VERALLGEMEINERUNGEN DER JORDANSCHEN NORMALFORM

THOMAS WILLWACHER

Wir haben bisher die Jordansche Normalform betrachtet für Endomorphismen $F \in \mathrm{End}(V)$ eines endlich dimensionalen K-Vektorraumes V, so dass das charakteristische Polynom

$$P_F(t) = \pm(t-\lambda_1)^{r_1}\cdots(t-\lambda_k)^{r_k} \tag{1}$$

in Linearfaktoren zerfällt. Dies ist automatisch der Fall für alle algebraisch abgeschlossenen Körper K. Ausserdem gibt es für jeden Körper K einen algebraisch abgeschlossenen Körper $\bar{K} \supset K$. Zum Beweis verweisen wir auf die Algebra Vorlesung im nächsten Jahr. (Beispiel: $K = \mathbb{R}$, $\bar{K} = \mathbb{C}$.) Dann kann man den Endomorphismus F über $\bar{K}$ betrachten, also genauer den Endomorphismus des $\bar{K}$-Vektorraumes

$$F_{\bar{K}} := id_{\bar{K}} \otimes F : \bar{K} \otimes_K V \to \bar{K} \otimes_K V.$$

Auf diesen können wir unseren Satz über die Jordansche Normalform anwenden.

Trotzdem ist es hilfreich, daneben auch Normalformen über nicht algebraisch abgeschlossenen Körpern zu haben, die nicht verlangen, dass man den Körper K verlässt. Solche Normalformen wollen wir nun anschauen.

1. Erinnerung

Vorangestellt sei eine kurze Erinnerung an die Jordansche Normalform. Wir betrachten die Jordanmatrizen

$$J_k(\lambda) := \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda \end{pmatrix} \in M(k\times k, K).$$

Ausserdem definieren wir, für A eine $k\times k$-Matrix

$$A^{\oplus s} = \begin{pmatrix} A & & & \\ & A & & \\ & & \ddots & \\ & & & A \end{pmatrix} \in M(ks\times ks, K)$$

die Blockdiagonalmatrix mit s Blöcken A auf der Blockdiagonalen. Dies ist gerade das Kroneckerprodukt von E_s und A, also sozusagen $E_s \otimes A$, oder äquivalent die (Matrix der) Abbildung

$$A \oplus \cdots \oplus A \in \mathrm{End}(K^k \oplus \cdots \oplus K^k)$$
$$(A \oplus \cdots \oplus A)(v_1 + \cdots + v_s) = A(v_1) + \cdots + A(v_s).$$

Die schon besprochenen Resultate über die Jordansche Normalform seien dann wie folgt zusammengefasst.

Theorem 1.1. *Sei F ein Endomorphismus des n-dimensionalen K-Vektorraums V (mit $n < \infty$), dessen charakteristisches Polynom in Linearfaktoren zerfällt*

$$P_F(t) = \pm(t-\lambda_1)^{r_1}\cdots(t-\lambda_k)^{r_k},$$

mit $\lambda_1, \dots, \lambda_k$ den paarweise verschiedenen Eigenwerten. Sei

$$V_j = \mathrm{Ker}\,(F - \lambda_j id_V)^{r_j}$$

der zum Eigenwert λ_j gehörende verallgemeinerte Eigenraum. Sei ausserdem

$$d_{jr} = \dim \mathrm{Ker}\,(F - \lambda_j id_V)^{r}$$

für $r = 1, 2, \dots,$ und setze $d_{jr} = 0$ für $r \le 0$. Dann gilt:

(i) Es gilt $\dim V_j = r_j$ und V zerfällt in die F-invarianten Untervektorräume V_j:

$$V = V_1 \oplus \cdots \oplus V_k\,.$$

(ii) Es gibt eine Basis $\mathcal{B}$ von V, so dass

$$M_{\mathcal{B}}(F) = \begin{pmatrix} J_1(\lambda_1)^{\oplus s_{11}} & & & & & & & \\ & J_2(\lambda_1)^{\oplus s_{12}} & & & & & & \\ & & \ddots & & & & & \\ & & & J_{r_1}(\lambda_1)^{\oplus s_{1r_1}} & & & & \\ & & & & J_1(\lambda_2)^{\oplus s_{21}} & & & \\ & & & & & \ddots & & \\ & & & & & & J_{r_k}(\lambda_k)^{\oplus s_{kr_k}} \end{pmatrix}.$$

Dabei gilt

$$s_{jr} = 2d_{jr} - d_{j(r+1)} - d_{j(r-1)} \tag{2}$$

für $j = 1, \ldots, k$ und alle r.

(iii) Das Minimalpolynom ist

$$M_F(t) = (t - \lambda_1)^{l_1} \cdots (t - \lambda_k)^{l_k}$$

mit l_j dem grössten Index mit $s_{jl_j} \neq 0$.

Man beachte auch, dass (wie wir gezeigt haben) für $r > r_j$ gilt, dass $d_{jr} = d_{jr_j} = \dim V_j$. Insbesondere ist auch $s_{jr} = 0$ für $r > r_j$. Ausserdem gilt

$$r_j = \dim V_j = s_{j1} + 2s_{j2} + \cdots r_j s_{jr_j}.$$

Wir wollen auch kurz an die Konstruktion der Basis $\mathcal{B}$ erinnern. Wir brauchen dazu nur den Fall zu betrachten, dass alle Eigenwerte gleich sind, indem wir separat Basen auf jedem V_j definieren. Sei also der notationellen Einfachheit halber gleich $k = 1$ oben, und setze $s_r := s_{1r}$ und $G := (F - \lambda_1 id_V)$. Dann betrachten wir die Untervektorräume

$$\{0\} = U_0 \subset U_1 \subset \cdots \subset U_d = V$$

mit $U_r = \operatorname{Ker} G^r$ und für ein $d \leq n$ so dass $U_d = V$. Dann wählen wir $w_1^{(d)}, \ldots, w_{s_d}^{(d)} \in U_d = V$ so, dass die Bilder in U_d/U_{d-1} eine Basis von U_d/U_{d-1} bilden. Dann wählen wir $w_1^{(d-1)}, \ldots, w_{s_{d-1}}^{(d-1)} \in U_{d-1}$ so, dass die Bilder von

$$Gw_1^{(d)}, \ldots, Gw_{s_d}^{(d)}, w_1^{(d-1)}, \ldots, w_{s_{d-1}}^{(d-1)} \in U_{d-1}$$

eine Basis von U_{d-1}/U_{d-2} bilden, usw. Unsere Basis ist dann

$$\mathcal{B} = (w_1^{(1)}, \ldots, w_{s_1}^{(1)}, Gw_1^{(2)}, w_1^{(2)}, \ldots, Gw_{s_2}^{(2)}, w_{s_2}^{(2)}, \ldots, G^{d-1}w_{s_d}^{(d)}, \ldots, Gw_{s_d}^{(d)}, w_{s_d}^{(d)}).$$

Wir wollen die jeweils "höchsten" Vektoren $w_i^{(j)}$ auch *zyklische Vektoren* nennen, um sie von den anderen Basiselementen zu unterscheiden, die durch anwenden von G erhalten werden.

