8.2 Normalformen

Die im letzten Abschnitt eingeführten Begleitmatrizen und die am Ende des letzten Kapitels diskutierte Smith-Normalform einer Matrix bilden in diesem Abschnitt die zentralen Bausteine, um auch für nicht diagonalisierbare Matrizen über einem Körper zu einer Normalform zu gelangen. Konkret werden wir zu einem Endomorphismus erst eine sogenannte charakteristische Matrix definieren, für diese dann Elementarteiler ausrechnen und schließlich die Begleitmatrizen der Elementarteiler durch geschickte Wahl einer Basis als Diagonalblöcke in der Frobenius-Normalform wiedersehen. Diese Überlegungen führen uns dann auch zum Satz von Cayley-Hamilton, der uns im vorigen Abschnitt für einen Beweis fehlte.

Wir betrachten in diesem Abschnitt einen K-Vektorraum V der Dimension n mit Basis $\mathcal{C} = (c_1, \ldots, c_n)$ sowie einen K-Vektorraum-Endomorphismus $F \in End_K(V)$ mit $M_{\mathcal{C}}^{\mathcal{C}}(F) = A \in K^{n \times n}$.
We bereits gesehen, ist V dann auch ein $K[t]$-Modul mittels der Skalarmultiplikation

$$\begin{array}{rcl} \cdot_F : K[t] \times V & \longrightarrow & V \\ (p, v) & \longmapsto & p \cdot_F v = p(F)(v) \end{array}$$

Lemma 8.2.1 *Die Abbildung*

$$\begin{array}{rcl} \psi_F : K[t]^{n \times 1} & \longrightarrow & V \\ \underline{p} = \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix} & \longmapsto & \psi_F(\underline{p}) := \sum_{i=1}^n p_i(F)(c_i) \end{array}$$

ist ein $K[t]$-Modul-Epimorphismus.

Beweis: Die $K[t]$-Modul-Homomorphismuseigenschaft, d.h. die Verträglichkeit mit Addition und Skalarmultiplikation, lässt sich ohne Schwierigkeiten direkt nachrechnen. Einzig die Surjektivität benötigt etwas mehr Aufmerksamkeit. Betrachtet man jedoch die konstanten Polynome $0_{K[t]}$ und $1_{K[t]}$ unter der Skalarmultiplikation, so gilt

$$0_{K[t]} \cdot_F v = 0_{End_K(V)}(v) = 0_V \quad \forall v \in V \text{ und}$$

$$1_{K[t]} \cdot_F v = Id_{End_K(V)}(v) = v \quad \forall v \in V,$$

weswegen die Basiselemente aus $\mathcal{C}$ in $\mathrm{Im}(\psi_F) \subseteq V$ liegen:

$$c_i = \psi_F\left(\begin{pmatrix} 0_{K[t]} & \\ \vdots & \\ 0_{K[t]} & \\ 1_{K[t]} & \text{Pos. i} \\ 0_{K[t]} & \\ \vdots & \\ 0_{K[t]} & \end{pmatrix}\right).$$

Damit liegt auch das Erzeugnis von $\mathcal{C}$, also der ganze K-Vektorraum V, in $\mathrm{Im}(\psi_F)$, was zu zeigen war.

$\square$

Bemerkung 8.2.2 *Ist $V = K^{n\times 1}$, $A \in K^{n\times n}$, und $\mathcal{C} = (e_1, \ldots, e_n)$ die kanonische Basis von V, so ist $F = F_A$ und ψ_{F_A} bildet gerade die kanonische Basis $(e_1^*, \ldots, e_n^*)$ aus $K[t]$-Einheitsvektoren auf $\mathcal{C}$ ab.*

Wir betrachten nun einen weiteren $K[t]$-Modul-Homomorphismus, wobei wir die Matrix $tE_n - A \in K[t]^{n\times n}$ mit H abkürzen:

$$\begin{aligned} \Phi_H : K[t]^{n\times 1} &\longrightarrow K[t]^{n\times 1} \\ \underline{p} &\longmapsto \Phi_H(\underline{p}) := H\underline{p} \end{aligned}$$

Dass es sich um einen $K[t]$-Modul-Homomorphismus handelt, lässt sich direkt nachrechnen, was hier aus Platzgründen nicht ausgeführt wird. Hier unterstreichen wir die Vektoren in $K^{n\times 1}$ bzw. $K[t]^{n\times 1}$ ausnahmsweise, damit wir uns im ganzen folgenden Beweis stets daran erinnern, dass es sich hierbei um Vektoren handelt.

