Vorbemerkung: Das Einsetzen von quadratischen Matrizen in Polynome.

Im folgenden sei R ein kommutativer Ring und $R[T]$ der Polynomring mit Koeffizienten in R (dies ist wieder ein kommutativer Ring).

Matrizen mit Koeffizienten in R. Sei $M(n\times n, R)$ die Menge der $(n\times n)$-Matrizen mit Koeffizienten in R, dies ist bezüglich der Addition und Multiplikation von Matrizen (siehe LA I) ein Ring, der für $n \geq 2$ **nicht** kommutativ ist.

Sei $A \in M(n \times n, R)$. Sei $\phi(T) \in R[T]$, etwa $\phi(T) = \sum_{i=0}^{n} c_i T^i$ mit $c_i \in R$. Wir können die Potenzen A^i der Matrix A bilden, wir können diese Potenzen A^i mit c_i multiplizieren (Skalar-Multiplikation), wir können die entstandenen Matrizen $c_i A^i$ addieren: wir erhalten $\sum_{i=0}^{n} c_i A^i$. Dieser Vorgang entspricht dem üblichen Einsetzen von Zahlen in Polynomen, hier haben wir "die Matrix A in das Polynom $\chi(T)$ eingesetzt", man schreibt demnach $\chi(A) = \sum_{i=0}^{n} c_i A^i$.

Beachte: Es ist $A^1 = A$ und $A^0 = E_n$ (die $(n \times n)$-Einheitsmatrix). Insbesondere gilt: Ist ϕ ein konstantes Polynom, etwa $\phi = c$ mit $c \in R$, so ist $\phi(A) = cE_n$, also die Skalarmatrix, deren Diagonalkoeffizienten alle gleich c sind. Wichtig ist, dass die Skalarmatrizen mit unserer Matrix A kommutieren!

Sei $A \in M(n \times n, R)$. Die Abbildung $e_A \colon R[T] \to M(n \times n, R)$ mit $e_A(\phi) = \phi(A)$ für $\phi = \phi(T) \in R[T]$ ist ein Ring-Homomorphismus.

Beweis: Zu zeigen ist für $\phi_1, \phi_2 \in R[T]$:

$$\begin{aligned} e_A(\phi_1 + \phi_2) &= e_A(\phi_1) + e_A(\phi_2), \\ e_A(\phi_1\phi_2) &= e_A(\phi_1)e_A(\phi_2), \\ e_A(1) &= E_n. \end{aligned}$$

Man rechnet dies sofort nach — man muss dabei natürlich die Definition der Addition und der Multiplikation von Polynomen verwenden (und auch die Kommutativität des Rings R).

Lemma. *Ist $A \in M(n \times n, R)$ und $\phi \in R[T]$, so gilt $\phi(A) \cdot A = A \cdot \phi(A)$.* Oder auch allgemeiner: Sind $\phi, \psi \in R[T]$, so gilt $\phi(A)\psi(A) = \psi(A)\phi(A)$.

Beweis: Sei $\phi = \sum_{i=0}^{n} c_i A^i$ mit $c_i \in R$. Natürlich gilt:

$$A^i \cdot A = A^{i+1} = A \cdot A^i, \quad \text{und} \quad c_i A = A c_i,$$

also auch

$$\phi(A) \cdot A = \sum_{i=0}^{n} c_i A^i \cdot A = A \cdot \sum_{i=0}^{n} c_i A^i = A \cdot \phi(A).$$

Sucht man also zu einer Matrix $A \in M(n\times n, R)$ Matrizen B, die mit A vertauschen, so wird man als erstes an Matrizen der Form $B = \phi(A)$ mit $\phi \in R[T]$ denken!

