Satz 3.9. *Sei $\alpha \in \mathrm{End}(V)$ mit $\mu_\alpha = p^m$ mit irreduziblem $p \in K[X]$ und $d := \mathrm{Grad}(p)$.*

1. *In der Situation von Satz 3.8 mit $\mu_\alpha = \chi_\alpha$ sind sämtliche α-invariante Teilräume von V gegeben durch $V_i = \langle p(\alpha)^i(v)\rangle_\alpha = \mathrm{Kern}\, p(\alpha)^{m-i}$ für jeden zyklischen Erzeuger v von V. Es gilt $n := \dim V = md$.*

2. *Ist v ein zyklischer Erzeuger von V, so ist $B = (b_1, \ldots, b_n) \in V^n$ definiert durch $b_1 = v$ und*
$$b_{di+r} := \alpha^{r-1}p(\alpha)^i(v) \qquad 1 \le r \le d, 0 \le i < n/d.$$
eine Basis von V, wobei mit $\nu = p(\alpha)$ gilt $B =$
$$(\underbrace{v, \alpha(v), \ldots, \alpha^{d-1}(v)}_{b_1,\ldots,b_d}, \underbrace{\nu(v), \alpha(\nu(v)), \ldots, \alpha^{d-1}(\nu(v)}_{b_{d+1},\ldots,b_{2d}}, \ldots, \underbrace{\nu^{m-1}(v), \ldots, \alpha^{d-1}(\nu^{m-1}(v)))}_{b_{(n-d)+1},\ldots,b_n}$$
und es gilt
$${}^B\alpha^B = J(p^m),$$
wobei $J(p^m)$ der verallgemeinerte JORDAN-Block von p^m ist, definiert durch
$$J(p^m) := \begin{pmatrix} M_p & 0 & 0 & \ldots & 0 & 0 \\ N_d & M_p & 0 & \ldots & 0 & 0 \\ 0 & N_d & M_p & \ldots & 0 & 0 \\ 0 & 0 & N_d & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & N_d & M_p \end{pmatrix} \in K^{md\times md}$$
mit $0 \in K^{d\times d}$ die Nullmatrix vom Grad d und
$$N_d := \underline{d} \times \underline{d} \to K : (i,j) \mapsto \begin{cases} 1 & \text{falls } (i,j) = (1,d) \\ 0 & \text{sonst} \end{cases}.$$
Man beachte $J_m(a) = J((X-a)^m)$.

Beweis. 1) Ist $W \le V$ ein α-invarianter Teilraum, so gibt es ein größtes i mit $W \le V_i$. Aber für jedes $w \in W - V_{i+1}$ ist $\langle w\rangle_\alpha$ ein α-invarianter Teilraum von V, der in V_i enthalten ist, und α induziert auf V_i/V_{i+1} einen Endomorphismus mit Minimalpolynom p. Andererseits ist das Minimalpolynom von $w + V_{i+1}$ ein Teiler von p, und damit gleich p. Daher ist $(w + V_{i+1}, \alpha(w) + V_{i+1}, \ldots, \alpha^{d-1}(w) + V_{i+1})$ linear unabhängig, d.h. $\dim(W/V_{i+1}) = d = \dim(V_i/V_{i+1})$, also $\langle w\rangle_\alpha = V_i$.
Wir zeigen nun, dass $V_i = \mathrm{Kern}\, p(\alpha)^{m-i}$. Wir sehen sofort, dass $p^i(\alpha)(v) \in \mathrm{Kern}(p(\alpha)^{m-i})$, da $p(\alpha)^{m-i}(p^i(\alpha)(v)) = p^m(\alpha)(v) = 0$. Also $V_i \le \mathrm{Kern}(p(\alpha)^{m-i})$. Weiter ist offensichtlich $\mathrm{Kern}\, p(\alpha)^{m-i}$ ein α-invarianter Teilraum von V. Da nach dem Homomorphiesatz $\dim(V/\mathrm{Kern}\, p(\alpha)^{m-i}) = \dim(\mathrm{Bild}(p(\alpha))^{m-i}) = \dim(V_{m-i})$ ist $\dim(\mathrm{Kern}\, p(\alpha)^{m-i}) = md - \dim(V_{m-i}) = md - (m-i)d = id = \dim(V_i)$.
2) Die $b_{r+di} + V_{i+1}$ für festes i liefern eine Basis von V_i/V_{i+1}, die gerade die Begleitmatrix von p als Matrix liefert. Damit folgt die Basiseigenschaft von B sehr leicht. Solange $r \ne d$

