# §23. Die Jordan'sche Normalform

Wir suchen für einen trigonalisierbaren Endomorphismus unter seinen dreiecksförmigen Darstellungsmatrizen eine Darstellungsmatrix, die in einem gewissen Sinne eindeutig bestimmt ist. Dies ist die sog. Jordan'sche Normalform dieses Endomorphismus, die sich auch durch eine besonders einfache Bauart auszeichnet.

$V$ sei ein $K$–Vektorraum und $f \in \mathrm{End}_K(V)$. Ist dann $U \subseteq V$ ein $f$–invarianter Unterraum von $V$, so gilt $f(U) \subseteq U$, so daß $f$ durch Einschränkung auf $U$ eine $K$–lineare Abbildung $f|U : U \longrightarrow U \quad , \quad u \longmapsto f(u) \in U$ definiert.

**(23.1) SATZ:** $V$ sei ein $n$–dimensionaler $K$–Vektorraum und $f \in \mathrm{End}_K(V)$. $U_1, \ldots, U_s$ seien $f$–invariante Untervektorräume von $V$ mit $\dim_K(U_i) = n_i$ und Basen $B_i \subseteq U_i \quad (i = 1, \ldots, s)$. Ist dann $V$ die direkte Summe der Untervektorräume $U_i$, so ist $B := \bigcup_{i=1}^{s} B_i$ eine Basis von $V$, bzgl. der $f$ die Darstellungsmatrix

$$\mathrm{M}_B^B(f) = \begin{pmatrix} M_1 & & & O \\ & M_2 & & \\ & & \ddots & \\ O & & & M_s \end{pmatrix} \in \mathrm{M}_n(K)$$

hat. Dabei gilt:

**a)** $M_i = \mathrm{M}_{B_i}^{B_i}(f|U_i) \in \mathrm{M}_{n_i}(K) \qquad$ für alle $i = 1, \ldots, s$

**b)** $p_f = \prod_{i=1}^{s} p_{f|U_i}$

**BEM:** $\mathrm{M}_B^B(f)$ ist in dem obigen Satz eine sog. **Blockdiagonalmatrix**, die wir auch mit $\mathrm{Diag}(M_1, M_2, \ldots, M_s)$ bezeichnen werden.

**Strategie:** Zerlege $V$ in eine direkte Summe von $f$–invarianten Untervektorräumen $U_i$, die sich selbst nicht weiter zerlegen lassen. Dann hat $f|U_i$ bzgl. einer geeigneten Basis als Darstellungsmatrix eine besonders einfach gebaute Matrix, eine sog. Jordan-Matrix.

**(23.2) DEF:** $V$ sei ein $K$–Vektorraum und $f \in \mathrm{End}_K(V)$. $V$ heißt **unzerlegbar bzgl.** $\boldsymbol{f}$, wenn gilt:

**i)** $V \neq 0 \qquad$ **ii)** $V$ ist **nicht** direkte Summe zweier $f$–invarianter Unterräume $\neq 0$ von $V$.

**BEM:** ii) bedeutet: Ist $V = U_1 \oplus U_2$, wobei $U_1$ und $U_2$ $f$–invariante Unterräume sind, so folgt $U_1 = 0$ oder $U_2 = 0$.

**Beispiele:** **a)** Im Falle $\dim_K(V) = 1$ ist $V$ unzerlegbar bzgl. $f$.

**b)** Ein weiteres Beispiel für einen unzerlegbaren Vektorraum findet man in Aufg. 66.

**(23.3) SATZ:** $V$ sei ein $n$–dimensionaler $K$–Vektorraum und $f \in \mathrm{End}_K(V)$. Dann gibt es $f$–invariante Untervektorräume $U_1, \ldots, U_s$ von $V$, so daß gilt:

**a)** $V = \bigoplus_{i=1}^{s} U_i$

**b)** $U_i$ ist unzerlegbar bzgl. $f|U_i$ für alle $i = 1, \ldots, s$.

**(23.4) HILFSSATZ:** Sei $f \in \mathrm{End}_K(V)$. Dann gilt:

**a)** $O \subseteq \mathrm{Kern}(f) \subseteq \mathrm{Kern}(f^2) \subseteq \mathrm{Kern}(f^3) \subseteq \ldots \subseteq V$ (Kernfolge)