4.5 Invariante Unterräume.

4.5.1. Definition. Sei $A \in M(n \times n, K)$. Ein Unterraum $U \subseteq K^n$ heißt *A-invariant,* falls gilt: Ist $u \in U$, so ist $Au \in U$. Entsprechend wird definiert: Ist $f\colon V \to V$ ein Endomorphismus eines Vektorraums V, so heißt ein Unterraum $U \subseteq V$ *f-invariant,* falls $f(u) \in U$ für alle $u \in U$ gilt. (Statt von A-invarianten oder f-invarianten Unterräumen spricht man auch einfach von *invarianten* Unterräumen, wenn klar ist, welche Matrix A oder welcher Endomorphismus f gemeint ist.)

Ist v ein Eigenvektor von $f\colon V \to V$, so ist der von v erzeugte Unterraum $L(v)$ ein f-invarianter Unterraum von V, der eindimensional ist. Auch umgekehrt gilt: Ist U ein eindimensionaler f-invarianter Unterraum von V, so ist jeder von Null verschiedene Vektor $v \in U$ ein Eigenvektor. *Eindimensionale A-invariante oder f-invariante Unterräume sind also gerade Unterräume, die von einem Eigenvektor erzeugt werden.* Die Betrachtung invarianter Unterräume verallgemeinert demnach das Arbeiten mit Eigenvektoren.

4.5.2. Lemma (a) *Sind $A, B \in M(n \times n, K)$ mit $AB = BA$, so ist*

$$\mathrm{Kern}(l_B) = \{v \in K^n \mid Bv = 0\}$$

ein A-invarianter Unterraum.

Beweis: Sei $Bu = 0$. Dann ist auch $B(Au) = 0$, denn $BAu = ABu = A0 = 0$.

(b) *Ist $A \in M(n \times n, K)$ und $\phi \in K[T]$, so ist $\mathrm{Kern}(l_{\phi(A)})$ ein A-invarianter Unterraum.*

Beweis: Die obige Vorbemerkung besagt gerade, dass $B = \phi(A)$ mit A kommutiert.

Ein ganz wichtiges Verfahren, um A-invariante Unterräume zu finden ist also das folgende: Man nimmt ein Polynom $\phi \in K[T]$, bildet $B = \phi(A)$ und bestimmt $\mathrm{Kern}(l_B)$.

4.5.3. Hauptsatz. *Sei $A \in M(n \times n, K)$, sei $\phi \in K[T]$ mit*

$$\phi(A) = 0. \tag{1}$$

Es sei

$$\phi = \phi_1 \cdots \phi_t$$

mit paarweise teilerfremden Polynomen $\phi_1, \cdots, \phi_t$. Für jedes i setze

$$U_i = \{v \in K^n \mid \phi_i(A)v = 0\}.$$

Dann gilt:

$$K^n = \bigoplus\nolimits_{i=1}^{t} U_i.$$

Warum ist man an solchen Unterräumen U_i interessiert? In 4.5.2 wurde gezeigt: *Alle Unterräume U_i sind A-invariant.* Man erhält also eine direkte Summen-Zerlegung $K^n = \bigoplus U_i$ mit **A-invarianten** Unterräumen U_i. Nun sehe man sich 4.5.4 an.

Beweis des Hauptsatzes: Nach Definition von U_i gilt:

$$u \in U_i \iff \phi_i(A)v = 0. \tag{2}$$

Setze $\phi_i^* = \phi_1 \cdots \phi_{i-1}\phi_{i+1} \cdots \phi_t$, also gilt

$$\phi = \phi_i \phi_i^*. \tag{3}$$

(4) Ist $v \in V$, so ist $\psi_i(A)\phi_i^*(A)v \in U_i$.

Beweis:

$$\phi_i(A)\psi_i(A)\phi_i^*(A)v = \psi_i(A)\phi_i(A)\phi_i^*(A)v \underset{(3)}{=} \psi_i(A)\phi(A)v \underset{(1)}{=} 0.$$

Es gilt:

$$i \neq j \implies \phi_i \mid \phi_j^*, \tag{5}$$

also

$$u_i \in U_i,\ j \neq i \implies \phi_j^*(A)(u_i) = 0. \tag{6}$$

Beweis: Wegen $i \neq j$, ist ϕ_i ein Teiler von ϕ_j^* (Aussage (5)), also verwende (2).