ist, haben wir wegen $b_{r+1+di} = \alpha(b_{r+di})$ keine weiteren Einträge in der Spalte $r + di$. Für den Rest beachte man

$$\alpha(b_d) = \alpha^d(b_1) = \underbrace{p(\alpha)(b_1)}_{b_{d+1}} - a_0 b_1 - a_1 b_2 - \cdots - a_{d-1} b_d,$$

wobei $p = X^d + a_{d-1}X^{d-1} + \cdots + a_0$ ist. Durch Heranmultiplizieren von $p(\alpha)^i$ ergeben sich schließlich die restlichen Spalten der Matrix. Übung: Zeige die lineare Unabhängigkeit der b_i etwas genauer. q. e. d.

Wir halten fest, die Basis, die uns die gewünschte Basis von V liefert, kann aus jedem Element $b_1 \in V - \operatorname{Bild} p(\alpha)$ durch Anwendung der Basis

$$B_{p,m} := (1, X, X^2, \ldots, X^{d-1}, p, Xp, \ldots, X^{d-1}p, \ldots, p^{m-1}, Xp^{m-1}, \ldots, X^{d-1}p^{m-1})$$

von $K[X]_{\mathrm{Grad}<dm}$ erhalten werden. Dabei heißt Anwenden, dass man für X den Endomorphismus α einsetzt und dann den dadurch erhaltenen Endomorphismus anwendet.

Der Schlüssel zur Normalform der Matrizen liegt im folgenden Lemma.

Lemma 3.10. *Sei $\alpha \in \mathrm{End}(V)$ mit $\mu_\alpha = p^m$ mit irreduziblem $p \in K[X]$. Ist $v \in V$ mit Minimalpolynom $\mu_{\alpha,v} = p^m$ und $w \in V$ beliebig, so existiert ein $w_1 \in \langle v\rangle_\alpha + \langle w\rangle_\alpha$ mit*

$$\langle v\rangle_\alpha + \langle w\rangle_\alpha = \langle v\rangle_\alpha \oplus \langle w_1\rangle_\alpha$$

Beweis. Sei $\mu_{\alpha,w} = p^r$. Sei $W = \langle w\rangle_\alpha$. Dann ist $\mu_{\alpha_{|W}} = \chi_{\alpha_{|W}}$ also sind die einzigen α-invarianten Teilräume von W die Räume $W_i = \mathrm{Bild}(p^i(\alpha_{|W}))$ mit $\{0\} = W_r \leq W_{r-1} \leq \cdots \leq W_1 \leq W$. Also liegt genau dann keine direkte Summe vor, wenn $S = \langle w\rangle_\alpha \cap \langle v\rangle_\alpha \neq \{0\}$. Da dieser Teilraum aber ein α-invarianter Teilraum von W ist, liegt genau dann keine direkte Summe vor, wenn $p^{r-1}(\alpha)(\langle w\rangle_\alpha)$ in S enhalten ist. Da aber ebenfalls $\langle v\rangle_\alpha$ nur α-invariante Teilräume der Form $p(\alpha)^j(\langle v\rangle_\alpha)$ hat, ist dies der Fall genau dann, wenn $p(\alpha)^{m-1}(\langle v\rangle_\alpha)$ in S enthalten ist, also $p^{r-1}(\alpha)(\langle w\rangle_\alpha) = p(\alpha)^{m-1}(\langle v\rangle_\alpha)$ ist.
Genau in diesem Fall finden wir (durch Lösen eines linearen Gleichungssystems) ein $q \in K[X]$, welches kleineren Grad als p hat, so dass $q(\alpha)p(\alpha)^{m-1}(v) = p(\alpha)^{r-1}(w)$. Da $r \leq m$ gilt, können wir w durch $w' := w - q(\alpha)p(\alpha)^{m-r}(v)$ ersetzen. Für w' ist das Minimalpolynom gleich p^s mit $s < r$. Offenbar gilt $\langle v\rangle_\alpha + \langle w\rangle_\alpha = \langle v\rangle_\alpha + \langle w'\rangle_\alpha$. Nach höchstens r derartigen Schritten liegt eine direkte Summe vor. q. e. d.