$V \supseteq \mathrm{Bild}(f) \supseteq \mathrm{Bild}(f^2) \supseteq \mathrm{Bild}(f^3) \supseteq \ldots \supseteq O$ (Bildfolge)

**b)** Ist $V$ endlichdimensional, so gibt es $r, s \in \mathbb{N}$ mit

$\mathrm{Kern}(f^k) = \mathrm{Kern}(f^r)$ $(\forall\, k \geq r)$ und $\mathrm{Bild}(f^l) = \mathrm{Bild}(f^s)$ $(\forall\, l \geq s)$

**(23.5) LEMMA:** $V$ sei ein endlichdimensionaler $K$–Vektorraum, und $f$ sei ein $K$–Endomorphismus von $V$. Es gelte $f^n = o$ für ein $n \geq 1$, und $W$ sei ein direktes Komplement von $\mathrm{Kern}(f^{n-1})$ in $V$ (d.h. $V = \mathrm{Kern}(f^{n-1}) \oplus W$). Dann gilt:

**a)** Die Einschränkungen

$$W \xrightarrow{f} f(W) \xrightarrow{f} f^2(W) \xrightarrow{f} \ldots \xrightarrow{f} f^{n-1}(W)$$

sind alle Isomorphismen.

**b)** Die Summe $S_W := W + f(W) + f^2(W) + \ldots + f^{n-1}(W)$ ist direkt, und $S_W$ ist ein $f$–invarianter Untervektorraum von $V$.

**c)** $S_W$ besitzt ein $f$–invariantes Komplement in $V$.

**(23.6) SATZ:** Sei $V$ ein $K$–Vektorraum und $f \in \mathrm{End}_K(V)$. Ist dann $V$ unzerlegbar bzgl. $f$ und gilt $f^n = o$ und $f^{n-1} \neq o$, so folgt:

**a)** $\dim_K(V) = n$

**b)** Es gibt eine Basis $B$ von $V$ mit

$$\mathrm{M}_B^B(f) = \begin{pmatrix} 0 & 1 & 0 & \ldots & \ldots & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & 0 & 1 & 0 \\ \vdots & & & & 0 & 1 \\ 0 & \ldots & \ldots & \ldots & \ldots & 0 \end{pmatrix} =: J_n \in \mathrm{M}_n(K)$$

**Bew:** **a)** Sei $W$ ein direktes Komplement von $\mathrm{Kern}(f^{n-1})$ in $V$. Wegen $f^{n-1} \neq o$ ist $\mathrm{Kern}(f^{n-1}) \neq V$, also $\underline{W \neq O}$. Nach (23.5b) ist

$$S_W \;=\; W \oplus f(W) \oplus f^2(W) \oplus \ldots \oplus f^{n-1}(W)$$

ein $f$–invarianter Unterraum von $V$, zu dem es nach (23.5c) ein $f$–invariantes Komplement $V'$ in $V$ gibt:

$$V \;=\; S_W \oplus V'\,.$$

Nach Voraussetzung ist $V$ unzerlegbar bzgl. $f$. Da $S_W \neq O$ wegen $W \neq O$ gilt, folgt also $V' = O$ und damit

$$(\star) \quad V \;=\; S_W \;=\; W \oplus f(W) \oplus f^2(W) \oplus \ldots \oplus f^{n-1}(W)$$

Es soll nun gezeigt werden, daß $\dim_K(W) = 1$ ist. Dazu führen wir einen indirekten Beweis:

**Annahme:** $\dim_K(W) \geq 2$

Dann ist $W$ die direkte Summe zweier Unterräume $W'$ und $W''$, die beide $\neq O$ sind. Mit (23.5a) ergibt sich

$$\begin{aligned} S_W &= (W' \oplus W'') \oplus (f(W') \oplus f(W'')) \oplus (f^2(W') \oplus f^2(W'')) \oplus \ldots \oplus (f^{n-1}(W') \oplus f^{n-1}(W'')) \\ &= [\,W' \oplus f(W') \oplus f^2(W') \oplus \ldots \oplus f^{n-1}(W')\,] \oplus [\,W'' \oplus f(W'') \oplus f^2(W'') \oplus \ldots \oplus f^{n-1}(W'')\,] \\ &= S_{W'} \oplus S_{W''} \end{aligned}$$