Und es gilt auch:

(7) Die Polynome ϕ_i, ϕ_i^* sind teilerfremd.

Daraus folgt:

$$\mathrm{ggT}(\phi_1^*, \ldots, \phi_t^*) = 1. \tag{8}$$

Beweis: Sei ζ irreduzibles Polynom, das alle ϕ_i^* teilt. Das Polynom ζ teilt also ϕ, also ein ϕ_i, aber dann nicht ϕ_i^*. (Beachte: ϕ_i, ϕ_i^* sind teilerfremd).

Nach Bézout gibt es Polynome ψ_i mit

$$\sum \psi_i \phi_i = 1. \tag{9}$$

(10) Ist $v \in V$, so ist

$$v = E_n v \underset{(9)}{=} \sum_{i=1}^{t} \psi_i(A)\phi_i^*(A)v \in \sum_i U_i$$

(wegen (4)).

Sei nun umgekehrt $u_i \in U_i$ mit $\sum u_i = 0$. Wende $\phi_s^*(A)$ an:

$$0 = \phi_s^*(A)(\sum u_i) = \sum_i \phi_s^*(A)u_i \underset{(6)}{=} \phi_s^*(A)u_s. \tag{11}$$

Also

$$u_s = E_n u_s \underset{(9)}{=} \sum_i \psi_i(A)\phi_i^*(A)u_s \underset{(6)}{=} \psi_s(A)\phi_s(A)^* u_s \underset{(11)}{=} 0$$

Damit ist gezeigt:

$$V = \bigoplus U_i$$

Zusatz: Ist $u_j \in U_j$, so gilt:

$$\psi_i(A)\phi_i^*(A)u_j = \delta_{ij}u_i \tag{12}$$

(dabei ist δ_{ij} das "Kronecker-Delta", also $\delta_{ii} = 1$ und $\delta_{ij} = 0$ für $i \neq j$).

Beweis: Ist $i \neq j$, so ist dies Aussage (6); ist i = j, so ist dies Bezout (9) zusammen mit (6).

> **Einschub: Projektionsabbildungen.** Sind U, U' Unterräume von V mit $V = U \oplus U'$, und definiert man $p : V \to V$ durch $p(u + u') = u$ für $u \in U$ und $u' \in U'$, so ist p wohl-definiert und linear, und das Bild von p ist U, der Kern ist U'. Und es ist $p^2 = p$. Man nennt p die *Projektionsabbildung* mit Bild U und Kern U'.

Man sieht also: *Die Abbildung $\psi_i(A)\phi_i^*(A)$ ist Projektionsabbildung mit Bild U_i und Kern $\bigoplus_{j\neq i} U_j$.*

Spezialfall. *Sei $A \in M(n \times n, K)$, seien $\lambda_1, \ldots, \lambda_t$ paarweise verschiedene Elemente von K, sei*

$$\phi = (T - \lambda_1)^{n_1}(T - \lambda_2)^{n_2} \cdots (T - \lambda_t)^{n_t} \in K[T]$$

und es gelte

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

Für jedes i setze

$$U_i = \{v \in K^n \mid (A - \lambda_i)^{n_i} v = 0\}.$$

Dann gilt:

$$K^n = \bigoplus_{i=1}^{t} U_i.$$

Beweis: Setzt man $\phi_i = (T - \lambda_i)^{n_i}$, so sind die Polynome ϕ_i paarweise teilerfremd, es sind also die Voraussetzungen des Hauptsatzes erfüllt. Wegen $\phi_i(A) = (A - \lambda_i)^{n_i}$ sind die Unterräume U_i genau diejenigen, die im Hauptsatz betrachtet werden.

