1 x + \ldots + a_{n_0} x^{n_0} \qquad \text{für alle } x \in K.$$

Man sieht sofort, dass die Abbildung $\mathcal{P}(K) \to P(K)$, $f \mapsto \dot{f}$, linear ist. Wir haben in Kapitel 2, §3, gesehen, dass diese Abbildung injektiv ist, wenn $|K| = \infty$.
Aber: Ist $K = \mathbb{F}_2$ und $f = (0, 1, 1, 0, 0, \ldots) \in \mathcal{P}(K)$, so $\dot{f}(x) = x + x^2 = 0$ für $x \in K$.

Also: Es besteht ein Unterschied zwischen Polynomen in $\mathcal{P}(K)$ und Polynomfunktionen in $P(K)$. Ein Polynom in $\mathcal{P}(K)$ ist zunächst einmal nur die Folge $f = (a_0, a_1, \ldots, a_{n_0}, 0, \ldots)$ seiner Koeffizienten (also keine Funktion $K \to K$).

Lemma 4.6

Mit $*\colon \mathcal{P}(K) \times \mathcal{P}(K) \to \mathcal{P}(K)$ ist $\mathcal{P}(K)$ ein kommutativer Ring mit Einselement e.
Sind $f, g \in \mathcal{P}(K)$ und $s \in K$, so gilt $(sf) * g = s(f * g) = f * (sg)$.

Beweis. In den Übungen wird gezeigt, dass $*$ assoziativ und kommutativ ist; die Distributivregeln folgen auch unmittelbar. Seien nun $f = (a_n)$ und $g = (b_n)$ gegeben. Sei $f * g = (c_n)$, mit $c_n = \sum_{i=0}^{n} a_i b_{n-1}$ für $n \geqslant 0$. Sei $(sf) * g = (c_n')$. Nun ist $sf = (sa_n)$, also $c_n' = \sum_{i=0}^{n} (sa_i) b_{n-i} = \sum_{i=0}^{n} s(a_i b_{n-i}) = sc_n$. Also folgt $(sf) * g = s(f * g)$. Völlig analog folgt $f * (sg) = s(f * g)$. □

Bemerkung 4.7

Seien $f = (a_0, a_1, \ldots, a_{n_0}, 0, \ldots) \in \mathcal{P}(K)$ und $g = (b_0, b_1, \ldots, b_{m_0}, 0, \ldots) \in \mathcal{P}(K)$.
Mit $X = (0, 1, 0, 0, \ldots)$ können wir schreiben:

$$f = \sum_{i=0}^{n_0} a_i X^i \qquad \text{und} \qquad g = \sum_{j=0}^{m_0} b_j X^j.$$

Die Tatsache, dass $\mathcal{P}(K)$ ein Ring ist und die Formeln in Lemma 4.6 gelten, bedeutet dann, dass wir $f * g$ einfach wie folgt ausmultiplizieren können:

$$f * g = \sum_{i=0}^{n_0} \sum_{j=0}^{m_0} (a_i X^i) * (b_j X^j) = \sum_{i=0}^{n_0} \sum_{j=0}^{m_0} (a_i b_j)(X^i * X^j) = \sum_{i=0}^{n_0} \sum_{j=0}^{m_0} a_i b_j X^{i+j}.$$

(Man braucht sich also gar nicht an die genaue Definition von $*$ zu erinnern.)

Wollen wir für die Unbestimmte X irgendwelche Werte x einsetzen, so müssen wir diese x miteinander multiplizieren, addieren und mit Skalaren aus K multiplizieren können. Dies führt auf folgende

Definition 4.8.
Sei K ein Körper und $\mathcal{A}$ ein K-Vektorraum. Zusätzlich sei eine Verknüpfung $\bullet\colon \mathcal{A} \times \mathcal{A} \to \mathcal{A}$ gegeben, so dass $(\mathcal{A}, +, \bullet)$ ein Ring mit 1 ist. Dann heißt $\mathcal{A}$ eine ***K-Algebra***, wenn $(s\alpha) \bullet \beta = s(\alpha\beta) = \alpha \bullet (s\beta)$ für alle $\alpha, \beta \in \mathcal{A}$ und $s \in K$ gilt.