Lemma 8.2.3 *Φ_H ist injektiv und* $\mathrm{Im}(\Phi_H) = \ker(\psi_F)$.

Beweis: Schritt 1: Injektivität von Φ_H
Jedes $\underline{q} \in K[t]^{n\times 1}$ lässt sich schreiben als

$$\underline{q} = \sum_{i=0}^{d} \underline{v_i} t^i$$

für $d = \max\{\deg(q_i) \mid 1 \leq i \leq n\}$ und für geeignete $\underline{v_0}, \ldots, \underline{v_d} \in K^{n\times 1}$. Ist $\underline{q} \in \ker(\Phi_H)$, so gilt

$$(tE_n - A)\sum_{i=0}^{d} \underline{v_i} t^i = \underline{0}.$$

Betrachtet man diese Gleichung gradweise, so ergibt sich im Grad $i+1$ in t:

$$\underline{v_i} t^{i+1} = (A\underline{v_{i+1}})t^{i+1}$$

für jedes $-1 \leq i \leq d$, wobei natürlich die fehlenden Vektoren als $\underline{v_{-1}} = \underline{0} = \underline{v_{d+1}}$ ergänzt sind. Insbesondere ist also $\underline{v_d} = A\underline{0} = \underline{0}$, weswegen mit demselben Argument $\underline{v_{d-1}} = \underline{0}$ gelten muss und dann durch Iteration schließlich $\underline{v_i} = \underline{0}$ für alle $0 \leq i \leq d$, womit die erste Behauptung $\ker(\Phi_H) = \{\underline{0}\}$ gezeigt ist.

Schritt 2: $\mathrm{Im}(\Phi_H) \subseteq \ker(\psi_F)$
Das Bild von Φ_H wird erzeugt von den Spalten der Matrix $(tE_n - A)$. Für die i-te Spalte $(1 \leq i \leq n)$ gilt

$$\psi_F\left(\begin{pmatrix} -a_{1i} \\ \vdots \\ -a_{(i-1)i} \\ t - a_{ii} \\ -a_{(i+1)i} \\ \vdots \\ -a_{ni} \end{pmatrix}\right) = F(c_i) - \sum_{j=1}^{n} a_{ji}c_j = \left(\sum_{j=1}^{n} a_{ji}c_j\right) - \left(\sum_{j=1}^{n} a_{ji}c_j\right) = 0$$

Schritt 3: $\mathrm{Im}(\Phi_H) \supseteq \ker(\psi_F)$
Sei nun $\underline{p} \in \ker(\psi_F)$. Dann gibt es ein $d \in \mathbb{N}_0$ und $\underline{w_0}, \ldots, \underline{w_d} \in K^{n\times 1}$ mit

$$\underline{p} = \sum_{i=0}^{d} \underline{w_i} t^i.$$

Schreiben wir nun

$$\underline{w_d} t^d = (tE_n - A)t^{d-1}\underline{w_d} + At^{d-1}\underline{w_d},$$

so erhalten wir

$$\underline{p} = (tE_n - A)t^{d-1}\underline{w_d} + (\underline{w_{d-1}} + A\underline{w_d})t^{d-1} + \sum_{i=0}^{d-2} \underline{w_i} t^i.$$

Iterieren wir dieses Vorgehen, indem wir nun $(\underline{w_{d-1}} + A\underline{w_d})$ umschreiben, dann den neuen Faktor vor t^{d-2} und so weiter bis wir schließlich auch den Faktor von t^1 umgeschrieben haben, so erreichen wir die Form

$$\underline{p} = (tE_n - A)\left(\sum_{i=0}^{d-1} \underline{w_i}' t^i\right) + \underline{w_0}',$$

wobei $\underline{w_0}', \ldots, \underline{w_d}' \in K^{n\times 1}$ die in diesem Prozess entstandenen, neuen Vektoren sind. Nach Schritt 2 wissen wir jedoch, dass sowohl $\underline{p} \in \ker(\psi_F)$ als auch $(tE_n - A)\sum_{i=0}^{d-1} \underline{w_i}' t^i \in \ker(\psi_F)$, womit

$$\underbrace{\underline{p}}_{\in\ker(\psi_F)} - \underbrace{(tE_n - A)\sum_{i=0}^{d-1} \underline{w_i}' t^i}_{\in\ker(\psi_F)} = \underline{w_0}' \in \ker(\psi_F).$$