4.5.4. Wofür braucht man A-invariante Unterräume?

Offensichtlich gilt: Sei U ein A-invarianter Unterraum von K^n der Dimension m. Sei $u_1, \ldots, u_m$ eine Basis von U, setze sie durch $u_{m+1}, \ldots u_n$ zu einer Basis von K^n fort. Bezüglich dieser Basis $(u_1, \ldots, u_n)$ hat l_A eine Matrizen-Darstellung der Form

$$\begin{bmatrix} B & D \\ 0 & C \end{bmatrix},$$

dabei ist B eine $(m \times m)$-Matrix.

Sind A-invariante Unterräume U, V mit $U \cap V = 0$ und $U + V = K^n$ gegeben, so nennt man das Paar (U, V) eine *A-invariante Zerlegung* des K^n. Wählt man eine Basis $u_1, \ldots, u_m$ von U und eine Basis $u_{m+1}, \ldots, u_n$ von V, so hat l_A bezüglich dieser Basis $(u_1, \ldots, u_n)$ eine Matrizen-Darstellung der Form

$$\begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix},$$

dabei ist B wieder eine $(m \times m)$-Matrix.

Ist $f\colon V \to V$ ein Endomorphismus und lässt sich V als direkte Summe $V = \bigoplus_{i=1}^{s} U_i$ schreiben mit f-invarianten Unterräumen U_i, und wählen wir Basen der Unterräume U_{γ_i}, so erhalten wir als Vereinigung eine Basis von K^n und bezüglich dieser Basis hat f eine Matrizendarstellung der Form

$$A = \begin{bmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & A_s \end{bmatrix},$$

dabei ist A_i eine Matrizendarstellung der Einschränkung der Abbildung f auf U_i. Zur Abkürzung schreibt man $A = \bigoplus_{i=1}^{s} A_i$ oder auch nur $A = \bigoplus_i A_i$.

4.6. Der Satz von Cayley-Hamilton.

Im folgenden sei R ein kommutativer Ring und $R[T]$ der Polynomring mit Koeffizienten in R (auch dies ist wieder ein kommutativer Ring).

4.6.1. Matrizen mit Koeffizienten in R. Sei $M(n \times n, R)$ die Menge der $(n \times n)$-Matrizen mit Koeffizienten in R, dies ist bezüglich der Addition und Multiplikation von Matrizen (siehe LA I) ein Ring, der allerdings für $n \geq 2$ **nicht** kommutativ ist. Ist $A \in M(n \times n, R)$, so sei $\det A$ mit Hilfe der Leibniz-Formel (LA I, 2.4.19) definiert; det ist also eine Abbildung $M(n \times n, R) \to R$.

In der Vorlesung LA I wurde nur der Fall $R = K$ ein Körper betrachtet. Zwei Ergebnisse, die dort gezeigt wurden, gelten ganz allgemein auch für Matrizen mit Koeffizienten in einem beliebigen kommutativen Ring:

Produktformel für Determinanten. *Sind* $A, B \in M(n \times n, R)$, *so ist*

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

Zweitens definiert man für jede Matrix $A \in M(n \times n, R)$ die zugehörige *Adjunkte* A^* auf folgende Weise (wie im Fall $R = K$, siehe LA I, 2.4.22): Sei A eine $n \times n$ Matrix. Setze $A^* = (a'_{ij})_{ij}$ mit $a'_{ij} = (-1)^{i+j} \det A_{ji}$, dabei entsteht die Matrix A_{ji} aus der Matrix A, indem man die j-te Zeile und die i-te Spalte streicht (A_{ji} ist also eine $(n-1) \times (n-1)$-Matrix, wieder mit Koeffizienten in R. (Beim Bilden von A^* muss man die Indexvertauschung beachten!)

Satz (Laplace). *Sei* $A \in M(n \times n, R)$. *Es ist* $AA^* = A^*A = \det A \cdot E_n$.

Siehe LA I, 2.4.23. Dieser Satz fasst die verschiedenen Möglichkeiten, den Laplaceschen Entwicklungssatz auf die gegebene Matrix A anzuwenden, in einer einzigen Formel zusammen.