Also $V = S_{W'} \oplus S_{W''}$. Da beide Unterräume nach (23.5b) $f$–invariant sind, folgt $S_{W'} = O$ oder $S_{W''} = O$ und daraus $W' = O$ oder $W'' = O$. **Widerspruch!**. Also $\underline{\dim_K(W) = 1}$.

Wegen $W \cong f(W) \cong f^2(W) \cong \ldots \cong f^{n-1}(W)$ nach (23.5a) gilt dann auch $\dim_K(f^i(W)) = 1$ für alle $i = 0, 1, \ldots, n-1$, so daß sich aus $(\star)$ mit (13.15b) ergibt

$$\underline{\dim_K(V)} \;=\; \sum_{i=0}^{n-1} \dim_K(f^i(W)) \;=\; n \cdot 1 \;=\; \underline{n}$$

**b)** Nach Voraussetzung gibt es ein $v \in V$ mit $f^{n-1}(v) \neq o_V$. Dann ist nach Aufgabe 63b) die Menge $\{v, f(v), f^2(v), \ldots, f^{n-1}(v)\}$ linear unabhängig und damit sogar eine Basis von $V$, da $\dim_K(V) = n$. Setze:

$$B \;:=\; (\underbrace{f^{n-1}(v)}_{=:v_1}, \underbrace{f^{n-2}(v)}_{=:v_2}, \ldots, \underbrace{f(v)}_{=:v_{n-1}}, \underbrace{v}_{=:v_n})$$

Damit ergibt sich

$f(v_1) = f(f^{n-1}(v)) = f^n(v) = o_V$

$f(v_2) = f(f^{n-2}(v)) = f^{n-1}(v) = v_1$

$\vdots$

$f(v_n) = f(v) = v_{n-1}$.

Die Darstellungsmatrix von $f$ bzgl. dieser Basis ist also

$$\mathrm{M}_B^B(f) \;=\; \begin{pmatrix} 0 & 1 & 0 & \ldots & \ldots & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & 0 & 1 & 0 \\ \vdots & & & & 0 & 1 \\ 0 & \ldots & \ldots & \ldots & \ldots & 0 \end{pmatrix} \;=:\; J_n \in \mathrm{M}_n(K)$$

**(23.7) FOLG:** Sei $\dim_K(V) = n$ , $f \in \mathrm{End}_K(V)$ , $V$ sei unzerlegbar bzgl. $f$. Ist dann $f$ nilpotent, so gilt $f^n = o$ und $f^{n-1} \neq o$.

**Bew:** Sei $m := \min\{\, l \mid l \in \mathbb{N}, f^l = o \,\}$. Dann gilt $m \in \mathbb{N}, f^m = o$ und $f^{m-1} \neq o$. Mit (23.6a) folgt $m = \dim_K(V) = n$, womit die Behauptung bewiesen ist.

**(23.8) Fitting–Lemma**

$V$ sei ein endlichdimensionaler $K$–Vektorraum, und $f$ sei ein $K$–Endomorphismus von $V$. Ist dann $V$ unzerlegbar bzgl. $f$, so ist jeder Endomorphismus $g \in \mathrm{End}_K(V)$, für den $f \circ g = g \circ f$ gilt, entweder bijektiv oder nilpotent. Insbesondere ist $f$ selbst entweder bijektiv oder nilpotent.

**Bew:** Seien $B_l := \mathrm{Bild}(g^l)$ und $K_l := \mathrm{Kern}(g^l)$ für $l \in \mathbb{N}$. Um zu zeigen, daß dies $f$–invariante Unterräume von $V$ sind, wird die Voraussetzung $f \circ g = g \circ f$ benötigt. Induktiv folgt daraus $f \circ g^l = g^l \circ f$ für alle $l \in \mathbb{N}$.