Wir wollen uns ein numerisches Beispiel anschauen.

***Beispiel* 3.11.** Sei $K := \mathbb{Q}$ und

$$A := \begin{pmatrix} -6 & 6 & 0 & 6 \\ -4 & 6 & -1 & 3 \\ -6 & 12 & -3 & 3 \\ -6 & 12 & -3 & 3 \end{pmatrix} \in \mathbb{Q}^{4\times 4}$$

mit Minimalpolynom $\mu_A(X) = X^3$. Also $p = X$ und $p(A) = A$ in der obigen Notation. Der erste Standardbasisvektor e_1 hat X^3 als Minimalpolynom (siehe Rechnung). Nun versuchen wir V als direkte Summe von zyklischen Teilräumen zu schreiben. Zuerst wollen wir $\langle e_1\rangle_\alpha + \langle e_2\rangle_\alpha$ als direkte Summe schreiben.

Auch der zweite Standardbasisvektor e_2 hat X^3 als Minimalpolynom.

$$(e_1, Ae_1, A^2e_1) = \begin{pmatrix} 1 & -6 & -24 \\ 0 & -4 & -12 \\ 0 & -6 & -12 \\ 0 & -6 & -12 \end{pmatrix}, \quad (e_2, Ae_2, A^2e_2) = \begin{pmatrix} 0 & 6 & 72 \\ 1 & 6 & 36 \\ 0 & 12 & 36 \\ 0 & 12 & 36 \end{pmatrix}.$$

Nun wissen wir, dass falls $\langle e_1\rangle_\alpha \cap \langle e_2\rangle_\alpha \neq \{0\}$, so der Schnitt ein nicht-trivialer A-invarianter Teilraum, also ist $\langle p^2e_1\rangle_A = \langle p^2e_2\rangle_A$, da dies die beiden kleinsten nicht-trivialen A-invarianten Unterräume von $\langle e_1\rangle_\alpha$ und $\langle e_2\rangle_\alpha$ sind.

Wir lesen die lineare Abhängigkeit zwischen den letzten beiden Spalten ab, also $3A^2e_1 + 1A^2e_2 = 0$. In der Notation des Beweises des letzten Lemmas ist $m = r = 3$ und für $q(A) = -3$ ist $q(A)p(A)^2e_1 = p^2(A)e_2$. Also definieren wir $e_2' := e_2 - q(A)p^{3-3}(A)(e_1) = e_2 + 3e_1$. Nun müssen wir $\langle e_1\rangle_A + \langle e_2'\rangle_A$ als direkte Summe schreiben. Es ist

$$(e_2', Ae_2') = \begin{pmatrix} 3 & -12 \\ 1 & -6 \\ 0 & -6 \\ 0 & -6 \end{pmatrix}$$

Wir haben wieder eine lineare Abhängigkeit zwischen den letzten Spalten, $-\frac{1}{2}A^2e_1 + Ae_2' = 0$, wobei nun $m = 3$, $r = 2$ und $q(A) = -\frac{1}{2}$. Also ersetzen wir demgemäß e_2' durch $f := e_2' - q(A)p^{m-r}(A)e_1 = e_2' + \frac{1}{2}Ae_1$. Dieses hat Minimalpolynom X und ist aber linear abhängig von A^2e_1, so dass wir gezeigt haben

$$\langle e_1\rangle_A + \langle e_2\rangle_A = \langle e_1\rangle_A.$$

Um ganz V zu erzeugen, werden wir das ganze Spiel nun mit e_3 statt e_2 wiederholen um $\langle e_1\rangle_A + \langle e_3\rangle_A$ als direkte Summe zu schreiben und wir kommen dann nach zwei Schritten wirklich zu einem f' mit

$$\langle e_1\rangle_A + \langle e_3\rangle_A = \langle e_1\rangle_A \oplus \langle f'\rangle_A = \mathbb{Q}^{4\times 1}.$$

In der Tat ist (e_1, Ae_1, A^2e_1, f') eine Basis von $\mathbb{Q}^{4\times 1}$, bezüglich der φ_A die Matrix $\mathrm{Diag}(M_{X^3}, M_X)$ bekommt.