4.6.2. Noch einmal: Das charakteristische Polynom einer Matrix. Sei R ein kommutativer Ring, sei $A \in M(n \times n, R)$. Man nennt

$$\chi_A(T) = \det(T \cdot I_n - A)$$

(dabei ist $T \cdot I_n - A$ eine $(n \times n)$-Matrix mit Koeffizienten im Polynomring $R[T]$, und zwar die Differenz der Skalarmatrix $T \cdot I_n$ und der Matrix A, daher ist $\chi_A = \chi_A(T) \in R[T]$). das **charakteristische Polynom** von A.

4.6.3 Satz von Cayley-Hamilton (CAYLEY 1821-1895, HAMILTON 1805-1865). *Sei* R *ein kommutativer Ring, sei* $A \in M(n \times n, R)$. *Es ist* $\chi_A(A) = 0$.

Was besagt dieser Satz? Zur Matrix A ist das charakteristische Polynom $\chi_A(T)$ definiert; in dieses Polynom können wir die Matrix A einsetzen, also $\chi_A(A)$ bilden. Was wir erhalten, ist die **Null-Matrix**.

Beweis des Satzes von Cayley-Hamilton: Sei $B = T \cdot E_n - A$. Es ist $\det B$ das charakteristische Polynom von $\chi_A(T)$, dies sei das Polynom $\det B = \chi_A(T) = \sum_{i=0}^{n} c_i T^i$.

Wir bilden wie üblich die zu B komplementäre Matrix B^*. Beachte: die Koeffizienten von B^* sind Polynome mit Koeffizienten in R, und ihr Grad ist höchstens $n-1$. Also kann man B^* in der Form

$$B^* = \sum_{i=0}^{n-1} B_i T^i \quad \text{mit} \quad B_i \in M(n \times n, R)$$

schreiben. Der Satz von Laplace besagt:

$$(*) \qquad BB^* = \det(B) \cdot E_n = \chi_A \cdot E_n = \sum_{i=0}^{n} c_i E_n T^i.$$

Das Produkt BB^* können wir umformen (einsetzen, distributiv rechnen):

$$\begin{aligned} BB^* &= \big(E_n \cdot T - A\big)\Big(\sum_{i=0}^{n-1} B_i T^i\Big) \\ &= \sum_{i=0}^{n-1} B_i T^{i+1} - \sum_{i=0}^{n-1} AB_i T^i \\ &= B_{n-1}T^n + \sum_{i=1}^{n-1}\big(B_{i-1} - AB_i\big)T^i - AB_0 \end{aligned} \qquad (**)$$

Die Matrix BB^* ist in $(*)$ und $(**)$ auf zwei verschiedene Weisen nach Potenzen von T entwickelt worden. Koeffizienten-Vergleich liefert:

$$\begin{aligned} c_0E_n &= & -AB_0 \\ c_1E_n &= B_0 & - AB_1 \\ c_2E_n &= B_1 & - AB_2 \\ &\dots & \\ c_{n-1}E_n &= B_{n-2} & - AB_{n-1} \\ c_nE_n &= B_{n-1} & \end{aligned}$$

Wir multiplizieren die Gleichung $c_iE_n = \dots$ von links mit A^i und erhalten

$$\begin{aligned} c_0A^0 &= & -AB_0 \\ c_1A^1 &= AB_0 & - A^2B_1 \\ c_2A^2 &= A^2B_1 & - A^3B_2 \\ &\dots & \\ c_{n-1}A^{n-1} &= A^{n-1}B_{n-2} & - A^nB_{n-1} \\ c_nA^n &= A^nB_{n-1} & \end{aligned}$$

Wenn wir nun die Gleichungen addieren, erhalten wir links $\chi_A(A)$, und rechts die Null-Matrix.