Da darüber hinaus $\underline{w_0}' \in K^{n\times 1}$, gilt $\underline{w_0}' \in \ker(\psi_F) \cap K^{n\times 1}$. Wegen der linearen Unabhängigkeit der Elemente von $\mathcal{C}$ ist aber $\ker(\psi_F) \cap K^{n\times 1} = \{0\}$. Damit ist $\underline{w_0}' = 0$, womit die gewünschte Inklusion gezeigt und folglich auch die Gleichheit der Mengen bewiesen ist.

$\square$

Eine Situation wie in den vorigen Lemmata, wo zwei Abbildungen verkettet werden, von denen die erste injektiv und die zweite surjektiv ist sowie das Bild der ersten gleich dem Kern der zweiten ist, ist eine gut untersuchte und für die Praxis sehr relevante Standardsituation. Daher existieren dafür verkürzte Schreibweisen:

Notation 8.2.4 *Seien $\varphi : U \longrightarrow V$ und $\psi : V \longrightarrow W$ zwei Homomorphismen von Moduln über einem kommutativen Ring R, so sagt man, dass sie einen* **Komplex**

$$U \xrightarrow{\varphi} V \xrightarrow{\psi} W$$

bilden, falls $\mathrm{Im}(\varphi) \subseteq \ker(\psi)$. *Der Komplex heißt* **exakt**, *falls* $\mathrm{Im}(\varphi) = \ker(\psi)$. *Wir sprechen von einer* **kurzen exakten Sequenz**

$$0 \longrightarrow U \xrightarrow{\varphi} V \xrightarrow{\psi} W \longrightarrow 0,$$

falls die Verkettungen der Abbildungen an allen drei Positionen exakt sind. In diesem Fall ist φ injektiv, ψ surjektiv und $\mathrm{Im}(\varphi) = \ker(\psi)$.

Damit lässt sich die Aussage der obigen Lemmata zusammenfassen als:

Satz 8.2.5 *Sei V ein K-Vektorraum mit Basis $\mathcal{C} = (c_1, \dots, c_n)$ und sei $F \in End_K(V)$. Dann existiert eine kurze exakte Sequenz von $K[t]$-Moduln*

$$\begin{array}{ccccccccc} 0 & \longrightarrow & K[t]^{n\times 1} & \xrightarrow{\Phi_H} & K[t]^{n\times 1} & \xrightarrow{\psi_F} & V & \longrightarrow & 0 \\ & & \underline{p} & \longmapsto & H\underline{p} & & & & \\ & & & & \underline{q} & \longmapsto & \sum_{i=1}^n q_i(F)(c_i) & & \end{array}$$

Bemerkung 8.2.6 *Ist in der Notation der vorigen Lemmata und des Satzes $v \in V$, dann existiert wegen der Exaktheit der Sequenz in der rechten Position, d.h. wegen der Surjektivität von ψ_F, ein $\underline{p} \in K[t]^{n\times 1}$ mit $\psi_F(\underline{p}) = v$. Direkte Rechnung zeigt, dass dann auch gilt:*

$$\psi_F(\langle \underline{p} \rangle_{K[t]}) = K[t] \cdot_F v = U(F, v),$$

denn $\psi_F(t^i \underline{p}) = t^i \cdot_F v = F^i(v)$.

Die bisherigen Überlegungen haben uns lediglich geholfen, mehr über den Endomorphismus F zu erfahren und über seine Matrix bzgl. gegebener Basis $\mathcal{C}$ von V. Was geschieht nun, wenn wir das Ganze bzgl. einer anderen Basis betrachten bzw. die Basis wechseln möchten?