Aus der Konstruktion folgt auch die Formel (2). Genauer ist s_r eindeutig bestimmt durch die Rekursionsformel (absteigende Rekursion)

$$s_r = \dim U_r - \dim U_{r-1} - (s_{r+1} + \cdots + s_d).$$

Diese Rekursion wird gelöst durch $s_r = 2\dim U_r - \dim U_{r-1} - \dim U_{r+1}$, denn mit $d_r := \dim U_r = d_{1r}$ gilt

$$\begin{aligned} s_d &= \dim U_d - \dim U_{d-1} = 2\dim U_d - \dim U_{d-1} - \dim U_{d+1} && \checkmark \\ s_r &= d_r - d_{r-1} - (s_{r+1} + \cdots + s_d) \\ &= d_r - d_{r-1} - \underbrace{(-d_r + 2d_{r+1} - d_{r+2} + \cdots - d_{d-1} + 2d_d - d_{d+1})}_{\text{Teleskopsumme}} \\ &= d_r - d_{r-1} - (-d_r + d_{r+1} + d_d - \underbrace{d_{d+1}}_{=d_d}) \\ &= 2d_r - d_{r-1} - d_{r+1} && \checkmark. \end{aligned}$$

2. Reelle Jordansche Normalform ($K = \mathbb{R}$)

Wir betrachten als nächstes separat den Fall $K = \mathbb{R}$.

Satz 2.1 (s. Fischer, Theorem 1.3.10)**.** *Sei $f \in \mathbb{R}[t]$ ein relles Polynom vom Grad $n \geq 1$. Dann hat f eine Zerlegung*

$$f = a(t-\lambda_1)\cdots(t-\lambda_k)g_1\cdots g_m, \tag{3}$$

wobei $a \in \mathbb{R}$ und die $\lambda_j \in \mathbb{R}$ die rellen Nullstellen sind und die $g_j \in \mathbb{R}[t]$ normierte quadratische Polynome ohne relle Nullstellen sind. Insbesondere ist hier $n = k + 2m$, und falls n ungerade ist muss f mindestens eine relle Nullstelle haben.

Beweis. Zunächst ist f auch ein komplexes Polynom und zerfällt als solches in Linearfaktoren.

$$f = a(t-\lambda_1)\cdots(t-\lambda_k)(t-\mu_1)(t-\bar{\mu}_1)\cdots(t-\mu_m)(t-\bar{\mu}_m)$$

Hier ordnen wir die Nullstellen so, dass die rellen Nullstellen $\lambda_j \in \mathbb{R}$ zuerst stehen. Die nicht-rellen Nullstellen treten immer in konjugiert komplexen Paaren auf, mit jeweils gleicher Vielfachheit, deshalb können wir die Linearfaktoren wie oben anordnen. Es ist jeweils

$$g_j := (t-\mu_j)(t-\bar{\mu}_j) = t^2 - 2\operatorname{Re}(\mu_j)t + |\mu_j|^2 \in \mathbb{R}[t]$$

ein relles Polynom. Ausserdem muss a als Vorfaktor von t^n in f auch reell sein. □

Sei nun $F \in \operatorname{End}(V)$ ein Endomorphismus des endlich dimensionalen $\mathbb{R}$-Vektorraums V. Fasst man dann noch gleiche Faktoren in (3) zusammen hat man eine Faktorisierung des charakteristischen Polynoms

$$P_F(t) = \pm(t-\lambda_1)^{r_1}\cdots(t-\lambda_k)^{r_k}g_1^{q_1}\cdots g_m^{q_m}, \tag{4}$$

wobei $\lambda_1, \ldots, \lambda_k$ die paarweise verschiedenen rellen Nullstellen sind und $g_1, \ldots, g_m \in \mathbb{R}[t]$ paarweise verschiedene normierte quadratische relle Polynome ohne relle Nullstellen. Es gilt insbesondere

$$n = \dim V = r_1 + \cdots + r_k + 2(q_1 + \cdots + q_m).$$

Betrachte nun einen quadratischen Faktor g_j mit nicht-rellen Nullstellen $\mu_j, \bar{\mu}_j \in \mathbb{C} \setminus \mathbb{R}$. Sei $\mu_j = a_j + ib_j$ mit $a_j, b_j \in \mathbb{R}$. Dann ist

$$g_j = t^2 - 2\operatorname{Re}(\mu_j)t + |\mu_j|^2 = t^2 - 2a_jt + a_j^2 + b_j^2 = (t-a_j)^2 + b_j^2 = P_{A_j}(t)$$

das charakteristische Polynom der Matrix

$$A_j := \begin{pmatrix} a_j & b_j \\ -b_j & a_j \end{pmatrix}. \tag{5}$$

Beachte auch, dass diese Matrix eine Drehstreckung beschreibt, also von der Form cR ist mit $c = \sqrt{a_i^2 + b_i^2} \in \mathbb{R}_{>0}$, $R \in \mathrm{SO}(2)$.

Schliesslich definieren wir noch für $A \in M(2\times 2, \mathbb{R})$ die Block-Jordanmatrix

$$J_r(A) := \begin{pmatrix} A & E_2 & & & \\ & A & E_2 & & \\ & & \ddots & \ddots & \\ & & & \ddots & E_2 \\ & & & & A \end{pmatrix} \in M(2r\times 2r, \mathbb{R}).$$

Dann hat man die folgende relle Version der Jordanschen Normalform.

Theorem 2.2. *Sei F ein Endomorphismus des n-dimensionalen $\mathbb{R}$-Vektorraums V (mit $n < \infty$). Seien $\lambda_1, \ldots, \lambda_k, r_1, \ldots, r_k, g_1, \ldots, g_m, q_1, \ldots, q_m$ wie in (4). Sei*

$$V_j = \operatorname{Ker}(F - \lambda_j id_V)^{r_j}$$

für $j = 1, \ldots, k$ und

$$V_j' = \operatorname{Ker}(g_j(F))^{q_j}.$$

für $j = 1, \ldots, m$. Sei ausserdem

$$d_{jr} = \dim \operatorname{Ker}(F - \lambda_j id_V)^r$$

$$d_{jr}' = \dim \operatorname{Ker}(g_j(F)^r)^r$$

für $r = 1, 2, \ldots,$ und setze $d_{jr} = d_{jr}' = 0$ für $r \leq 0$. Dann gilt:

(i) Es gilt $\dim V_j = r_j$ *und* $\dim V_j' = 2q_j$ *und* V *zerfällt wie folgt in* F*-invariante Untervektorräume:*

$$V = V_1 \oplus \cdots \oplus V_k \oplus V_1' \oplus \cdots \oplus V_m' \,. \tag{6}$$