$v \in B_l \implies v = g^l(w)\ (w \in V) \implies f(v) = f(g^l(w)) = g^l(f(w)) \in B_l$ , d.h. $f(B_l) \subseteq B_l$

$v \in K_l \implies g^l(v) = o_V \implies o_V = f(g^l(v)) = g^l(f(v)) \implies f(v) \in Keg^l$ , also $f(K_l) \subseteq K_l$.

Nach (23.4b) werden die Folgen $(B_l)$ und $(K_l)$ stationär, so daß es ein $n \in \mathbb{N}$ gibt mit $B_l = B_n$ und $K_l = K_n$ für alle $l \geq n$. Wir behaupten

$$(\star) \quad V \;=\; B_n \oplus K_n$$

Da $V$ unzerlegbar bzgl. $f$ ist und da $B_n$ und $K_n$ $f$–invariante Unterräume von $V$ sind, muß einer dieser beiden Unterräume O sein:

**1. Fall:** $B_n = O \implies \mathrm{Bild}(g^n) = O \implies g^n = o$, d.h. $g$ ist nilpotent.

**2. Fall:** $K_n = O \implies g^n$ ist injektiv $\implies g^n$ ist ein Isomorphismus $\implies g$ ist ein Isomorphismus.

Bleibt also noch $(\star)$ zu zeigen. Sei $v \in V$ beliebig. Wegen $B_n = B_{2n}$ existiert $w \in V$ mit $g^n(v) = g^{2n}(w) \implies$

$$v \;=\; \underbrace{g^n(w)}_{\in B_n} + \underbrace{(v - g^n(w))}_{\in K_n} \in B_n + K_n$$

denn es ist $g^n(v - g^n(w)) = g^n(v) - g^{2n}(w) = g^n(v) - g^n(v) = o_V$. Aus $V \subseteq B_n + K_n \subseteq V$ folgt $\underline{V = B_n + K_n}$.

$v \in B_n \cap K_n \implies v = g^n(w)$ mit $w \in V$ und $o_V = g^n(v) = g^{2n}(w) \implies w \in K_{2n} = K_n \implies o_V = g^n(w) = v \implies \underline{B_n \cap K_n = O}$.

Damit ist $(\star)$ bewiesen.

**(23.9) SATZ:** $V$ sei ein $n$–dimensionaler $K$–Vektorraum. Der $K$–Endomorphismus $f : V \longrightarrow V$ besitze einen Eigenwert $\lambda \in K$. Ist dann $V$ unzerlegbar bzgl. $f$, so gibt es eine Basis $B$ von $V$ mit

$$\mathrm{M}_B^B(f) \;=\; \begin{pmatrix} \lambda & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 \\ 0 & \dots & \dots & \lambda \end{pmatrix} \;=\; J_n(\lambda) \in \mathrm{M}_n(K)\ .$$

$J_n(\lambda)$ heißt $(n \times n)$–**Jordan-Matrix** (oder **Jordan–Block**) zum Eigenwert $\lambda$ von $f$.

**Bew:** Sei $g := f - \lambda\,\mathrm{id}_V$ . Dann ist $g \in \mathrm{End}_K(V)$ , und es gilt

$g \circ f \;=\; (f - \lambda\,\mathrm{id}_V) \circ f \;=\; f \circ f - \lambda f \;=\; f \circ (f - \lambda\,\mathrm{id}_V) \;=\; f \circ g$

Wegen $\mathrm{Kern}(g) = \mathrm{Kern}(f - \lambda\,\mathrm{id}_V) = \mathrm{Eig}(f, \lambda) \neq O$ ist $g$ **kein** Isomorphismus. Nach dem Fitting–Lemma (23.8) ist $g$ deshalb **nilpotent**. Da die $f$–invarianten Unterräume von $V$ mit den $g$–invarianten Unterräume übereinstimmen, ist $V$ auch unzerlegbar bzgl. $g$ . Nach (23.7) gilt dann $g^n = o$ und $g^{n-1} \neq o$, so daß nach (23.6) eine Basis $B$ von $V$ existiert mit

$$\mathrm{M}_B^B(g) \;=\; \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix} \;=\; J_n(0)$$