Erinnerung 8.2.7 *Zwei quadratische Matrizen $A, B \in K^{n\times n}$ heißen ähnlich, falls es eine invertierbare Matrix $P \in Gl(n, K)$ gibt mit*

$$B = P^{-1}AP.$$

Sei V ein n-dimensionaler K-Vektorraum. Matrizen die denselben Endomorphismus von V bzgl. unterschiedlicher Basen darstellen sind ähnlich, die invertierbare Matrix P aus der obigen Formel ist dabei die genau die Basiswechselmatrix. Stellt umgekehrt eine Matrix A einen Endomorphismus von V bzgl. einer gegebenen Basis dar und ist eine Matrix B ähnlich zu A, so gibt es eine Basis von V, bzgl. derer B darstellende Matrix des Endomorphismus ist. Das Problem der Diagonalisierung war das Problem der Existenz und des Auffindens einer Basis bzgl. derer die darstellende Matrix Diagonalgestalt hat.

Wie schon bei der Diagonalisierung ist auch hier der zentrale Punkt der Betrachtung die Wahl der Basis von V. Ehe wir jedoch dazu kommen können, müssen wir zuerst nocheinmal die gerade eingeführte exakte Sequenz bzgl. Basiswechsel des freien $K[t]$-Moduls $K[t]^{n\times 1}$ untersuchen. Diese haben direkte Auswirkungen auf die Matrix der Abbildung Φ_H, die wir zu diesem Zweck explizit angeben. Zu beachten ist dabei, dass bei den folgenden Sequenzen die $K[t]$-Modulstruktur auf V nur schwer sichtbar zu machen ist, weswegen die Isomorphien

$$V \cong \operatorname{coker}(tE_n - A)) = K[t]^{n\times 1}/\operatorname{Im}(tE_n - A)$$

bzw.

$$V \cong \operatorname{coker}(P(tE_n - A)Q) = K[t]^{n\times 1}/\operatorname{Im}(P(tE_n - A)Q)$$

mit $P, Q \in Gl(n, K[t])$ für den letzten Modul verwendet werden.[1] Die Wahl einer geeigneten Basis von V wird sich dann in einem weiteren Schritt aus der Wahl einer Basis von $K[t]^{n\times 1}$ gewinnen lassen, bzgl. derer die Matrix von Φ_H besonders einfache Gestalt hat.

Es gilt das folgende kommutative Diagramm für $P, Q \in Gl(n, K[t])$:

$$\begin{array}{ccccccccc}
0 & \longrightarrow & K[t]^{n\times 1} & \xrightarrow{tE_n - A} & K[t]^{n\times 1} & \xrightarrow{E_n} & K[t]^{n\times 1}/\operatorname{Im}(tE_n - A) & \longrightarrow & 0 \\
 & & Q \uparrow\cong & & \cong\downarrow P & & \downarrow\cong & & \\
0 & \longrightarrow & K[t]^{n\times 1} & \xrightarrow{P(tE_n - A)Q} & K[t]^{n\times 1} & \xrightarrow{E_n} & K[t]^{n\times 1}/\operatorname{Im}(P(tE_n - A)Q) & \longrightarrow & 0
\end{array}$$

Durch geeignete Wahl von P und Q ist es also wie in Satz 7.3.8 möglich, eine Matrix $P(tE_n - A)Q$ in Smith-Normal-Form zu erreichen. Hierbei handelt es sich um eine Diagonalmatrix, deren Diagonalelemente wegen der Injektivität Φ_H alle von Null verschieden sind. Dabei erfüllen die Elementarteiler natürlich die Teilbarkeitsbedingungen $s_i \mid s_{i+1}$ für jedes $1 \leq i < n$. Ausserdem ist das charakteristische Polynom von F gerade das Produkt dieser Elementarteiler.
Insbesondere finden wir normierte Polynome $g_1, \ldots, g_k \in K[t]$ mit $\deg(g_i) > 0$, so dass die Elementarteiler von folgender Gestalt sind mit geeigneten

[1] Für den Fall, dass der Kokern noch nicht früher definiert war, fügen wir hier nochmals die Definition ein. Ist $\phi : M \longrightarrow N$ ein Homomorphismus von Moduln über einem Ring R, so bezeichnet man mit $\operatorname{coker}(\phi)$ den Faktormodul $N/\operatorname{Im}(\phi)$ und spricht vom Kokern der Abbildung ϕ.