(ii) Es gibt eine Basis $\mathcal{B}$ *von* V*, so dass*

(7)

$$M_{\mathcal{B}}(F) = \begin{pmatrix} J_1(\lambda_1)^{\oplus s_{11}} & & & & & & & & & & \\ & \ddots & & & & & & & & & \\ & & J_{r_k}(\lambda_k)^{\oplus s_{kr_k}} & & & & & & & & \\ & & & J_1(A_1)^{\oplus s'_{11}} & & & & & & & \\ & & & & J_2(A_1)^{\oplus s'_{12}} & & & & & & \\ & & & & & \ddots & & & & & \\ & & & & & & J_{q_1}(A_1)^{\oplus s'_{1q_1}} & & & & \\ & & & & & & & J_1(A_2)^{\oplus s'_{2q_1}} & & & \\ & & & & & & & & \ddots & & \\ & & & & & & & & & J_{q_m}(A_m)^{\oplus s'_{mq_m}} \end{pmatrix}$$

mit A_j *wie in* (5)*. Dabei gilt*

$$s_{jr} = 2d_{jr} - d_{j(r+1)} - d_{j(r-1)} \tag{8}$$

für $j = 1, \ldots, k$ *und alle* r *und*

$$2s'_{jr} = 2d'_{jr} - d'_{j(r+1)} - d'_{j(r-1)} \tag{9}$$

für $j = 1, \ldots, m$ *und alle* r*.*

(iii) Das Minimalpolynom ist

$$M_F(t) = (t - \lambda_1)^{l_1} \cdots (t - \lambda_k)^{l_k} g_1^{l_1'} \cdots g_m^{l_m'}$$

mit l_j *bzw.* l_j' *den grössten Indizes mit jeweils* $s_{jl_j} \neq 0$ *bzw.* $s'_{jl_j'} \neq 0$*.*

Bevor wir zum Beweis kommen wollen wir noch bemerken:

Korollar 2.3. *Zwei relle Matrizen sind ähnlich (über* $\mathbb{R}$*) genau dann wenn ihre charakteristischen Polynome übereinstimmen und zusätzlich alle Zahlen* d_{jr}*,* d'_{jr} *wie im Theorem.*

Beweis. Sind die Matrizen ähnlich, so sind die charakteristischen Polynome gleich. Die Zahlen d_{jr}, d'_{jr} sind jeweils Dimensionen von Kernen von Polynomen in den Matrizen, und damit auch invariant unter Ähnlichkeit. (Ausführlich: ist $A = SBS^{-1}$ und P ein Polynom, so ist $P(A) = SP(B)S^{-1}$ ähnlich zu $P(B)$ und hat damit gleichen Rang.)

Stimmen umgekehrt die angegebenen Daten überein, so haben die Matrizen nach dem Theorem die gleiche reelle Jordansche Normalform (zu der sie jeweils ähnlich sind), und sind damit offensichtlich ähnlich. □

Korollar 2.4. *Zwei relle Matrizen sind ähnlich über* $\mathbb{R}$ *genau dann wenn Sie ähnlich über* $\mathbb{C}$ *sind.*

Beweis. Alle rellen Matrizen sind auch komplexe Matrizen, deshalb impliziert Ähnlichkeit über $\mathbb{R}$ Ähnlichkeit über $\mathbb{C}$.

Seien nun umgekehrt die rellen Matrizen A und B ähnlich über $\mathbb{C}$. Also $A_{\mathbb{C}} = SB_{\mathbb{C}}S^{-1}$ mit $S \in \mathrm{GL}(n, \mathbb{C})$, wobei wir hier temporär ganz pedantisch $A_{\mathbb{C}} = A$ für die relle Matrix A schreiben, einfach aufgefasst als komplexe Matrix. Wir wollen auch erinnern, dass für jede Matrix C gilt, dass der (reelle) Rang von C und der (komplexe) Rang von $C_{\mathbb{C}}$ gleich sind, also insbesondere

$$\dim_{\mathbb{R}} \operatorname{Ker} C = \dim_{\mathbb{C}} \operatorname{Ker} C_{\mathbb{C}}.$$

Dann gilt also für jedes relle Polynom P

$$\dim_{\mathbb{R}}(\operatorname{Ker} P(A)) = \dim_{\mathbb{C}}(\operatorname{Ker} P(A_{\mathbb{C}})) = \dim_{\mathbb{C}}(\operatorname{Ker} P(B_{\mathbb{C}})) = \dim_{\mathbb{R}}(\operatorname{Ker} P(B)) \,.$$

Insbesondere sind also die Invarianten d_{jr} und d'_{jr} für A und B gleich. Also folgt die Aussage aus Korollar 2.3. (Hinweis: Wir werden die pedantische Notation $A_{\mathbb{C}}$ später normalerweise nicht mehr verwenden und schreiben einfach A, ausser das führt zu Verwirrung.) □

Korollar 2.5. *Jede relle Matrix ist ähnlich zu ihrer Transponierten.*

Beweis. Für jedes Polynom P gilt $P(A^T) = P(A)^T$, und Transposition ändert den Rang nicht (denn Zeilenrang = Spaltenrang). Deswegen sind wiederum die Invarianten d_{jr} und d'_{jr} für A und A^T gleich. □

Schliesslich:

Beweis vom Theorem (z.T. nur skizziert). Durch Wahl einer (rellen) Basis reicht es den Fall zu betrachten, dass $V = \mathbb{R}^n$ und F durch eine Matrix $A \in M(n\times n, \mathbb{R})$ gegeben ist. Wir betrachten zunächst A als komplexe Matrix und F (bzw. genauer $id_{\mathbb{C}} \otimes_{\mathbb{R}} F$) als Endomorphismus von $\mathbb{C}^n \cong \mathbb{C} \otimes_{\mathbb{R}} \mathbb{R}^n$. Wir wenden dann Theorem 1.1 an und betrachten die komplexe JNF von A bzw. F. Das charakteristische Polynom faktorisert dabei als

$$P_F = \pm(t-\lambda_1)^{r_1}\cdots(t-\lambda_k)^{r_k} g_1^{q_1}\cdots g_m^{q_m} = \pm(t-\lambda_1)^{r_1}\cdots(t-\lambda_k)^{r_k}(t-\mu_1)^{q_1}(t-\bar\mu_1)^{q_1}\cdots(t-\mu_m)^{q_m}(t-\bar\mu_m)^{q_m},$$

wobei jeweils $\mu_j, \bar\mu_j$ die Nullstellen von g_j sind. Wir wählen nun die Basis $\mathcal{B}$ aus Theorem 1.1 (-man erinnere sich an deren Konstruktion im vorherigen Abschnitt-) wie folgt aus. Zunächst ist für jede relle Matrix C der Kern von $C_{\mathbb{C}}$ einfach die Komplexifizierung von $\operatorname{Ker} C$. (Konkret: Jede Basis von $\operatorname{Ker} C \subset \mathbb{R}^n$ ist auch eine (komplexe) Basis von $\operatorname{Ker} C_{\mathbb{C}} \subset \mathbb{C}^n$.) Entsprechend kann man die Basis der Räume $\operatorname{Ker}(A_{\mathbb{C}} - \lambda_j E_n)^{r_j}$ genau wie oben reell wählen, indem man einfach die Vektoren $w_j^{(i)}$ reell wählt.

Ausserdem wählen wir die Basis von

$$\operatorname{Ker}(A_{\mathbb{C}} - \bar\mu_j E_n)^{q_j} = \operatorname{Ker}\overline{(A_{\mathbb{C}} - \mu_j E_n)^{q_j}} = \overline{\operatorname{Ker}(A_{\mathbb{C}} - \mu_j E_n)^{q_j}}$$

einfach als die komplex konjugierte Basis zu der von $\operatorname{Ker}(A_{\mathbb{C}} - \mu_j E_n)^{q_j}$.