4.6.4. Satz. *Ist ψ ein normiertes Polynom mit $\psi(A) = 0$, so ist χ_A ein Teiler von ψ^n.*

Beweis: Es sei $\psi = T^r + d_{r-1}T^{r-1} + \dots + d_1T + d_0$. Statt E_n schreiben wir einfach E. Wir bilden r Matrizen $B_0, \dots, B_{r-1}$, sie werden induktiv mit fallendem Index, konstruiert:

$$\begin{aligned} B_{r-1} &= E \\ B_{r-2} &= AB_{r-1} + d_{r-1}E \\ B_{r-3} &= AB_{r-2} + d_{r-2}E \\ &\dots \\ B_0 &= AB_1 + d_1E \end{aligned}$$

Dies können wir folgendermaßen umschreiben:

$$\begin{aligned} B_{r-1} &= E \\ B_{r-2} - AB_{r-1} &= d_{r-1}E \\ B_{r-3} - AB_{r-2} &= d_{r-2}E \\ &\cdots \\ B_0 - AB_1 &= d_1E \end{aligned}$$

Nun multiplizieren wir jeweils mit einer geeigneten Potenz von A:

$$\begin{aligned} A^rB_{r-1} &= A^r \\ A^{r-1}B_{r-2} - A^rB_{r-1} &= d_{r-1}A^{r-1} \\ A^{r-2}B_{r-3} - A^{r-1}B_{r-2} &= d_{r-2}A^{r-2} \\ &\cdots \\ A^2B_1 - A^3B_2 &= d_2A \\ AB_0 - A^2B_1 &= d_1A \end{aligned}$$

und addieren die Gleichungen auf. Links erhalten wir AB_0, rechts erhalten wir fast $\psi(A)$, es fehlt nur der konstante Term. Also

$$AB_0 + d_0E = \sum_{i=1}^{r} d_iA^i + d_0E = \psi(A) = 0.$$

also

$$AB_0 = -d_0E.$$

Wir bilden die folgende Matrix $B(T) \in M(n \times n, R[T])$

$$B = B_{r-1}T^{r-1} + B_{r-2}T^{r-2} + \cdots + B_1T + B_0.$$

Behauptung: Es gilt

$$(*) \qquad (ET - A) \cdot B = \psi(T) \cdot E$$

(rechts steht die Diagonalmatrix, deren Hauptdiagonalkoeffizienten alle gleich $\psi(T)$ sind).

Beweis:

$$\begin{aligned} (ET - A) \cdot B &= \big(B_{r-1}T^r + B_{r-2}T^{r-1} + \cdots + B_1T^2 + B_0T\big) \\ &\quad - \big(AB_{r-1}T^{r-1} + AB_{r-2}T^{r-2} + \cdots + AB_1T + AB_0\big) \\ &= B_{r-1}T^r + (B_{r-1} - AB_{r-2})T^{r-1} + \cdots + (B_0 - AB_1)T + AB_0 \\ &= ET^r + d_{r-1}ET^{r-1} + \cdots d_1ET + d_0E \\ &= \psi(T) \cdot E. \end{aligned}$$

Für beide Seiten von $(*)$ bilden wir die Determinante und verwenden den Produktsatz für Determinanten:

$$\det(ET - A) \cdot (\det B) = \det((ET - A) \cdot B) = \det(\psi(T) \cdot E) = \psi(T)^n.$$

Also sehen wir: $\chi_A = \det(ET - A)$ ist Teiler von ψ^n.

4.7. Das Minimalpolynom einer quadratischen Matrix.

4.7.1. Lemma. *Seien $\phi, \psi \in K[T]$ normiert mit $\phi(A) = 0$ und $\psi(A) = 0$. Dann gilt auch* $(\mathrm{ggT}(\phi, \psi))(A) = 0$.