$\varepsilon_1, \ldots, \varepsilon_n \in K^*$:

$$\begin{array}{lcll} s_i & := & \varepsilon_i & \text{für } 1 \leq i \leq n-k \\ s_i & := & \varepsilon_i g_{i+k-n} & \text{für } n-k+1 \leq i \leq n \end{array}.$$

Nennen wir wie in Satz 7.3.8 bzw. Korollar 7.3.12 die Basen $\mathcal{B} = (v_1, \ldots, v_n)$ und $\mathcal{C} = (w_1, \ldots, w_n)$, so ist wegen der speziellen Struktur der Elementarteiler $\mathcal{D} = (w_1, \ldots, w_{n-k}, g_1 w_{n-k+1}, \ldots, g_k w_n)$ eine Basis von $\operatorname{Im}(\Phi_H)$ (Denken Sie daran, dass ε_i eine Einheit ist für jedes $1 \leq i \leq n$; dann ist mit Korollar 7.3.12 klar, warum die Basis des Bildes so aussieht). Aus $\mathcal{D}$ können wir leicht ein $K[t]$-Erzeugendensystem von $K[t]^{n\times 1}/\operatorname{Im}(\Phi_H)$ gewinnen, nämlich $(w_{n-k+1}, \ldots, w_n)$ (Nachrechnen!). Damit ist ein K-Vektorraum-Erzeugendensystem von $K[t]^{n\times 1}/\operatorname{Im}(\Phi_H)$ gegeben durch

$$(w_{n-k+1}, tw_{n-k+1}, \ldots, t^{\deg(g_1)-1} w_{n-k+1}, w_{n-k+2}, \ldots, t^{\deg(g_k)-1} w_n).$$

Bei näherer Betrachtung sehen wir unter Zuhilfenahme von Bemerkung 8.2.6, dass $\psi_F(t^\ell w_{n-k+i}) = F^\ell(\psi_F(w_{n-k+i}))$ bzw. $\psi_F(\langle w_{n-k+i}\rangle_{K[t]}) = U(F, \psi_F(w_{n-k+i}))$. Unsere Konstruktion hat also eine Zerlegung des K-Vektorraums V in zyklische Unterräume geliefert:

$$V = \bigoplus_{i=1}^{k} U(F, a_i),$$

wobei $a_i = \psi_F(w_{n-k+i})$ für alle $1 \leq i \leq k$. Mit den vorigen Überlegungen wurde insbesondere eine Basis $\mathcal{D}'$ von V konstruiert, die sich aus den ebenfalls konstruierten Basen $\mathcal{D}_i$ der F-zyklischen Unterräume $U(F, a_i)$ zusammensetzt. Nach Satz 8.1.8 gilt

$$M_{\mathcal{D}_i}^{\mathcal{D}_i}(F\mid_{U(F,a_i)}) = B(g_i).$$

Setzen wir dies zusammen so erhalten wir die folgende Blockdiagonalmatrix

$$M_{\mathcal{D}'}^{\mathcal{D}'}(F) = \begin{pmatrix} B(g_1) & 0 & \ldots & 0 \\ 0 & B(g_2) & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & B(g_k) \end{pmatrix}.$$

Definition 8.2.8 *Die soeben konstruierte Bockdiagonalmatrix heißt die* **Frobenius-Normalform** *des Endomorphismus F.*