Folglich $\quad \mathrm{M}_B^B(f) \;=\; \mathrm{M}_B^B(g + \lambda\,\mathrm{id}_V) \;=\; \mathrm{M}_B^B(g) + \mathrm{M}_B^B(\lambda\,\mathrm{id}_V) \;=$

$$\begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix} + \begin{pmatrix} \lambda & & & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & & & \lambda \end{pmatrix} \;=\; \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix} \;=\; J_n(\lambda)$$

Mit Hilfe von (23.3) können wir jetzt den allgemeinen Fall behandeln:

**(23.10) SATZ:** **Jordan'sche Normalform (ca. 1870)**

$V$ sei ein $n$–dimensionaler $K$–Vektorraum. $f \in \mathrm{End}_K(V)$ sei trigonalisierbar. Dann gibt es eine Basis $B$ von $V$, Zahlen $n_1, n_2, \dots, n_s \in \mathbb{N}$ $(1 \le s \le n)$ und Elemente $\lambda_1, \lambda_2, \dots, \lambda_s \in K$, so daß gilt:

$$\mathrm{M}_B^B(f) \;=\; \begin{pmatrix} J_{n_1}(\lambda_1) & & & 0 \\ & J_{n_2}(\lambda_2) & & \\ & & \ddots & \\ 0 & & & J_{n_s}(\lambda_s) \end{pmatrix} \;\in\; \mathrm{M}_n(K)$$

Diese Matrix heißt die **Jordan'sche Normalform (JNF)** von $f$.

**Bew:** Nach (23.3) gibt es $f$–invariante Unterräume $U_i \subseteq V$ mit folgenden Eigenschaften:

$$V \;=\; \bigoplus_{i=1}^{s} U_i \quad , \quad U_i \text{ ist unzerlegbar bzgl. } f|U_i =: f_i$$

Aus (23.1b) folgt, daß $p_{f_i}$ ein Teiler von $p_f$ ist, also in Linearfaktoren zerfällt. $f_i$ besitzt also mindestens einen Eigenwert $\lambda_i \in K$ , der auch Eigenwert von $f$ ist .

Sei $\dim_K(U_i) =: n_i$ . Nach (23.9) gibt es eine Basis $B_i$ von $U_i$ mit

$$\mathrm{M}_{B_i}^{B_i}(f_i) \;=\; J_{n_i}(\lambda_i)$$

Nach (23.1) ist dann $B := \bigcup_{i=1}^{s} B_i$ eine Basis von $V$ , bzgl. der gilt

$$\mathrm{M}_B^B(f) \;=\; \begin{pmatrix} J_{n_1}(\lambda_1) & & & 0 \\ & J_{n_2}(\lambda_2) & & \\ & & \ddots & \\ 0 & & & J_{n_s}(\lambda_s) \end{pmatrix} \;\in\; \mathrm{M}_n(K)$$

**(23.11) BEM:** **a)** In der Hauptdiagonale der JNF von $f$ stehen die Eigenwerte von $f$. $\lambda_1, \ldots, \lambda_s$ müssen nicht paarweise verschieden sein. Es ist $s \geq$ Anzahl der paarweise verschiedenen Eigenwerte von $f$.

**b)** Die JNF ist eindeutig bestimmt bis auf die Reihenfolge der Jordan–Blöcke.

**c)** Eine Diagonalmatrix ist eine JNF, bei der alle Jordan–Blöcke das Format $(1 \times 1)$ haben.

Jeder Endomorphismus (und jede Matrix) über $\mathbb{C}$ sind trigonalsierbar, daher:

**(23.12) FOLG:** **a)** Jeder $\mathbb{C}$ Endomorphismus $f : V \longrightarrow V$ $(\dim_{\mathbb{C}}(V) < \infty)$ besitzt eine JNF als Darstellungsmatrix bzgl. einer geeigneten Basis.