Beispiel 4.9. Sei K Körper.
(a) $\mathcal{A} = M_n(K)$, mit $\bullet =$ Matrixmultiplikation. Dann ist $\mathcal{A}$ ein Ring, und man sieht sofort, dass $(sA)B = s(AB) = A(sB)$ gilt für alle $A, B \in M_n(K)$ und $s \in K$.
(b) $K[X]$ selbst ist eine K-Algebra; siehe Lemma 4.6.
(c) Ist K' Körper mit $K \subseteq K'$ (mit verträglichen Verknüpfungen wie in Kapitel 2, §1), so ist K' nicht nur ein K-Vektorraum, sondern eine K-Algebra.
Zum Beispiel ist $\mathbb{R}$ eine $\mathbb{Q}$-Algebra. Insbesondere: $K' = K$ ist eine K-Algebra.

Einsetzen: Sei $\mathcal{A}$ eine beliebige K-Algebra (mit Multiplikation $\bullet$) und $\alpha \in \mathcal{A}$ fest.

Wir definieren nun eine Abbildung $\Phi_\alpha\colon \mathcal{P}(K) \to \mathcal{A}$.
Sei $f = (a_n) \in \mathcal{P}(K)$ und $n_0 \geqslant 0$ mit $f = (a_0, a_1, \ldots, a_{n_0}, 0, \ldots)$. Dann setze

$$(*) \qquad \Phi_\alpha(f) = a_0 1_{\mathcal{A}} + a_1\alpha + \ldots + a_{n_0}\alpha^{n_0} = \sum_{i=0}^{n_0} a_i\alpha^i,$$

(Hier definiere $\alpha^0 := 1_{\mathcal{A}}$, $\alpha^1 := \alpha$ und $\alpha^i := \alpha \bullet \alpha^{i-1}$ für $i \geqslant 2$.) Beachte auch, dass dies unabhängig von der Wahl von n_0 ist. Setzen wir wie oben $e := (1, 0, 0, \ldots)$ und $X := (0, 1, 0, 0, \ldots)$, so gilt $\Phi_\alpha(e) = 1_{\mathcal{A}}$ und $\Phi_\alpha(X) = \alpha$.

Anders ausgedrückt: Schreiben wir f in der Form $f = a_0 + a_1X + \ldots + a_{n_0}X^{n_0}$, so erhalten wir $\Phi_\alpha(f) \in \mathcal{A}$, indem $\alpha \in \mathcal{A}$ für X "eingesetzt" wird; daher heißt Φ_α auch ***Einsetzungs-Homomorphismus***: Kurzschreibweise $f(\alpha) := \Phi_\alpha(f) \in \mathcal{A}$.

Satz 4.10. Sei $\alpha \in \mathcal{A}$ und $\Phi_\alpha\colon \mathcal{P}(K) \to \mathcal{A}$ wie oben definiert.
Dann ist Φ_α linear und es gilt $\Phi_\alpha(f * g) = \Phi_\alpha(f) \bullet \Phi_\alpha(g)$ für alle $f, g \in \mathcal{P}(K)$.

Beweis. Zuerst zeigen wir, dass Φ_α linear ist. Seien $f = (a_n)$ und $g = (b_n)$ in $\mathcal{P}(K)$. Sei $f = (a_0, a_1, \ldots, a_{n_0}, 0, \ldots)$ und $g = (b_0, b_1, \ldots, b_{n_0}, 0, \ldots)$ mit $n_0 \geqslant 1$.
Dann ist $f + g = (a_0 + b_0, a_1 + b_1, \ldots, a_{n_0} + b_{n_0}, 0, \ldots)$ und damit

$$\Phi_\alpha(f + g) = \sum_{i=0}^{n_0}(a_i + b_i)\alpha^i = \sum_{i=0}^{n_0}(a_i\alpha^i + b_i\alpha^i) = \Big(\sum_{i=0}^{n_0} a_i\alpha^i\Big) + \Big(\sum_{i=0}^{n_0} b_i\alpha^i\Big)$$