Bemerkung 8.2.9 *In den obigen Ausführungen haben wir nebenbei gezeigt, dass das Minimalpolynom g_F von F gerade g_k ist, also bis auf Assoziiertheit der größte Elementarteiler von $tE_n - A$: Wir haben gesehen, dass $V \cong \bigoplus_{i=1}^k K[t]/\langle g_i \rangle_{K[t]}$. Damit annuliert die (Skalar-)Multiplikation mit g_k alle Summanden (Denken Sie an die Teilbarkeitsbedingungen der Elementarteiler!). Andererseits annulliert kein echter Faktor von g_k den Summanden $K[t]/\langle g_k \rangle_{K[t]}$ in der obigen Summe, so dass g_k in der Tat der Erzeuger des Annulator-Ideals von $\bigoplus_{i=1}^k K[t]/\langle g_i \rangle_{K[t]}$ ist. Da nun Ψ_F (Der Homomorphismus vom Beginn von Abschnitt 8.1, nicht ψ_F aus diesem Abschnitt – Groß-Klein-Schreibung beachten!) gerade das Einsetzen von F für t ausführt, ist damit $\Psi_F(g_k) = 0_{End_K(V)}$ und g_k der Erzeuger des Hauptideals* $\ker(\Psi_F)$, *also das Minimalpolynom von F.*

Bemerkung 8.2.10 *Besteht die Frobenius-Normalform von F aus nur einem Block, der dann nach Konstruktion Begleitmatrix zu $g_1 \in K[t]$ ist, so gilt $g_{B(g_1)} = g_1$ ist das Minimalpolynom zu F.*

Satz 8.2.11 *(Cayley-Hamilton) Sei V ein n-dimensionaler K-Vektorraum mit Basis $\mathcal{C} = (c_1, \ldots, c_n)$und sei $F \in End_K(V)$. Sei $\chi_F \in K[t]$ das charakteristische Polynom von F. Dann gilt:*

$$\chi_F(F) = 0_{End_K(V)}.$$

Beweis: Wir bleiben weiter in der Notation der obigen Ausführungen. Nach Wahl der Basis $\mathcal{D}'$ für V ergibt sich für das charakteristische Polynom von F (bis auf Assoziiertheit) gerade:

$$\chi_F = \prod_{i=1}^k g_i.$$

Wegen $g_k(F) = 0_{End_K(V)}$ gilt daher auch $\chi_F(F) = 0_{End_K(V)}$.

□

Korollar 8.2.12 *Seien K, V, F und χ_F wie zuvor und sei g_F das Minimalpolynom von F. Dann gilt:*

$$g_F \mid \chi_F$$

Beweis: Wegen $\chi_F(F) = 0_{End_K(V)}$ gilt $\chi_F \in \langle g_k \rangle_K[t]$.

□

Normalformen führt man in der Regel ein, um große Matrizen übersichtlicher gestalten zu können. Da diese Matrixgrößen aber leicht den Rahmen eines Skripts sprengen, beschränken wir uns hier auf ein Minibeispiel, das zumindest das Vorgehen illustrieren kann.

Beispiel 8.2.13 *Sei $K = \mathbb{Q}^2$, $\mathcal{C} = (e_1, e_2)$ und $F \in End_K(V)$ gegeben durch*

$$M_{\mathcal{C}}^{\mathcal{C}}(F) = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} =: A.$$

Dann ist die charakteristische Matrix dazu

$$tE_2 - A = \begin{pmatrix} t-1 & 0 \\ 0 & t-2 \end{pmatrix}.$$

Dann ist die Smith-Normal-Form dieser Matrix:

$$\begin{pmatrix} 1 & 0 \\ 0 & t^2 - 3t + 2 \end{pmatrix}$$

Die Begleitmatrix zu $t^2 - 3t + 2$ und damit die Frobenious-Normalform zu F ist dann

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

Ob die Normalform im obigen Beispiel wirklich schöner ist als die ursprüngliche Matrix, liegt im Auge des Betrachters. Die Frobenius-Normalform zielt in jedem Fall darauf, möglichst wenige, möglichst große Blöcke zu schaffen.

Bemerkung 8.2.14 *Sei $\chi_F \in K[t]$ das charakteristische Polynom eines K-Vektorraum-Endomorphismus $F \in End_K(V)$ des n-dimensionalen K-Vektorraums V. Enthält χ_F keinen seiner Faktoren in einer zweiten oder höheren Potenz, so sind die Elementarteiler $s_1, \ldots, s_{n-1}$ Einheiten und s_n ist assoziiert zu χ_F, denn es muss gelten $s_i \mid s_{i+1}$ für alle $1 \leq i < n$. Damit ist die zugehörige Frobenius-Normalform gerade die Begleitmatrix zu χ_F – selbst wenn χ_F zufällig in Linearfaktoren zerfallen sollte.*

Das Nachrechnen der vorigen Bemerkung bleibt dem Leser als Übungsaufgabe überlassen.