Sei nun $v \in \mathbb{C}^n$ ein zu einem Jordanblock der Grösse s (zu Eigenwert μ_j) gehörender zyklischer Vektor in unserer Basis. Das heisst $(A-\mu_j E_n)^s v = 0$ und $v, (A-\mu_j E_n)v, (A-\mu_j E_n)^2 v, \ldots, (A-\mu_j E_n)^{s-1}v$ sind Teile unserer Basis, die den zu dem Jordanblock gehörenden s-dimensionalen (A-invarianten) Untervektorraum aufspannen. Setzten wir $v_j := (A-\mu_j E_n)^{s-j} v$ (mit $j = 1, \ldots, s$) so gilt also

$$Av_j = \mu_j v_j + v_{j-1}$$

mit $v_0 := 0$. Entsprechend sind dann $\bar v = \bar v_s, (A-\bar\mu_j E_n)\bar v = \bar v_{s-1}, \ldots, (A-\bar\mu_j E_n)^{s-1}\bar v = \bar v_1$ auch Teile unserer Basis, die zu einem Jordanblock zum Eigenwert $\bar\mu_j$ gehören, und wir haben entsprechend

$$A\bar v_j = \bar\mu_j \bar v_j + \bar v_{j-1}.$$

Die im Theorem 2.2 gesuchte relle Basis erhalten wir nun, indem wir jeweils diese komplex konjugierten Paare $v_j, \bar v_j$ von Basisvektoren ersetzen durch die rellen Basisvektoren

$$e_j := \operatorname{Re}(v_j) = \frac{1}{2}(v_j + \bar v_j) \in \mathbb{R}^n \subset \mathbb{C}^n$$
$$f_j := \operatorname{Im}(v_j) = \frac{1}{2i}(v_j - \bar v_j) \in \mathbb{R}^n \subset \mathbb{C}^n.$$

Offensichtlich ist (e_j, f_j) eine Basis von $\operatorname{span}(v_j, \bar v_j)$, deshalb erhalten wir wieder eine Basis, wenn wir $v_j, \bar v_j$ aus $\mathcal{B}$ jeweils durch e_j, f_j ersetzen. Ausserdem gilt dann mit $\mu_j = a_j + ib_j$

$$\begin{aligned} Ae_j &= \operatorname{Re}(\mu_j v_j + v_{j-1}) = a_j e_j - b_j f_j + e_{j-1} \\ Af_j &= \operatorname{Im}(\mu_j v_j + v_{j-1}) = b_j e_j + a_j f_j + f_{j-1}. \end{aligned} \tag{10}$$

Sei $\mathcal{A}$ die so konstruierte Basis des $\mathbb{C}^n$ aus rellen Vektoren. Dann ist $\mathcal{A}$ auch eine Basis des $\mathbb{R}^n$. Ausserdem folgt dann aus (10) gerade (7), bei entsprechender Ordnung der Basisvektoren. Damit ist der erste Teil von Aussage (ii) von Theorem 2.2 gezeigt.

Ausserdem gibt es nach Konstruktion gleich viele Jordanblöcke der Form $J_r(A_j)$ in der rellen JNF von A wie Jordanblöcke der Form $J_r(\mu_j)$ in der komplexen JNF von A. Die Formel (9) folgt damit direkt aus der Rechnung

$$\begin{aligned} \dim_{\mathbb{R}} \operatorname{Ker} g_j(A)^r &= \dim_{\mathbb{C}} \operatorname{Ker} g_j(A_{\mathbb{C}})^r \\ &= \dim_{\mathbb{C}} \operatorname{Ker}(A_{\mathbb{C}} - \mu_j E_n)^r (A_{\mathbb{C}} - \bar\mu_j E_n)^r \\ &= \dim_{\mathbb{C}} \operatorname{Ker}(A_{\mathbb{C}} - \mu_j)^r + \dim_{\mathbb{C}} \operatorname{Ker}(A_{\mathbb{C}} - \bar\mu_j)^r \\ &= 2\dim_{\mathbb{C}} \operatorname{Ker}(A_{\mathbb{C}} - \mu_j), \end{aligned}$$

wobei die 3. Gleichheit folgt, indem man einfach A durch die (zu A ähnliche, komplexe) JNF ersetzt (ÜA). Die Formel (8) folgt gleich wie im komplexen Fall. Alternativ kann man für A in reller JNF auch die Zahlen d_{jr} und d'_{jt} gleich explizit ausrechnen und sieht, dass (8) und (9) gelten. Dies zeigt dann auch, dass die Formel (8) und (9) für jede Basis $\mathcal{B}$ gelten, die die Matrix in die Form (7) bringt.

Damit ist Aussage (ii) gezeigt. Ausserdem folgt aus der Rechnung die Formel für die Dimensionen in (i).

Aussage (iii) folgt dann einfach aus der Tatsache, dass $M_A = M_{A_{\mathbb{C}}}$ ist - die Minimalpolynome von A als reller und komplexer Matrix sind gleich (ÜA oder Linalg I).

Es fehlt noch der Beweis der direkten Summenzerlegung (6) in (i). Hierzu kann man am einfachsten direkt annehmen, dass A schon in reller JNF (also in der Form (7)) vorliegt. Es ist dann

$$g_j(J_r(A_i)) = \begin{pmatrix} g_j(A_i) & * & * \\ & \ddots & * \\ & & g_j(A_i) \end{pmatrix} \in M(2r \times 2r, \mathbb{R})$$

eine obere Block-Dreiecksmatrix. Für $i = j$ sind die Einträge auf der Diagonalen = 0 (wegen des Theorems von Caley-Hamilton), und ansonsten invertierbar (ÜA). Also ist

$$g_j(J_r(A_i))^r = \begin{cases} 0 & \text{für } i = j \\ \text{invertierbar} & \text{für } i \neq j \end{cases}.$$

Also ist (für A in reeller JNF)

$$\mathrm{Ker}(g_j(A)^{r_j})$$

eine Blockdiagonalmatrix von gleicher Blockdiagonalform wie (7), aber mit verschwindenden Diagonalblöcken anstelle der $J_r(A_j)$, und mit invertierbaren Diagonalblöcken, sonst. Der Kern ist also aufgespannt von den Einheitsvektoren mit Einsen an den Stellen der $J_r(A_j)$, und die Zerlegung (6) folgt. □

3. Jordansche Normalform für allgemeine Körper K

Es gibt verschiedene Normalformen für Matrizen über einem allgemeinen Körper K. Wir wollen hier eine anschauen, die sehr nah an der Jordanschen Normalform ist. In der Algebra Vorlesung wird dies (wahrscheinlich) noch einmal aufgegriffen werden, in folgender Sprache: Wie wir schon gesehen haben definiert jedes $F \in \mathrm{End}(V)$ (V ein endlich dimensionaler K-VR) einen Homomorphismus von Ringen mit Eins $K[t] \to \mathrm{End}(V)$, definiert durch $P \mapsto P(F)$. In der Sprache der Algebra sagt man auch, dass V damit ein $K[t]$-Modul ist. In der Algebra werden Sie sehr allgemein die endlich erzeugten Moduln von Hauptidealringen klassifizieren, was in unserem Beispiel und in unserer Sprache gerade interpretiert werden kann als die Herleitung einer Normalform für Endomorphismen.