und die rechte Seite ist gleich $\Phi_\alpha(f) + \Phi_\alpha(g)$.
Sei nun $s \in K$. Dann ist $sf = (sa_0, sa_1, \ldots, sa_{n_0}, 0, \ldots)$ und damit

$$\Phi_\alpha(sf) = \sum_{i=0}^{n_0}(sa_i)\alpha^i = \sum_{i=0}^{n_0} s(a_i\alpha^i) = s\sum_{i=0}^{n_0} a_i\alpha^i = s\Phi_\alpha(f).$$

Also ist Φ_α linear. Nun betrachte das Produkt $f * g$. Mit Bemerkung 4.7 folgt

$$\Phi_\alpha(f * g) = \Phi_\alpha\Big(\sum_{i=0}^{m_0}\sum_{j=0}^{n_0} a_ib_jX^{i+j}\Big) = \sum_{i=0}^{m_0}\sum_{j=0}^{n_0} a_ib_j\Phi_\alpha(X^{i+j}) = \sum_{i=0}^{n_0}\sum_{j=0}^{n_0} a_ib_j\alpha^{i+j}.$$

Analog erhalten wir durch Ausmultiplizieren in der K-Algebra $\mathcal{A}$:

$$\Phi_\alpha(f) \bullet \Phi_\alpha(g) = \sum_{i=0}^{n_0}\sum_{j=0}^{n_0}(a_i\alpha^i) \bullet (b_j\alpha^j) = \sum_{i=0}^{n_0}\sum_{j=0}^{n_0} a_ib_j(\alpha^i \bullet \alpha^j) = \sum_{i=0}^{n_0}\sum_{j=0}^{n_0} a_ib_j\alpha^{i+j}.$$ $\square$

Beispiel 4.11. Sei $\mathcal{A} = K$ und $c \in K$ fest.
Für $f = (a_0, a_1, \ldots, a_{n_0}, 0, 0, \ldots) \in \mathcal{P}(K)$ ist dann

$$f(c) = \Phi_c(f) = a_0 + a_1c + \ldots + a_{n_0}c^{n_0} = \dot{f}(c),$$

wobei $\dot{f} \in P(K)$ die zu f gehörige Polynomfunktion ist, siehe Bemerkung 4.5.
Analog zu Kapitel 2, Lemma 3.2, gibt es auch ein Horner-Schema für Polynome:
Sei $0 \neq f \in K[X]$ mit $n = \text{Grad}(f) \geqslant 1$. Ist $c \in K$, so gilt $f = (X - c) * g + f(c)$ mit $0 \neq g \in K[X]$ und $\text{Grad}(g) = n - 1$. Ist $f(c) = 0$, so heißt c eine ***Nullstelle*** von f.
Aus Kapitel 2, Folgerung 3.3, folgt sofort, dass f höchstens n Nullstellen hat.

Beispiel 4.12. Sei $\mathcal{A} = M_n(K)$ und $A \in M_n(K)$ fest.
Für $f = (a_0, a_1, \ldots, a_{n_0}, 0, 0, \ldots) \in \mathcal{P}(K)$ ist dann

$$f(A) = \Phi_A(f) = a_0I_n + a_1A + \ldots + a_{n_0}A^{n_0} \in M_n(K).$$

Sei z.B. $f = 2X^3 + X - 3$ und $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$. Nun ist $f = (X - 1) * (2X^2 + 2X + 3)$.

Also folgt mit Satz 4.10 auch $f(A) = (A - I_2)(2A^2 + 2A - 3I_3) = \ldots = \begin{bmatrix} -1 & 5 \\ 5 & 4 \end{bmatrix}$.