Wenn wir die einzelnen Elementarteiler in Faktoren zerlegen, dann ist es möglich eine feinere Normalform zu erreichen. Ausgehend von einer Frobenius-Normalform skizzieren wir im folgenden diese Konstruktion:

Seien K, V, F und $\mathcal{D}'$ sowie $g_1, \ldots, g_k$ wie zuvor bei der Frobenius-Normalform verwendet bzw. konstruiert und sei

$$V = \bigoplus_{i=1}^{k} U(F, a_i)$$

die dort konstruierte Zerlegung von V in F-zyklische Unterräume. Wir betrachten nun einen (beliebigen, aber dann festen) dieser Unterräume genauer: $U(F, a_i)$ mit zugehörigem, bereits vollständig faktorisierten Minimalpolynom

$$g_i = \prod_{j=1}^{r_i} p_{ij}^{m_{ij}}$$

Dann ist

$$K[t]/\langle g_i \rangle_{K[t]} \cong \bigoplus_{j=1}^{r_i} K[t]/\langle p_{ij}^{m_{ij}} \rangle_{K[t]}$$

nach dem Chinesischen Restsatz und V besitzt eine Zerlegung als

$$V = \bigoplus_{i=1}^{k} \bigoplus_{j=1}^{r_i} K[t]/\langle p_{ij}^{m_{ij}} \rangle_{K[t]}.$$

Für jeden auftretenden Doppelindex ij gilt aber

$$K[t]/\langle p_{ij}^{m_{ij}} \rangle_{K[t]} = U(F, \frac{g_i}{p_{ij}^{m_{ij}}} a_i),$$

wie man direkt nachrechnen kann. Somit gilt

$$V = \bigoplus_{i=1}^{k} \bigoplus_{j=1}^{r_i} U(F, \frac{g_i}{p_{ij}^{m_{ij}}} a_i)$$

und wir können aus den Basen der einzelnen F-zyklischen Unterräume wieder eine Basis $\mathcal{D}''$ von V zusammensetzen, bzgl. derer der Endomorphismus F

dann die darstellende Matrix

$$M_{\mathcal{D}''}^{\mathcal{D}''}(F) = \begin{pmatrix} B(p_{11}^{m_{11}}) & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & B(p_{kr_k}^{m_{kr_k}}) \end{pmatrix}$$

besitzt.

Beispiel 8.2.15 *(Fortsetzung von Beispiel 8.2.13) Wir hatten die Begleitmatrix zu $t^2 - 3t + 2$ und damit die Frobenious-Normalform zu F gefunden als*

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

Nun ist $t^2 - 3t + 2 = (t-1) \cdot (t-2)$, weswegen wir den F-zyklischen Unterraum $U(F, t^2 - 3t + 2)$ weiter zerlegen können als

$$U(F, t^2 - 3t + 2) = U(F, t-1) \oplus U(F, t-2),$$

was zu einer neuen Basis $\mathcal{D}''$ (aus Eigenvektoren zu den Eigenwerten 1 und 2) führt und damit zu der Matrix

$$M_{\mathcal{D}''}^{\mathcal{D}''}(F) = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

Das fühlt sich doch schon besser an, oder?

Definition 8.2.16 *Die soeben konstruierte Bockdiagonalmatrix heißt die* **Weierstraß-Normalform** *des Endomorphismus F.*

Zum Abschluss des Kapitels schauen wir uns nun noch einen klassischen Spezialfall im Detail an, um nochmals eine andere Normalform zu erhalten: Zerfällt χ_F vollständig in Linearfaktoren, wie es z.B. über $\mathbb{C}$ stets der Fall ist, so bietet sich eine Blockdiagonalmatrix mit anderer Blockgestalt an.

Lemma 8.2.17 *Sei $f = (t - \lambda)^m$. Dann ist die Begleitmatrix $B(f)$ ähnlich zu der Matrix*

$$J_m(\lambda) = \lambda E_m + N_m = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & & \dots & \lambda & 1 \\ 0 & 0 & & \dots & 0 & \lambda \end{pmatrix},$$

wobei N_m eine ad-hoc Notation für die Matrix darstellt, die über Null als Eintrag hat außer in der oberen Nebendiagonale, wo jeder Eintrag 1 *ist.*

Dieses Lemma beweist man durch direktes Ausrechnen der Frobenius-Normalform von $J_m(\lambda)$.