Die Themen hier sind z.T. ein Ausblick auf die Algebra Vorlesung im nächsten Jahr, und sind (allenfalls mit Ausnahme von 3.1) nicht Gegenstand der Prüfung.

3.1. **Begleitmatrizen.** Sei $f = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0 \in K[t]$ ein Polynom vom Grad $n \geq 0$. Dann ist die *Begleitmatrix* von f die folgende $n \times n$ Matrix:

$$B_f := \begin{pmatrix} -a_{n-1} & 1 & 0 & \cdots & 0 \\ -a_{n-2} & 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ -a_1 & 0 & \cdots & \cdots & 1 \\ -a_0 & 0 & \cdots & \cdots & 0 \end{pmatrix}.$$

Satz 3.1. *Das charakteristische und das Minimalpolynom von B_f sind gleich f, bis auf Vorzeichen*

$$(-1)^n P_{B_f} = M_{B_f} = f.$$

Beweis. Die Aussage $P_{B_f} = (-1)^n f$ war Inhalt einer Übungsserie (genauer in der Form $P_{B_f^T} = (-1)^n f$, aber natürlich gilt $P_{B_f^T} = P_{B_f}$).

Ausserdem gilt allgemein, dass M_{B_f} das charakteristische Polynom P_{B_f} teilt. Also müssen wir nur noch zeigen, dass $\deg M_{B_f} = n$. Die ist aber gleichbedeutend mit der Aussage, dass die Matrizen

$$E_n = B_f^0, B_f = B_f^1, B_f^2, \ldots, B_f^{n-1}$$

linear unabhängig sind. Da aber die Einträge von B_f oberhalb der ersten Nebendiagonalen verschwinden gilt für $j = 0, 1, \ldots, n-1$

$$B_f^j = \begin{pmatrix} * & \cdots & * & 1 & 0 & \cdots & 0 \\ * & \cdots & * & * & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ * & \cdots & * & * & * & \cdots & 1 \\ \vdots & \ddots & & & & \ddots & \vdots \\ * & \cdots & * & * & * & \cdots & * \end{pmatrix},$$

wobei hier die 1 in der ersten Zeile an der $j+1$-ten Stelle steht.

Sei nun $A := \sum_{j=0}^{n-1} \lambda_j B_f^j = 0$, und wir wollen zeigen, dass $\lambda_1 = \cdots = \lambda_n = 0$. Zunächst ist wegen der Vorbemerkung der $(1,n)$-Eintrag von A gerade λ_{n-1}, also $\lambda_{n-1} = 0$. Aber dann ist der $(1,n-1)$-Eintrag von A gerade λ_{n-2}, also auch $\lambda_{n-2} = 0$. So fährt man fort und sieht, dass alle $\lambda_j = 0$ sind, also die Potenzen von B_f linear unabhängig wie verlangt. □

3.2. **Teilbarkeitstheorie von Polynomen.** Wir wollen nun einige Eigenschaften des Ringes der Polynome $K[t]$ besprechen. Wir beginnen mit einer Repetition einiger Aussagen aus LA1 (siehe auch Fischer 4.5.5):

- Sei R ein kommutativer Ring und $f, g \in R \setminus \{0\}$. Dann heisst f *teilt* g, oder in Symbolen $f \mid g$, falls ein $h \in R$ existier mit $fh = g$. Falls f das Element g nicht teilt schreiben wir $f \nmid g$.
- Ein Ideal $I \subset R$ des kommutativen Ringes R ist eine Unterring, für den zusätzlich $IR \subset I$ gilt. Äquivalent, und expiziter: Für $f, g \in I$, $h \in R$ ist $f - g \in I$ und $fh \in I$.
  Für $I_1, I_2 \subset R$ Ideale sind $I_1 \cap I_2$ und $I_1 + I_2$ auch wieder Ideale.
- Für jeden Körper K ist der Polynomring $K[t]$ ein Hauptidealring. Konkret gibt es zu jedem Ideal $\{0\} \neq I \subset K[t]$ ein eindeutiges normiertes Polynom $M_I \in I$ so dass $I = M_I R = \{M_I f \mid f \in K[t]\}$. Konkret ist M_I das normierte Polynom kleinsten Grades in I und heisst Minimalpolynom von I.
- Aus der Übung: Für jeden Körper K ist $K[t]$ ein faktorieller Ring, d.h., er hat keine Nullteiler: aus $fg = 0$ folgt $f = 0$ oder $g = 0$.

Für Polynome $f, g \in K[t]$ (mit $f, g \neq 0$) definieren wir den grössten gemeinsamen Teiler $\mathrm{ggT}(f,g) \in K[t]$ von f und g als das normierte Polynom vom grössten Grad, das sowohl f als auch g teilt. Dies ist auch das Minimalpolynom von $fK[t] + gK[t]$, also $\mathrm{ggT}(f,g) = M_{fK[t]+gK[t]}$ (ÜA). Insbesondere existieren in diesem Fall also $a, b \in K[t]$ so dass

$$\mathrm{ggT}(f,g) = af + bg.$$

Praktisch berechnet man den grössten gemeinsamen Teiler mit dem Euklidischen Algorithmus, siehe dazu die Übungsserie. Man kann die Definition auch auf mehrere Polynome ausweiten. Dann ist $\mathrm{ggT}(f_1, \ldots, f_n)$ das normierte Polynom vom grössten Grad, dass alle f_j teilt, oder äquivalent das Minimalpolynom des Ideals $f_1K[t] + \cdots + f_nK[t]$. Dann gibt es also wieder $a_1, \ldots a_n \in K[t]$ so dass

$$\mathrm{ggT}(f_1, \ldots, f_n) = a_1 f_1 + \cdots + a_n f_n,$$

und diese sind auch mit dem Euklidischen Algorithmus berechenbar (s. Serie).

Ebenso definieren wir das kleinste gemeinsame Vielfache $\mathrm{kgV}(f,g)$ als das Polynom vom kleinsten Grad, dass sowohl von f als auch von g geteilt wird. Es ist dann (per Definition) $\mathrm{kgV}(f,g) = M_{fK[t]\cap gK[t]}$ das Minimalpolynom des Schnittes der von f und g erzeugten Ideale. Analog definiert man $\mathrm{kgV}(f_1, \ldots, f_n)$ für mehrere Arguments.

Lemma 3.2. *Für ein Polynom $f \in K[t]$ positiven Grades sind äquivalent:*

- *Aus $f = gh$ folgt, dass ein $0 \neq c \in K$ existiert so dass $f = cg$ oder $f = ch$. (Also hat f keine nichttrivialen Teiler, also nur die Teiler c und cf für $0 \neq c \in K$.Man sagt auch f ist* irreduzibel.*)*
- *Aus $f \mid gh$ (für $0 \neq g, h \in K[t]$) folgt, dass $f \mid g$ oder $f \mid h$. (Man sagt in diesem Fall auch f ist* prim.*)*

Beweis. "⇒": Sei f irreduzibel und $f \mid gh$, also $fr = gh$ für ein $r \in K[t]$. Wir nehmen an, f würde weder g noch h teilen. Dann muss gelten $\mathrm{ggT}(f,g) = \mathrm{ggT}(f,h) = 1$. (Denn der grösste gemeinsame Teiler ist insbesondere ein Teiler von f, und f hat nur Teiler der Form c oder cf, aber cf ist kein Teiler von g, h, da f keiner ist.) Also gibt es $a, b, a', b' \in K[t]$ so dass $af + bg = 1 = a'f + b'h$. Damit ist dann

$$\begin{aligned} 1 = 1 \cdot 1 &= (af + bg)(a'f + b'h) \\ &= aa'f^2 + afb'h + bga'f + a'b' \underbrace{gh}_{=rf} \\ &= (aa'f + ab'h + bga' + a'b'r)f. \end{aligned}$$

Also wäre 1 durch f teilbar, ein Widerspruch zu $\deg f > 0$.