§5 Das Minimalpolynom einer Matrix und der Beweis von Satz 3.9

Sei $A \in M_n(K)$. Sei $d \geqslant 1$ minimal, so dass es $a_0, a_1, \ldots, a_{d-1} \in K$ gibt mit $A^d + a_{d-1}A^{d-1} + \ldots + a_1 A + a_0 I_n = 0_{n\times n}$. Dann heißt $d = d(A)$ der Minimalgrad von A; die a_i sind dabei eindeutig bestimmt (siehe Kapitel 1, §4).
Dann definieren wir das ***Minimalpolynom*** von A als

$$\tilde{\mu}_A := a_0 + a_1 X + \ldots + a_{d-1}X^{d-1} + X^d \in K[X].$$

Die zugehörige Polynomfunktion

$$\mu_A \colon K \to K, \qquad x \mapsto \tilde{\mu}_A(x) = a_0 + a_1 x + \ldots + a_{d-1}x^{d-1} + x^d,$$

ist dann die Minimalgleichung für A, wie ursprünglich in Kapitel 1, §4, definiert.

Bemerkung 5.1. Mit den obigen Bezeichnungen gilt:

(a) $\operatorname{Grad}(\tilde{\mu}_A) = d(A)$ und $\tilde{\mu}_A(A) = 0_{n\times n}$.

(b) $\lambda \in K$ ist ein Eigenwert von A genau dann, wenn $\tilde{\mu}_A(\lambda) = 0$.

(c) Sei $0 \neq f \in K[X]$ beliebig mit $f(A) = 0_{n\times n}$. Dann gilt $d(A) \leqslant \operatorname{Grad}(f)$. Ist hier $d(A) = \operatorname{Grad}(f)$, so folgt $f = c\tilde{\mu}_A$ mit einem $0 \neq c \in K$.

Beweis. (a) Dies ist klar nach Definition von $\tilde{\mu}_A$. (b) ist Kapitel 1, Satz 4.9.
(c) Sei $m := \operatorname{Grad}(f) \geqslant 0$ und $f = b_0 + b_1 X + \ldots + b_m X^m$ mit $b_m \neq 0$. Dann ist $0_{n\times n} = f(A) = b_0 I_n + b_1 A + \ldots + b_m A^m$. Nach Multiplikation mit b_m^{-1} erhalten wir eine Gleichung $A^m + b'_{m-1}A^{m-1} + \ldots + b'_1 A + b_0 I_n = 0_{n\times n}$ mit $b'_i := b_m^{-1}b_i \in K$, also folgt $d(A) \leqslant m$. Ist $d(A) = m$, so folgt aus der Eindeutigkeit der a_i, dass $a_i = b'_i$ gilt für $0 \leqslant i \leqslant m-1$. Also ist $b_i = b_m a_i$ für $0 \leqslant i \leqslant m-1$ und damit $f = b_m \tilde{\mu}_A$. □

Bemerkung 5.2. Sind $A, B \in M_n(K)$ ähnlich, so gilt $\tilde{\mu}_A = \tilde{\mu}_B$ (Beweis siehe Ü15).

Definition 5.3. Sei $0 \neq f \in K[X]$.
Gibt es $c_1, \ldots, c_r \in K$ und ein $0 \neq c \in K$ mit $f = c(X - c_1)*(X - c_2)*\ldots*(X - c_r)$, so heißt f ***zerfallend***. Sind die c_i alle verschieden, so heißt f ***einfach-zerfallend***.

Zum Beispiel sei $f = X^2 + 1 \in K[X]$. Für $K = \mathbb{R}$ ist f nicht zerfallend; für $K = \mathbb{C}$ ist $f = (X + i) * (x - i)$ einfach-zerfallend. Für $K = \mathbb{F}_2$ ist $f = (X + 1)^2$ zerfallend, aber nicht einfach-zerfallend. (Hier Matrizen und Polynome in GAP zeigen.)