Beweis: Verwende Bézout! Es gilt Polynome u, v mit $\mathrm{ggT}(\phi, \psi) = u\phi + v\psi$, also

$$(\mathrm{ggT}(\phi, \psi))(A) = u(A)\phi(A) + v(A)\psi(A) = 0.$$

Es folgt: Ist μ ein normiertes Polynom kleinsten Grads mit $\mu(A) = 0$, so ist μ ein Teiles jedes normierten Polynoms ϕ mit $\phi(A) = 0$, insbesondere also eindeutig bestimmt. Man nennt μ das Minimal-Polynom zur Matrix A.

4.7.2. Das Minimal-Polynom einer quadratischen Matrix. Sei K ein Körper und $A \in M(n \times n, K)$. Sei $\mu_A \in K[T]$ ein normiertes Polynom kleinstmöglichen Grads mit $\mu_A(A) = 0$.

Ist $\phi \in K[T]$ ein Polynom mit $\phi(A) = 0$, so ist μ_A ein Teiler von ϕ.

Beweis: Teile ϕ durch μ_A mit Rest, etwa $\phi = q\mu_A + r$ und setze die Matrix A ein, wir erhalten $r(A) = 0$. Da nach Voraussetzung μ_A den kleinstmöglichen Grad hat, muss $r = 0$ gelten.)

Daraus folgt:

(1) *Das Polynom μ_A ist eindeutig bestimmt* (man nennt μ_A das *Minimalpolynom* von A).

(2) *Ähnliche Matrizen haben das gleiche Minimalpolynom.* (Seien A, B ähnliche Matrizen, also $B = S^{-1}AS$ mit einer invertierbaren $(n \times n)$-Matrix S. Es ist $0 = \mu_B(B) = \mu_B(S^{-1}AS) = S^{-1}\mu_B(A)S$, also $\mu_B(A) = 0$ und demnach ist μ_A ein Teiler von μ_B. Entsprechend ist aber μ_B auch ein Teiler von μ_A).

(3) *Das Minimalpolynom μ_A ist ein Teiler des charakteristischen Polynoms χ_A.* (Dies ist der Satz von Cayley-Hamilton 4.6.3.)

(4) *Das charakteristischen Polynom χ_A ist ein Teiler von μ_A^n, wobei μ_A das Minimalpolynom von A ist.* (Dies ist Satz 4.6.4.)

(5) *Das charakteristische Polynom und das Minimalpolynom haben also die gleichen irreduziblen Faktoren, allerdings möglicherweise mit verschiedenen Multiplizitäten.*

(6) Genauer: *Ist*

$$\chi_A = \phi_1^{n_1} \cdot \phi_2^{n_2} \cdots \phi_t^{n_t}$$

mit paarweise verschiedenen irreduziblen normierten Polynomen ϕ_i, und mit Exponenten $n_i \geq 1$, so ist

$$\mu_A = \phi_1^{m_1} \cdot \phi_2^{m_2} \cdots \phi_t^{m_t}$$

und es gilt:

$$1 \leq m_i \leq n_i, \quad \text{und} \quad 1 \leq n_i \leq n \cdot m_i.$$

Beispiele.

(1) Für die Skalarmatrix $A = cE_n$ mit $c \in K$ ist das Minimalplynom $\mu_A = T - c$ und das charakteristische Polynom $\chi_A = (T-c)^n$.

(2) Für die reelle Diagonal-Matrix $\begin{bmatrix} 1 & & & \\ & 2 & & \\ & & 3 & \\ & & & 3 \end{bmatrix}$ lauten Minimalplynom μ_A und charakteristisches Polynom χ_A wie folgt:

$$\mu_A = (T-1)(T-2)(T-3) \quad \text{und} \quad \chi_A = (T-1)(T-2)(T-3)^2.$$

(3) Die Matrix

$$J(n) = (a_{ij}) = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix} \in M(n \times n, K)$$

wird durch die Gleichungen $a_{i,i+1} = 1$ für $1 \leq i < n$ und $a_{ij} = 0$ für alle i, j mit $j \neq i+1$, beschrieben. Für $J(n)$ gilt: $\mu_{J(n)} = \chi_{J(n)} = T^n$.