Beachten Sie hier, dass die Matrix N_m nilpotent ist, wobei m gerade der Index der Nilpotenz von N_m ist, d.h. $N_m^i \neq 0_{mm}$ für alle $1 \leq i < m$, aber $N_m^m = 0_{mm}$.

Definition 8.2.18 *Die im Lemma 8.2.17 eingeführte Matrix wird als* **Jordanblock** *der Größe m zum Eigenwert λ bezeichnet.*

Bemerkung 8.2.19 *Betrachten wir nun die Bedeutung des Lemmas 8.2.17 für die Basiswahl. Wir starten mit dem F-zyklischen Unterraum $U(F, a)$, wobei $\chi_F = g_F = f = (t-\lambda)^m$ aus dem Lemma ist. Zum Erhalt der Weierstraß-Normalform als darstellender Matrix von F ist hierbei (wie auf den vorigen Seiten ausgeführt) die Basis*

$$\mathcal{D}'' = (v, F(a), \dots, F^{m-1}(a))$$

gewählt.
Wählen wir stattdessen die Basis

$$\mathcal{D}_J = ((F - \lambda Id_V)^{m-1}(a), \dots, (F - \lambda Id_V)(a), a),$$

so ergibt sich gerade J_m als darstellende Matrix. Am besten in einem Beispiel nachrechnen.

Wenden wir uns nun einer Matrix in Weierstraß-Normalform zu, deren charakteristisches Polynom vollständig in Linearfaktoren zerfällt. Dann hat die Matrix Blockdiagonalgestalt mit Blöcken, die Begleitmatrizen von Polynomen der Form $p_{ij}^{m_{ij}} = (t - \lambda_{ij})_{ij}^m$ sind. Für jeden dieser Blöcke erlaubt das vorige Lemma (zusammen mit der Bemerkung) eine Wahl einer neuen Basis, so dass der Block von der Form $J_{m_{ij}}(\lambda_{ij})$ ist. Wir haben damit in der speziellen Situation eine neue Normalform erhalten, die der Diagonalgestalt so nahe kommt, wie es nur geht:

Satz 8.2.20 *Sei V ein n-dimensionaler K-Vektorraum mit Basis $\mathcal{C}$ und $F \in End_K(V)$ mit darstellender Matrix $M_{\mathcal{C}}^{\mathcal{C}}(F) = A \in K^{n \times n}$, so dass das charakteristische Polynom χ_F in Linearfaktoren zerfällt, d.h.*

$$\chi_F = \prod_{i=1}^{\ell} (t - \lambda_i)^{m_i}$$

mit paarweise verschiedenen Eigenwerten $\lambda_i \in K$ *von* F *für* $1 \leq i \leq \ell$. *Dann gilt (in der Notation dieses Abschnitts)*

a) Es gibt eine Basis $\mathbb{C}'$ *von* V, *so dass* $M_{\mathcal{C}'}^{\mathcal{C}'}(F)$ *eine Blockdiagonalmatrix ist, die auf ihrer Diagonale die Blöcke*

$$J_{m_{11}}(\lambda_1), \ldots, J_{m_{1k_1}}(\lambda_1), \ldots, J_{m_{\ell 1}}(\lambda_\ell), \ldots, J_{m_{\ell k_\ell}}(\lambda_\ell)$$

enthält. Diese Blockdiagonalmatrix wird als die **Jordan-Normalform** *von* F *bezeichnet.*

b) Für $1 \leq i \leq \ell$ *gilt:*

$$k_i = \dim_K(Eig(F, \lambda_i)),$$

d.h. die geometrische Vielfachheit des Eigenwerts λ_i *von* F.

c) Für $1 \leq i \leq \ell$ *gilt:*

$$\sum_{j=1}^{k_i} m_{ij} = m_i,$$

d.h. die algebraische Vielfachheit des Eigenwerts λ_i *von* F.