Sei $0 \neq f \in K[X]$ mit $n = \operatorname{Grad}(f) \geqslant 1$. Sei $c \in K$ eine Nullstelle von f, also $f(c) = 0$. Mit dem Horner-Schema für Polynome folgt, dass es ein $0 \neq g \in K[X]$ gibt mit $\operatorname{Grad}(g) = n - 1$ und $f = (X - c) * g$. Jede Nullstelle von f führt also zur Abspaltung eines Faktors $X - c$ von f. Nun gilt:

Lemma 5.4. Sei $0 \neq f \in K[X]$ und $r := \operatorname{Grad}(f) \geqslant 1$.
Genau dann ist f einfach-zerfallend, wenn die zugehörige Polynomfunktion $\dot{f} \colon K \to K$ genau r paarweise verschiedene Nullstellen in K hat.

Beweis. Sei zuerst f einfach-zerfallend, also $f = c(X - c_1) * \ldots * (X - c_r)$ mit $0 \neq c \in K$ und paarweise verschiedenen $c_i \in K$. Dann ist

$$\dot{f}(x) = f(x) = \Phi_x(f) = c\Phi_x(X - c_1)\cdots\Phi_x(X - c_r) = c(x - c_1)\cdots(x - c_r)$$

für alle $x \in K$; also hat $\dot{f}$ genau r paarweise verschiedene Nullstellen.
Umgekehrt besitze $\dot{f}$ genau r paarweise verschiedene Nullstellen; seien diese $c_1, \ldots, c_r \in K$. Dann ist $f(c_i) = \dot{f}(c_i) = 0$ für $1 \leqslant i \leqslant r$. Argumentiere dann so wie in Aufgabe 11.5(c), wobei man obige Abspaltungsregel wiederholt benutzt. □

Lemma 5.5. Sei $A \in M_n(K)$ diagonalisierbar.
Dann ist $\tilde{\mu}_A \in K[X]$ einfach-zerfallend, also gilt auch die Bedingung in Satz 3.9.

Beweis. Sei B Basis eine von K^n aus Eigenvektoren von A; insbesondere gibt es Eigenwerte. Seien $\lambda_1, \ldots, \lambda_r \in K$ die (verschiedenen) Eigenwerte von A.
Wir setzen $f := (X - \lambda_1) * \ldots * (X - \lambda_r) \in K[X]$. Dann ist $\operatorname{Grad}(f) = r \geqslant 1$.
Da $K[X]$ kommutativ ist, können wir die Faktoren in f beliebig vertauschen. Für $1 \leqslant i \leqslant r$ sei $g_i \in K[X]$ das Produkt der $r - 1$ Faktoren $X - \lambda_j$ mit $j \neq i$. Dann ist

$$f = (X - \lambda_i) * g_i = g_i * (X - \lambda_i) \quad \text{und} \quad \operatorname{Grad}(g_i) = r - 1.$$

Sei nun $v \in B$ beliebig; dann gibt es ein i, so dass v Eigenvektor mit Eigenwert λ_i ist, also $Av = \lambda_i v$ und damit $(A - \lambda_i I_n)v = 0_n$. Es folgt $f(A) = g_i(A)(A - \lambda_i I_n)$, also $f(A)v = g_i(A)((A - \lambda_i I_n)v) = 0_n$. Da dies für alle $v \in B$ gilt, folgt $f(A) = 0_{n\times n}$, und damit $d := \operatorname{Grad}(\tilde{\mu}_\varphi) \leqslant \operatorname{Grad}(f) = r$.
Da $\tilde{\mu}_\varphi(\lambda_i) = 0$ gilt für $1 \leqslant i \leqslant r$ und $\tilde{\mu}_\varphi$ höchstens d Nullstellen hat, folgt $d = r$.
Also erhalten wir $\tilde{\mu}_\varphi = cf$ mit $0 \neq c \in K$; siehe Bemerkung 5.1. □

Satz 5.6 (Charakterisierung diagonalisierbarer Abbildungen, vgl. Satz 3.9)

Sei $A \in M_n(K)$. Dann sind äquivalent:

(a) A ist diagonalisierbar.

(b) Das Minimalpolynom $\tilde{\mu}_A \in K[X]$ ist einfach-zerfallend.

(c) Es gibt ein einfach-zerfallendes Polynom $0 \neq f \in K[X]$ mit $f(A) = 0_{n \times n}$.