**b)** Jede Matrix $M \in \mathrm{M}_n(\mathbb{C})$ ist ähnlich zu einer JNF.

Es lassen sich Ergebnisse über Matrizen mit Hilfe der Jordan-Normalform beweisen. Dies wollen wir an einigen Beispielen zeigen. Eine trigonalisierbare Matrix $M$ ist ähnlich zu einer Blockdiagonalmatrix, in deren Hauptdiagonale Jordan–Matrizen stehen. Für das Rechnen mit Blockdiagonalmatrizen gelten die folgenden Rechenregeln, die sich leicht verifizieren lassen:

**(23.13) LEMMA:** Seien $M = \mathrm{Diag}(M_1, \ldots, M_s)$ , $N = \mathrm{Diag}(N_1, \ldots, N_s) \in \mathrm{M}_n(K)$ Blockdiagonalmatrizen mit $M_i, N_i \in \mathrm{M}_{n_i}(K)$ $(\forall\, i = 1, 2, \ldots, s)$ . Dann gilt:

**a)** $M \overset{+}{\cdot} N = \mathrm{Diag}(M_1 \overset{+}{\cdot} N_1, \ldots, M_s \overset{+}{\cdot} N_s)$

**b)** ${}^tM = \mathrm{Diag}({}^tM_1, \ldots, {}^tM_s)$

**c)** $M_i \in \mathrm{GL}_{n_i}(K)\ \ \forall\, i = 1, \ldots, s \quad \Longrightarrow \quad M \in \mathrm{GL}_n(K)$ und $M^{-1} = \mathrm{Diag}(M_1^{-1}, \ldots, M_s^{-1})$

**d)** $M_i \approx N_i \quad (\forall\, 1, \ldots, s) \quad \Longrightarrow \quad M \approx N$

**e)** $\det(M) = \det(M_1) \cdot \ldots \cdot \det(M_s)$

**f)** $p_M \;=\; p_{M_1} \cdot \ldots \cdot p_{M_s}$

**g)** Für jedes Polynom $p \in K[T]$ gilt $p(M) = \mathrm{Diag}(p(M_1), \ldots, p(M_s))$ .

**(23.14) SATZ:** Für eine Matrix $M \in \mathrm{M}_n(\mathbb{C})$ gilt $M \approx {}^tM$.

**Bew:** Nach (23.12b) ist $M$ ähnlich zu einer JNF, d.h.

$$M \approx \mathrm{Diag}(J_{n_1}(\lambda_1), \ldots, J_{n_s}(\lambda_s)) =: J$$

Ist $P := \begin{pmatrix} 0 & & 1 \\ & / & \\ 1 & & 0 \end{pmatrix} \in \mathrm{M}_r(K)$, so gilt $P \cdot P = E_r$, also $P = P^{-1}$. Man rechnet sofort nach: $P \cdot J_r(\lambda) \cdot P = {}^tJ_r(\lambda)$, d.h. $J_r(\lambda) \approx {}^tJ_r(\lambda)$. Mit (23.13d) folgt

$$M \approx J = \mathrm{Diag}(J_{n_1}(\lambda_1), \ldots, J_{n_s}(\lambda_s)) \approx \mathrm{Diag}({}^tJ_{n_1}(\lambda_1), \ldots, {}^tJ_{n_s}(\lambda_s)) = {}^tJ$$

Aus $M \approx J$ folgt sofort ${}^tM \approx {}^tJ$, so daß sich mit dem obigen Ergebnis die Behauptung $M \approx {}^tM$ ergibt.

**(23.15) SATZ: Satz von Cayley–Hamilton (für Matrizen)**

Sei $M \in \mathrm{M}_n(K)$ und $p_M \in K[T]$ das charakteristische Polynom von $M$. Dann gilt

$$p_M(M) = O \in \mathrm{M}_n(K) \quad \text{(Nullmatrix).}$$

**Bew: 1. Fall:** $M$ ist eine Jordan–Matrix, d.h.

$$M = J = J_n(\lambda) = \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix} = \begin{pmatrix} \lambda & & & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & & & \lambda \end{pmatrix} + \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix} = \lambda E_n + N,$$

wobei $N$ als strikt obere Dreiecksmatrix nilpotent ist und $N^n = O$ gilt. Das charakteristische Polynom von $J$ ist