"⇐": Sei nun f prim und nehme an, f wäre nicht irreduzibel, also $f = gh$ für $g, h \in K[t]$ von jeweils positivem Grad. Dann ist insebsondere

$$\deg g < \deg f \qquad \deg h < \deg f. \tag{11}$$

Dann teilt f insbesondere gh. Da f prim ist folgt $f \mid g$ oder $f \mid h$, im Widerspruch zu (11). □

Als Folgerung könne wir dann zeigen:

Satz 3.3 (Primfaktorzerlegung für Polynome)**.** *Sei $0 \neq f \in K[t]$ ein Polynom. Dann gibt es irreduzible normierte Polynome (positiven Grades) $p_1, \ldots, p_n \in K[t]$ (mit $n \geq 0$) und ein $0 \neq c \in K$ so dass*

$$f = cp_1 \cdots p_n .$$

Diese Faktorisierung ist eindeutig, bis auf der Anordnung der p_j.

Man sagt auch: $K[t]$ ist ein *faktorieller Ring.*

Beweis: Existenz: Mit Induktion über $\deg f$. Ist $\deg f = 0$ so ist $c := f$ und nichts zu zeigen. Ist f schon irreduzibel, so wählt man c so, dass $p_1 := \frac{1}{c} f$ normiert ist und hat dann $f = cp_1$. Ist f nicht irreduzibel, so können wir schreiben $f = gh$ mit $g, h \in K[t]$ von kleinerem Grad als f. Wendet man die Induktionsannahme an auf g, h, so erhält man direkt die gewünschte Faktorisierung.

Eindeutigkeit: Seien zwei Faktoriserungen in irreduzible Polynome gegeben:

$$cp_1 \cdots p_n = f = dq_1 \cdots q_m .$$

Zunächst sind c und d jeweils die Koeffizienten der führenden Potent von t auf beiden Seiten, also $c = d$. Ausserdem sieht man, dass $p_1 \mid q_1 \cdots q_m$. Da $q_1, \ldots, q_m$ aber irreduzibel, also nach dem Lemma auch prim sind, gibt es ein j so dass $p_1 \mid q_j$. Damit gilt aber (wieder wegen der Irreduzibilität von q_j, und der Normiertheit von p_1, q_j) dass $p_1 = q_j$. Durch umordnen der qj können wir ausserdem o.B.d.A. annehmen, dass $j = 1$ ist. Dann ist aber

$$0 = p_1 \cdots p_n - q_1 \cdots q_m = p_1 (p_2 \cdots p_n - q_2 \cdots q_m),$$

also

$$p_2 \cdots p_n = q_2 \cdots q_m$$

da $K[t]$ nullteilerfrei ist. So fährt man nun analog mit p_2 (statt p_1) fort und erhält schlussendlich, dass (allenfalls nach Umordnen) $p_j = q_j$ für alle j und $m = n$. □

Definition 3.4. *Ein Ideal $I \subset R$ in einem kommutativen Ring R heisst* maximal*, falls $I \neq R$ und es kein Ideal J gibt so dass*

$$I \subsetneq J \subsetneq R.$$

I heisst prim*, falls $I \neq R$ und für alle $f, g \in R$ gilt $fg \in I \Rightarrow f \in I \vee g \in I$.*

In unserem Fall ($R = K[t]$) folgt direkt aus der Definition, dass ein Ideal ein Primideal ist, genau dann wenn sein Minimalpolynom prim ist. Ebenso ist das Ideal maximal genau dann wenn sein Minimalpolynom keine nicht-trivialen Teiler hat, also irreduzibel ist. Also sind für $K[t]$ die Primideale und die maximalen Ideale identisch. (Für allgemeinere Ringe ist nicht jedes Primideal maximal.)

Es gibt noch ein weiteres hilfreiches Kriterium.

Lemma 3.5. *Sei $I \subsetneq R$ ein Ideal im kommutativen Ring mit Eins R. Dann gilt:*

- *I ist maximal genau dann wenn R/I mit der von R geerbten Addition und Multiplikation ein Körper ist.*
- *I ist genau dann prim wenn R/I nullteilerfrei ist.*

Beweis. Zunächst sei allgemein bemerkt, dass für jedes Ideal $I \subset R$ der Quotient

$$R/I = R/\sim$$

durch die Äquivalenzrelation

$$f \sim g \Leftrightarrow f - g \in I$$

definiert ist. Wir schreiben $f + I \in R/I$ für die Äquivalenzklasse von $f \in R$. Die Addition und Multiplikation von R induzieren eine Addition und Multiplikation auf R/I durch

$$(f + I) + (g + I) = (f + g) + I \qquad (f + I)(g + I) = (fg) + I.$$

Man zeige als Übungsaufgabe, dass diese Operationen wohldefiniert sind, das heisst unabhängig von der Wahl der Repräsentanten f, g, und ausserdem die Ringaxiome von R erben. Ist ausserdem $I \neq R$ und hat R ein Einselement $1 \in R$, so ist $1 + I$ ein Einselement im Ring R/I und ausserdem $1 + I \neq 0 + I$.

Die Aussage, dass R/I nullteilerfrei ist, ist dann

$$(a + I)(b + I) = 0 + I \Rightarrow (a + I) = 0 + I \vee (b + I) = 0 + I.$$

Dies ist gleichbedeutend mit

$$ab \in I \Rightarrow a \in I \vee b \in I,$$

was aber gerade heisst, das I prim ist.

Sei nun I maximal. Wir müssen dann zeigen dass R/I sogar ein Körper ist, also zusätzlich, dass jedes Element $f + I \neq 0 + I$ ein multiplikatives Inverses hat. Da $f \neq I$ ist das ideal $fR + I$ echt grösser als I. Also gilt wegen der Maximalität von I dass $fR + I = R$. Das heisst aber, es gibt insbesondere $g \in R$ und $h \in I$ so dass $1 = fg + h$, oder äquivalent

$$1 + I = fg + I = (f + I)(g + I).$$

Also ist $g + I$ das multiplikative Inverse zu $f + I$.

Sei zuletzt umgekehrt R/I ein Körper und wir wollen zeigen, dass I maximal ist. Sei also $J \supsetneq I$ ein weiteres Ideal, und wir müssen zeigen, dass $J = R$, oder äquivalent $1 \in J$. Sei $f \in J \setminus I$ und wähle ein multiplikatives Inverses $g + I$ von $f + I$. Dann ist aber $fg - 1 \in I \subset J$ und daher auch $1 \in J$, denn mit $f \in J$ ist auch $fg \in J$. □

3.3. **Jordansche Normalform.** Wir definieren nun die Matrix

$$\tilde{E}_n := \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \end{pmatrix} \in M(n \times n, K)$$

mit einem Eintrag 1 unten links und sonst nur Nullen. Ausserdem definieren wir für $A \in M(n \times n, K)$

$$\tilde{J}_r(A) := \begin{pmatrix} A & \tilde{E}_n & & & \\ & A & \tilde{E}_n & & \\ & & \ddots & \ddots & \\ & & & \ddots & \tilde{E}_n \\ & & & & A \end{pmatrix} \in M(nr \times nr, K).$$

Für ein irreduzibles normiertes Polynom f nennen wir $\tilde{J}_r(B_f)$ den zu f gehörigen Jordanblock der Grösse nr.