Beweis. "(a) $\Rightarrow$ (b)" ist Lemma 5.5. Für "(b) $\Rightarrow$ (c)" setze einfach $f := \tilde{\mu}_\varphi$.
Nun zu "(c) $\Rightarrow$ (a)": Sei $f = c(X - \lambda_1) * \ldots * (X - \lambda_r)$, mit $0 \neq c \in K$ und paarweise verschiedenen $\lambda_1, \ldots, \lambda_r \in K$. Wie oben gibt es $0 \neq g_k \in K[X]$ mit $f = (X - \lambda_k) * g_k = g_k * (X - \lambda_k)$ und Grad$(g_k) = r - 1$ für $1 \leqslant k \leqslant r$. Beachte:
Weil $\lambda_1, \ldots, \lambda_r$ paarweise verschieden sind, ist $g_k(\lambda_k) \neq 0$ und $g_k(\lambda_i) = 0$ für $i \neq k$.
Behauptung: $\quad g := g_1(\lambda_1)^{-1} g_1 + \ldots + g_r(\lambda_r)^{-1} g_r = 1.$
Dazu: Wegen $g_k(\lambda_i) = 0$ für $i \neq k$ folgt $g(\lambda_i) = 1$ für alle i. Also hat $g - 1 \in K[X]$ die r Nullstellen $\lambda_1, \ldots, \lambda_r$. Wegen $g \neq 0$ und Grad$(g) \leqslant r - 1$ folgt $g - 1 = 0$.

Wir haben also $1 = a_1 g_1 + \ldots + a_r g_r$ mit $a_k := g_k(\lambda_k)^{-1} \in K$ für $1 \leqslant k \leqslant r$.
Nächste Behauptung: $\quad V_k := \{g_k(A)v \mid v \in K^n\} \subseteq E_A(\lambda_k) \quad$ für $1 \leqslant k \leqslant r$.
Dazu: Sei $v \in K^n$ beliebig und $v' := g_k(A)v \in V_k$.
Wegen $f = (X - \lambda_k) * g_k$ und $f(A) = 0_{n \times n}$ folgt

$$0_n = f(A)v = (A - \lambda_k I_n)\big(g_k(A)v\big) = (A - \lambda_k I_n)v',$$

also $Av' = \lambda_k v'$, d.h., $v' \in E_A(\lambda_k)$.

Schließlich zeigen wir: $\quad K^n = E_A(\lambda_1) + \ldots + E_A(\lambda_r)$.
Dazu: Sei $v \in V$ beliebig. Wegen $1 = a_1 g_1 + \ldots + a_r g_r$ (siehe oben) folgt

$$v = I_n v = (a_1 g_1 + \ldots + a_r g_r)(A)v = a_1 g_1(A)v + \ldots + a_r g_r(A)v,$$

und die rechte Seite ist in $V_1 + \ldots + V_r$ enthalten. Wegen $V_k \subseteq E_A(\lambda_k)$ für $1 \leqslant k \leqslant r$ folgt damit auch $K^n = E_A(\lambda_1) + \ldots + E_A(\lambda_r)$.

Also wird K^n von Eigenvektoren von A erzeugt. Aber dann gibt es auch eine Basis von Eigenvektoren, also ist A diagonalisierbar. $\square$

Beispiel 5.7. Sei $K = \mathbb{C}$ und $A \in M_n(\mathbb{C})$.

Behauptung: Gibt es ein $d \geqslant 1$ mit $A^d = I_n$, so ist A diagonalisierbar.
Dazu betrachte $f := X^d - 1 \in \mathbb{C}[X]$. Es gilt $f(A) = 0_{n \times n}$. Nullstellen von f ?
Für $0 \leqslant k \leqslant d - 1$ sei $z_k := e^{2\pi i k/d} = \cos(2\pi k/d) + i \sin(2\pi k/d) \in \mathbb{C}$. Dann $z_k^n = e^{2\pi i k} = 1$, also $f(z_k) = 0$ für $0 \leqslant k \leqslant d - 1$, und damit

$$f = (X - z_0) * (X - z_1) * \ldots * (X - z_{d-1}).$$

Nach Satz 5.6 ist A diagonalisierbar; die Eigenwerte sind in $\{z_0, z_1, \ldots, z_{d-1}\}$.