Unsere Aufgabe ist es, diese Normierung auf den Fall von mehr als zwei Summanden zu übertragen.

Satz 3.12. *Sei $\alpha \in \mathrm{End}(V)$ mit $\mu_\alpha = p^m$ mit irreduziblem $p \in K[X]$. Dann existiert ein $k \in \mathbb{N}$ und Vektoren $v_1, \ldots, v_k \in V$ mit*

$$V = \bigoplus_{i=1}^{k} \langle v_i \rangle_\alpha$$

Beweis. Unser Beweis ist gleichzeitig ein Algorithmus, wie man diese Zerlegung herstellen kann. Offenbar existieren $w_1, \ldots, w_n \in V$ mit

$$V = \langle w_1 \rangle_\alpha + \cdots + \langle w_n \rangle_\alpha.$$

Wir nennen das kleinste r mit $p(\alpha)^r(w_i) = 0$ die α-Länge $\lambda(w_i)$ von w_i. Ist d der Grad von p, so liegt offenbar genau dann eine direkte Summe vor, wenn

$$\sum_i \lambda(w_i) = \mathrm{Dim}(V)/d.$$

Im allgemeinen wird die Summe also größer als $\mathrm{Dim}(V)/d$ sein, was nach unseren Vorbemerkungen sich zu einer nicht trivialen $K[X]/pK[X]$-linearen Abhängigkeit von $(p(\alpha)^{\lambda(w_1)-1}(w_1), \ldots, p(\alpha)^{\lambda(w_n)-1}(w_n)) \in (\mathrm{Kern}(p(\alpha)))^n$ führt, sagen wir

$$\sum q_i(\alpha) p(\alpha)^{\lambda(w_i)-1}(w_i) = 0$$

mit $q_i(X) \in K[X]$ vom Grad kleiner d. (Zu dieser Relation kommt man z.B., indem man auf eine allgemeine Relation $\sum_i r_i(\alpha)(w_i)$ eine geeignete Potenz von $p(\alpha)$ anwendet.) Sei

$$r := \min\{\lambda(w_i) | q_i(X) \neq 0\}$$

und j eine Stelle, wo das Minimum angenommen wird, also $r = \lambda(w_j), q_j \neq 0$. Dann setzen wir $w_i' := w_i$ für $i \neq j$ und

$$w_j' := \sum_{i=1}^{k} q_i(\alpha) p(\alpha)^{\lambda(w_i)-r}(w_i).$$

Wir sehen, dass der Koeffizient von w_j in w_j' gerade $q_j(\alpha)$ ist. Da sich $\overline{q_j} \in K[X]/pK[X]$ invertieren läßt, ist klar, dass nach wie vor gilt

$$V = \langle w_1' \rangle_\alpha + \cdots + \langle w_n' \rangle_\alpha.$$

Wir zeigen nun, dass $\lambda(w_j)' < r = \lambda(w_j)$. Denn

$$p(\alpha)^{r-1}(w_j') = \sum_{i=1}^{k} q_i(\alpha) p(\alpha)^{\lambda(w_i)-r+r-1}(w_i) = \sum_{i=1}^{k} q_i(\alpha) p(\alpha)^{\lambda(w_i)-1}(w_i) = 0,$$

aufgrund der obigen linearen Abhängigkeit.

Aber die Summe der α-Längen ist mindestens um Eins kleiner geworden. Dieses Verfahren terminiert nach endlich vielen Schritten mit einer direkten Summenzerlegung. q. e. d.