Theorem 3.6 (Jordansche Normalform)**.** . *Sei V ein endlich dimensionaler k-Vektorraum und $F \in \mathrm{End}(V)$. Sei*

$$P_F = \pm p_1^{r_1} \cdots p_k^{r_k} \tag{12}$$

die Faktorisierung des charakteristischen Polynoms in irreduzible normierte Polynome $p_1, \ldots, p_k$, die wir als paarweise verschieden annehmen können. Dann gibt es eine Basis $\mathcal{B}$ von V so dass $M_{\mathcal{B}}$ blockdiagonal ist, wobei die Blöcke auf der Blockiagonalen alle Jordanblöcke von der Form $\tilde{J}_r(B_{p_j})$ sind mit $j = 1, \ldots, k$ und $r = 1, \ldots, r_j$.

Sei dabei s_{jr} die Anzahl Jordanblöcke von der Grösse $\tilde{J}_r(B_{p_j})$. Dann gilt für alle j

$$\sum_{r=1}^{r_j} s_{jr} = r_j \deg p_j$$

und

$$s_{jr} = \frac{1}{\deg p_j} \left(2\dim \mathrm{Ker}(p_j(F)^r) - \dim \mathrm{Ker}(p_j(F)^{r+1}) - \dim \mathrm{Ker}(p_j(F)^{r-1}) \right). \tag{13}$$

Insbesondere ist also die Jordansche Normalform wieder eindeutig, bis auf Permutation der Diagonalblöcke.

Den Beweis werden wir in Teilschritte zerlegen.

Satz 3.7. *Sei V ein endlich dimensionaler K-Vektorraum und $F \in \mathrm{End}(V)$. Sei*

$$P_F = \pm p_1^{r_1} \cdots p_k^{r_k}$$

die Faktorisierung des charakteristischen Polynoms von F in irreduzible Polynome wie in (12)*. Sei*

$$V_j := \mathrm{Ker}\left(p_j^{r_j}(F) \right).$$

Dann zerfällt V in die direkte Summe der F-invarianten Untervektorräume V_j,

$$V = V_1 \oplus \cdots \oplus V_k.$$

Ferner gilt für das charakteristische Polynom der Einschränkung $F|V_j$ von F auf den invarianten Untervektorraum V_j dass

$$P_{F|V_j} = \pm p_j^{r_j}.$$

Insbesondere ist also $\dim_K(V_j) = r_j \deg p_j$.

Beweis. Betrachte die Polynome

$$\hat{P}_j := \prod_{i \neq j} p_i^{r_i},$$

das heisst im Vergleich zu P_F fehlt einfach der Faktor $p_j^{r_j}$. Wegen Caley-Hamilton gilt

$$0 = P(F) = \pm p_j^{r_j}(F)\hat{P}_j(F),$$

also ist Insbesondere

$$\operatorname{Im} \hat{P}_j(F) \subset \operatorname{Ker} p_j^{r_j}(F) = V_j.$$

Da nun die irreduziblen Faktoren p_j paarweise verschieden sind gilt

$$\operatorname{ggT}(\hat{P}_1, \ldots, \hat{P}_k) = 1.$$

Also gibt es $q_1, \ldots, q_k \in K[t]$ so dass

$$1 = \hat{P}_1 q_1 + \cdots + \hat{P}_k q_k.$$

Damit gilt dann für jedes $v \in V$

$$v = \left(\hat{P}_1(F)q_1(F) + \cdots + \hat{P}_k(F)q_k(F)\right)(v) = \sum_{j=1}^{k} \underbrace{\hat{P}_j(F)(q_j(F)(v))}_{\in V_j}. \tag{14}$$

Also gilt schonmal

$$V = V_1 + \cdots + V_k.$$

Wir müssen noch zeigen, dass die Summe direkt ist. Sei dazu also

$$v_1 + \cdots + v_k = 0 \tag{15}$$

mit $v_j \in V_j$, $j = 1, \ldots, k$. Beachte nun, dass gilt $V_j \subset \operatorname{Ker} \hat{P}_i(F)$ für $i \neq j$. Aus (14) folgt dann auch, dass für $v = v_j \in V_j$ gilt

$$q_j(F)\hat{P}_j(F)(v_j) = v_j.$$

($q_j(F)\hat{P}_j(F)$ ist also der Projektor auf V_j.) Wenden wir auf beide Seiten von (15) den Projektor $q_j(F)\hat{P}_j(F)$ an, so wird die Gleichung zu

$$0 = q_j(F)\hat{P}_j(F)(v_1 + \cdots + v_k) = q_j(F)\hat{P}_j(F)(v_j) = v_j.$$

Dies gilt für alle j, also haben wir die direkte Summenzerlegung gezeigt.

Betrachte nun $F_j := F|V_j \in \operatorname{End}(V_j)$. Nach Voraussetzung gilt $p_j(F_j)^{r_j} = 0$, also teilt das Minimalpolynom von F_j das Polynom $p_j^{r_j}$, also ist es selbst von der Form $p_j^{n_j}$. Das charakteristische Polynom teilt aber eine Potenz des Minimalpolynoms und muss daher auch die Form

$$P_{F|V_j} = \pm p_j^{m_j}$$

haben, für noch zu bestimmende m_j. Aus der direkten Summenzerlegung von F folgt aber

$$P_F = P_{F|V_1} \cdots P_{F|V_k},$$

und wegen der Eindeutigkeit der Primfaktorzerlegung liest man sofort ab dass $m_j = r_j$. □

Beweis von Theorem 3.6. Wir wählen die Basis als Vereinigung von Basen der V_j aus dem vorherigen Theorem. Wir können uns also zur Vereinfachung der Notation gleich auf den Fall einschränken, dass $P_F = \pm p^r$ nur einen irreduziblen Faktor hat, also $k = 1$. Definiere dann

$$U_j := \operatorname{Ker} p(F)^j.$$

Wir haben dann eine aufsteigende Kette von Untervektorräumen

$$\{0\} = U_0 \subset U_1 \subset \cdots \subset U_{r-1} \subset U_r = V = U_{r+1} = U_{r+2} = \cdots$$

Definiere ausserdem

$$W_j = U_j / U_{j-1}.$$

Da die Vektorräume U_j alle F-invariant sind induziert F wieder Endomorphismen von den W_j, die wir mit $\bar{F}$ bezeichnen. (Diese sind definiert durch $\bar{F}(v + U_{j-1}) := F(v) + U_{j-1}$.) Es gilt $p(\bar{F}) = 0$, denn