$$p_J \;=\; (\lambda - T)^n .$$

Hieraus folgt

$$p_J(J) \;=\; (\lambda E_n - J)^n \;=\; (-N)^n \;=\; (-1)^n N^n \;=\; O .$$

**2. Fall:** $M$ ist trigonalisierbar.

Dann gibt es Jordan–Matrizen $J_i = J_{n_i}(\lambda_i)$ mit der Eigenschaft

$$M \;\approx\; \mathrm{Diag}(J_1, \ldots, J_s) =: J .$$

Nach (23.13f) gilt

$$p_J \;=\; p_{J_1} \cdot \ldots \cdot p_{J_s} ,$$

so daß für alle $i = 1, \ldots, s$ unter Benutzung des Ergebnisses aus dem ersten Fall

$$p_J(J_i) \;=\; p_{J_1}(J_i) \cdot \ldots \cdot \underbrace{p_{J_i}(J_i)}_{=O} \cdot \ldots \cdot p_{J_s}(J_i) \;=\; O$$

folgt. Mit (23.13g) erhält man

$$p_J(J) \;=\; \mathrm{Diag}(p_J(J_1), \ldots, p_J(J_s)) \;=\; \mathrm{Diag}(O, \ldots, O) \;=\; O .$$

Es läßt sich leicht zeigen: Für Matrizen $A, B \in \mathrm{M}_n(K)$ und ein Polynom $p \in K[T]$ gilt: $A \approx B \implies p(A) \approx p(B)$. Wendet man dies auf $M \approx J$ an und berücksichtigt, daß ähnliche Matrizen dasselbe charakteristische Polynom haben, so ergibt sich

$$p_M(M) \approx p_M(J) = p_J(J) = O$$

Eine Matrix, die ähnlich zur Nullmatrix ist, ist selbst die Nullmatrix. Folglich

$$p_M(M) \;=\; O\,.$$

**3. Fall:** Sei jetzt $M \in \mathrm{M}_n(K)$ beliebig.

In der Algebra–Vorlesung wird gezeigt, daß es einen $K$ umfassenden Körper $L$ gibt, über dem jedes Polynom in Linearfaktoren zerfällt. $L$ heißt algebraischer Abschluß von $K$ (wir kennen $\mathbb{R} \subseteq \mathbb{C}$ als konkretes Beispiel für diese Aussage). Dann gilt $M \in \mathrm{M}_n(L)$ und $p_M \in L[T]$ zerfällt in Linearfaktoren, so daß $M$ als Matrix über dem Körper $L$ trigonalisierbar ist. Nach dem zweiten Fall gilt

$$p_M(M) \;=\; O\,.$$

Damit ist der Beweis beendet.

Sei nun $M \in \mathrm{M}_n(K)$ eine beliebige Matrix. Dann gibt es Polynome $p \in K[T]$ mit $p(M) = O$ (z.B. $p = p_M$). Die Menge

$$I_M \;:=\; \{\, p \mid p \in K[T]\,,\, p(M) = O \,\} \subseteq K[T]$$

ist ein Ideal in dem HIB $K[T]$ (nachprüfen!) und ist daher ein Hauptideal. Nach Aufgabe 58 wird $I_M$ von einem Polynom $q$ kleinsten Grades in $I_M$ erzeugt. Normiert man dieses Polynom, so erhält man ein eindeutig bestimmtes erzeugendes Element $q_M$ von $I_M$. Zusammenfassend erhalten wir das folgende Ergebnis:

**(23.16) SATZ:** Sei $M \in \mathrm{M}_n(K)$. Dann gilt:

**a)** Es existiert ein eindeutig bestimmtes normiertes Polynom $q_M \in K[T]$ kleinsten Grades mit $q_M(M) = O$. $q_M$ heißt das **Minimalpolynom von** $M$.