Allerdings sind diagonalisierbare Matrizen sehr speziell !

Dazu nur ein einfaches Beispiel: Sei $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \in M_4(K)$. Dann gilt:

$$A^2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \qquad A^3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \qquad A^4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = 0_{4 \times 4}.$$

Also $\tilde{\mu}_A = X^4 \rightsquigarrow A$ nicht diagonalisierbar! (Analog für beliebiges n anstelle $n = 4$.)

Eine Matrix $A \in M_n(K)$ heißt ***nilpotent***, wenn es ein $m \in \mathbb{N}$ gibt mit $A^m = 0_{n \times n}$.

Lemma 5.8. Sei $A \in M_n(K)$.

Genau dann ist A nilpotent, wenn $\tilde{\mu}_A = X^d$ für ein $d \geqslant 1$ gilt.

Beweis. Ist $\tilde{\mu}_A = X^d$, so folgt $A^d = \tilde{\mu}_A(A) = 0_{n \times n}$, also ist A nilpotent. Sei nun umgekehrt A nilpotent und $m \geqslant 1$ minimal mit $A^m = 0_{n \times n}$. Dann ist $m \geqslant d$ wobei $d = d(A) \geqslant 1$ der Minimalgrad ist; sei $a_0 I_n + a_1 A + \ldots + a_{d-1} A^{d-1} + A^d = 0_{n \times n}$ mit $a_0, a_1, \ldots, a_{d-1} \in K$. Nehmen wir an, es gibt ein i mit $a_i \neq 0$.
Sei $k = \min\{i \mid a_i \neq 0\}$, also $a_k A^k + a_{k+1} A^{k+1} + \ldots + a_{d-1} A^{d-1} + A^d = 0_{n \times n}$ mit $a_k \neq 0$ und $0 \leqslant k \leqslant d - 1 \leqslant m - 1$. Multipliziere mit A^{m-1-k} und erhalte

$$a_k A^{m-1} + a_{k+1} A^m + \ldots + a_{d-1} A^{d-1+m-k-1} + A^{d+m-k-1} = 0_{n \times n}.$$

Aber $A^m = A^{m+1} = \ldots = A^{d+m-k-1} = 0_{n \times n}$, also bleibt nur $a_k A^{m-1} = 0_{n \times n}$ übrig.
Wegen $a_k \neq 0$ folgt $A^{m-1} = 0_{n \times n}$, Widerspruch zur Minimalität von m.
Also gilt $a_0 = a_1 = \ldots = a_{d-1} = 0$ und damit ist $\tilde{\mu}_A = X^d$. $\square$

Also folgt mit Satz 5.6: $\quad A$ nilpotent und diagonalisierbar $\quad \Leftrightarrow \quad A = 0_{n \times n}$.

Schlussbemerkung.

(a) Ist $K = \mathbb{R}$, so kann man unter Verwendung von analytischen Eigenschaften (sup, inf, Grenzwerte, ...) stärkere Aussagen zeigen, zum Beispiel:
Ist $A \in M_n(\mathbb{R})$ symmetrisch, also $A = A^{\mathrm{tr}}$, so ist A diagonalisierbar; siehe LAAG2.
(b) Ebenso in LAAG2: Was kann man machen, wenn A nicht diagonalisierbar ist ?
Auch: Wie kann man Minimalpolynom effizient berechnen ?

Plan für LAAG2:

- Kapitel 4: Determinanten
- Kapitel 5: Dualraum, Skalarprodukte und multilineare Algebra
- Kapitel 6: Normalformen von Matrizen
- Kapitel 7: Affine und Euklidische Geometrie

Alles Gute für die Prüfung im März und auf Wiedersehen im Sommersemester !