$p(F)(U_j) \subset U_{j-1}$ per Definition der U_j. Ausserdem induziert aus dem gleichen Grund die Abbildung $p(F) \in \mathrm{End}(V)$ Abbildungen

$$\overline{p(F)} : W_j \to W_{j-1}.$$

Wir haben also eine Kette von linearen Abbildungen

$$\cdots \to W_{j+1} \xrightarrow{\overline{p(F)}} W_j \xrightarrow{\overline{p(F)}} W_{j-1} \to \cdots \tag{16}$$

Die Abbildungen $\overline{p(F)}$ sind ausserdem alle injektiv. (ÜA)

Nun kommt der entscheidende Schritt: Definiere den Körper (s. Lemma 3.5)

$$K_p := K[t]/pK[t] \supset K.$$

Die K-Vektorraumstruktur auf W_j kann man nun fortsetzen auf eine K_p-Vektorraumstruktur. Dabei ist die Multiplikation mit $f + pK[t] \in K_p$ auf $w = v + U_{j-1} \in W_j$ definiert durch

$$(f + pK[t])w = f(\bar{F})(w) = f(F)(v) + U_{j-1}.$$

Dies ist wohldefiniert wegen $p(\bar{F}) = 0$ wie gezeigt. (ÜA: Man zeige die K_p-Vektorraumaxiome). Ausserdem sind die Abbildungen $\overline{p(F)}$ offensichtlich K_p-linear, so dass (16) eine Kette von injektiven K_p-linearen Abbildungen von K_p-Vektorräumen ist. Wir wählen nun K_p-Basen der W_j wie folgt. Sei zunächst $(w_1, \ldots, w_{s_r})$ eine (K_p-)Basis von W_r. Dann ist die Familie $(\overline{p(F)}(w_1^r), \ldots, \overline{p(F)}(w_{s_r}^r))$ in W_{r-1} linear unabhängig wegen der Injektivität von $\overline{p(F)}$. Wir können sie also zu einer Basis

$$(\overline{p(F)}(w_1^r), \ldots, \overline{p(F)}(w_{s_r}^r), w_1^{r-1}, \ldots, w_{s_r}^r)$$

von W_{r-1} fortsetzen. Diese bilden wir mit $\overline{p(F)}$ nach W_{r-2} ab und setzten sie dort wieder zu einer Basis fort, usw. Die zur Fortsetzung zur Basis von W_j benötigten Vektoren nennen wir dabei $w_1^j, \ldots, w_{s_j}^j$, es gibt also s_j viele. Es gilt dann offensichtlich

$$s_j = \dim_{K_p} W_j - \dim_{K_p} W_{j-1}. \tag{17}$$

Man kann nun aus unserer K_p-Basis auch eine K-Basis ablesen. Betrachte z.B. zwecks Vereinfachung der Notation zuerst W_r mit K_p-Basis $w_1^r, \ldots, w_{s_r}^r$. Beachte auch, dass K_p ein $\ell := \deg p$-dimensionaler K-Vektorraum ist mit Basis den Äquivalenzklassen von $1, t, \ldots, t^{\ell-1}$. Dann hat also jedes Element $w \in W_r$ eine eindeutige Darstellung der Form

$$w = f_1(\bar{F})(w_1^r) + \cdots + f_r(\bar{F})(w_{s_r}^r) = \sum_{i=1}^{s_r} \sum_{n=0}^{\ell-1} a_{in} \bar{F}^n(w_i^r)$$

wobei die $f_i = \sum_{n=0}^{\ell-1} a_{in} t^n \in K[t]$ Polynome vom Grad $< \ell$ sind mit Koeffizienten $a_{in} \in K$. Man sieht also, dass jedes w wie oben eindeutig als K-Linearkombination der Vektoren $\bar{F}^n(w_i^r)$ geschrieben werden kann, diese bilden also eine K-Basis von W_r. Dies geht analog für die anderen W_j. Zusammengefasst: Aus der K_p-Basis erhält man eine K-Basis durch wirken mit $1, t, \ldots, t^{\ell-1}$, also in unserem Fall durch Anwenden von $1, \bar{F}, \ldots, \bar{F}^{\ell-1}$, und insbesondere gilt

$$\dim_K W_j = \ell \dim_{K_p} W_j.$$

Nun konstruieren wir unsere gesucht K-Basis $\mathcal{B}$ von V. Seien nun $v_i^j \in U_{j+1}$ Repräsentanten von den w_i^j oben, also

$$w_i^j = v_i^j + U_{j-1}.$$

Dann besteht $\mathcal{B}$ aus den Vektoren

$$F^\alpha(p(F)^\beta(v_i^j))$$

mit $j = 1, \ldots, r$, $i = 1, \ldots, s_j$, $\alpha = 0, \ldots, \ell - 1$ und $\beta = 0, \ldots, j - 1$. Die Äquivalenzklassen dieser Vektoren ergeben gerade die vorher konstruierte Basis von

$$W_1 \oplus \cdots \oplus W_r \cong V.$$

Man sieht wie im "klassischen" Fall (Beweis von Theorem 4.6.5 im Fischer), dass $\mathcal{B}$ tatsächlich eine Basis von V ist. (ÜA)

Wir prüfen, dass (bei geeigneter Ordnung der Elemente) die Matrix bzgl. dieser Basis die angegebene Form hat. Dazu wenden wir F auf unsere Basisvektoren an und berechnen die Bilder. Sei dazu noch $p = t^\ell + a_{\ell-1} t^{\ell-1} + \cdots + a_0$.

- Für $\alpha < \ell - 1$ gilt einfach
$$F(F^\alpha(p(F)^\beta(v_i^j))) = F^{\alpha+1}(p(F)^\beta(v_i^j))$$
und dies ist wieder ein Basisvektor.
- Für $\alpha = \ell - 1$ gilt
$$F(F^{\ell-1}(p(F)^\beta(v_i^j))) = F^\ell(p(F)^\beta(v_i^j)) = (p(F) - \sum_{n=0}^{\ell-1} a_n F^n)(p(F)^\beta(v_i^j))$$
$$= p(F)^{\beta+1}(v_i^j) - \sum_{n=0}^{\ell-1} a_n F^n(p(F)^\beta(v_i^j)),$$
und die Summe rechts ist wieder eine Linearkombination von Basisvektoren. (Für $\beta = j - 1$ ist der erste Term kein Basisvektor, aber dann $= 0$ wegen $p(F)^j(v_i^j) = 0$, denn $v_i^j \in U_j = \mathrm{Ker}\, p(F)^j$.)

Übersetzt man die Formeln oben zurück in Matrizenschreibweise, so erhält man gerade eine Blockdiagonalmatrix mit Jordanblöcken der Form $\tilde{J}_j(B_p)$ auf der Diagonalen. Genauer erhält man gerade s_j solcher Blöcke der Grösse $j\ell \times j\ell$. Wir müssen dann nur noch die Formel (13) prüfen. Wir bekommen aber aus (17) die Formel
$$\ell s_j = \ell \left(\dim_{K_p} W_j - \dim_{K_p} W_{j-1}\right) = \dim_K W_j - \dim_K W_{j-1}.$$
Mit $\dim W_i = \dim U_i - \dim U_{i-1}$ bekommt man daraus genau wie in Abschnitt 1 dass
$$\ell s_j = 2\dim_K U_j - \dim_K U_{j+1} - \dim_K U_{j-1},$$
also gerade (13).

□