(4) Entsprechend gilt für $J(\lambda, n) = J(n) + E_n$ mit $\lambda \in K$: Es ist $\mu_{J(\lambda,n)} = \chi_{J(\lambda,n)} = (T-\lambda)^n$.

4.7.3. Lemma. *Sei K ein Körper, sei $A \in M(n \times n, K)$. Hat das Minimalpolynom μ_A den Grad m, so gilt: Alle Potenzen A^i mit $i \in \mathbb{N}_0$ liegen im Unterraum $L(I, A, A^2, \ldots, A^{m-1})$ des Vektorraums $M(n \times n, K)$.*

(Erinnerung, LA I, 1.5.1.: Sind $v_1, \ldots, v_t$ Vektoren im Vektorraum V, so sei $L(v_1, \ldots, v_t)$ der von diesen Vektoren erzeugte Unterraum.)

Beweis: Sei $\mu_A(T) = \sum_{j=0}^m a_j T^j$ mit $a_j \in K$ und $a_m = 1$. Wegen $\mu_A(A) = 0$ gilt $\sum_{j=0}^m a_j A^j = 0$, also

$$A^m = -\sum_{j=0}^{m-1} a_j A^j.$$

Multiplizieren wir diese Gleichung mit A^s mit $s \geq 0$, so erhalten wir

$$A^{m+s} = -\sum_{j=0}^{m-1} a_j A^{j+s}.$$

Dies zeigt: Die Matrizen der Form A^{m+s} mit $s \geq 0$ lassen sich als Linearkombinationen von Matrizen der Form A^r mit $r < m + s$ schreiben. Induktiv sieht man, dass sich die

Matrizen der Form A^{m+s} mit $s \geq 0$ als Linearkombination der Matrizen der Form A^r mit $r < m$ schreiben lassen.

Folgerung. *Sei K ein Körper, sei $A \in M(n \times n, K)$. Das Minimalpolynom μ_A habe den Grad m. Ist $v \in K^n$, so sind die Vektoren $v, Av, \ldots, A^m v$ linear abhängig.*

Beweis: Es ist $A^m \in L(I, A, A^2, \ldots, A^{m-1})$, also $A^m = \sum_{i=0}^{m-1} c_i A^i$ mit $c_i \in K$. Dann ist $A^m v = \sum_{i=0}^{m-1} c_i A^i v$.

4.7.4. Lemma. *Seien $c_0, \ldots, c_{n-1} \in K$. Für die Matrix*

$$A = \begin{bmatrix} 0 & \cdots & \cdots & 0 & c_0 \\ 1 & \ddots & & \vdots & c_1 \\ 0 & 1 & \ddots & \vdots & c_2 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & c_{n-1} \end{bmatrix}$$

gilt:

$$\mu_A = \chi_A = T^n - \sum_{i=0}^{n-1} c_i T^i.$$

(Man nennt diese Matrix A die *Begleitmatrix* zum Polynom $T^n - \sum_{i=0}^{n-1} c_i T^i$).

Beweis: Man rechnet leicht nach $\chi_A = T^n - \sum_{i=0}^{n-1} c_i T^i$. Zu zeigen ist, dass das Minimalpolynom den Grad n hat. Es ist $Ae_i = e_{i+1}$, für $0 \leq i \leq n-1$, demnach ist $A^i e_1 = e_{i+1}$, ebenfalls für $0 \leq i \leq n-1$. Insbesondere sehen wir, dass die Vektoren $A^i e_1$ mit $0 \leq i \leq n-1$ linear unabhängig sind. Also hat das Minimalpolynom μ_A den Grad n. Da μ_A Teiler von χ_A ist, muss $\mu_A = \chi_A$ gelten.

Folgerung. *Ist ϕ ein normiertes Polynom in $K[T]$ vom Grad n, so gibt es $A \in M(n \times n, K)$ mit $\chi_A = \mu_A = \phi$.*