**b)** Für alle $p \in K[T]$ mit $p(M) = O$ folgt $q_M | p$. Insbesondere $q_M | p_M$.

**c)** $1 \leq \mathrm{grad}(q_M) \leq \mathrm{grad}(p_M) = n$.

**Bew:** **a)** s.o.

**b)** Hier bezeichnet $|$ die Teilbarkeitsbeziehung zwischen Polynomen aus $K[T]$:
$p \,|\, q \;:\Longleftrightarrow\; \exists\, r \in K[T] \;:\; q = p \cdot r\,.$

$p(M) = O \implies p \in I_M = (q_M) = K[T] \cdot q_M$, d.h. es existiert ein $r \in K[T]$ mit $p = r \cdot q_M$, also $q_M | p$.

**c)** Klar.

**(23.17) FOLG:** Seien $M \in \mathrm{M}_n(K)$ und $\lambda \in K$. Dann sind folgende Aussagen äquivalent:

**a)** $\lambda$ ist ein Eigenwert von $M$ (d.h. Nullstelle von $p_M$)

**b)** $\lambda$ ist Nullstelle des Minimalpolynoms $q_M$ von $M$.

**Bew:** ***a*) $\Longrightarrow$ *b*)**

Nach Aufgabe 54 ist $q_M(\lambda)$ Eigenwert der Matrix $q_M(M) = O$. Also existiert ein Vektor $o_n \neq v \in K^n$ mit

$$o_n \;=\; \underbrace{q_M(M)}_{=O} \cdot v \;=\; q_M(\lambda) \cdot v \;\overset{v \neq o_n}{\Longrightarrow}\; q_M(\lambda) = 0\,.$$

***b*) $\Longrightarrow$ *a*)**

Gelte $q_M(\lambda) = 0$. Nach (23.16b) existiert ein Polynom $r \in K[T]$ mit der Eigenschaft $p_M = q_M \cdot r$. Es folgt

$$p_M(\lambda) \;=\; \underbrace{q_M(\lambda)}_{=0} \cdot r(\lambda) \;=\; 0$$

**(23.18) FOLG:** Sei $M \in \mathrm{M}_n(K)$.

**a)** $p_M$ und $q_M$ haben dieselben Nullstellen, allerdings mit i.a. unterschiedlichen Vielfachheiten.

**b)** Ist $M$ trigonalisierbar und ist $p_M = (-1)^n \prod_{i=1}^{r}(T-\lambda_i)^{s_i}$ das charakteristische Polynom von $M$ (die $\lambda_i$'s seien paarweise voneinander verschieden), so ist

$$q_M \;=\; (-1)^n \prod_{i=1}^{r}(T-\lambda_i)^{t_i} \quad \text{mit } 1 \leq t_i \leq s_i \;\; (\forall\, i = 1, \ldots, r) \text{ das Minimalpolynom von } M\,.$$

**Beispiel:** Sei $M = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -2 & 4 \end{pmatrix} \in \mathrm{M}_4(\mathbb{R})$

Das charakteristische Polynom berechnet sich zu $p_M = (T-3)(T-2)^3$. Für das Minimalpolynom ergibt sich $q_M = (T-3)(T-2)^k$ mit $k \in \{1, 2, 3\}$.

<u>$k = 1$:</u> $q = (T-3)(T-2) \;\Longrightarrow\; q(M) = (M - 3E_4)(M - 2E_4) = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \neq O$.

<u>$k = 2$:</u> $q = (T-3)(T-2)^2 \;\Longrightarrow\; q(M) = (M - 3E_4)(M - 2E_4)^2 = O$. Also

$$\underline{q_M \;=\; (T-3)(T-2)^2}$$

Dies ist das normierte Polynom kleinsten Grades, das von $M$ annulliert wird.

**(23.19) SATZ:** Ist $M \in \mathrm{M}_n(K)$ eine <u>trigonalisierbare</u> Matrix, so sind folgende Aussagen äquivalent:

**a)** $M$ ist diagonalisierbar

**b)** Das Minimalpolynom von $M$ besitzt nur einfache Nullstellen.

**Historische Bemerkung:** Die Jordan'sche Normalform einer Matrix wurde um 1870 von dem französischen Mathematiker **Camille Jordan** (1838–1922) im Zusammenhang mit der Untersuchung von Systemen linearer Differentialgleichungen mit konstanten Koeffizienten gefunden.