9. Die Jordansche Normalform.

Etwas verkürzt ausgedrückt geht es in diesem Kapitel darum, zu einem $L \in \mathrm{End}(V)$, $\dim V = n < \infty$, eine Basis $\mathcal{G}$ zu finden, in der $\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(L)$ eine möglichst einfache Form hat.

(9.1) Def.:

(i) $L \in \mathrm{End}(V)$ heißt diagonalisierbar, falls es eine Basis $\mathcal{G} = (v_1, \ldots, v_n)$ von V aus Eigenvektoren v_i von L gibt.
(ii) $A \in K^{n \times n}$ heißt diagonalisierbar, falls es ein $B \in \mathrm{GL}_n(K)$ gibt, so daß $B^{-1}AB$ eine Diagonalmatrix ist (d.h. so daß $(B^{-1}AB)_{ij} = 0$ für $i \neq j$).

Bem.: 1) Ist $\mathcal{G} = (v_1, \ldots, v_n)$ Basis aus Eigenvektoren von L, so gilt

$$\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(L) = \begin{pmatrix} \lambda_1 & & & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

wobei λ_i der EW von L zum EV v_i ist.

2) Ist $L \in \mathrm{End}(V), \tilde{\mathcal{G}}$ beliebige Basis von V und $A := \mathrm{Mat}_{\tilde{\mathcal{G}}}^{\tilde{\mathcal{G}}}(L)$, so gilt:

$$L \text{ diagonalisierbar} \Leftrightarrow A \text{ diagonalisierbar}$$

(vgl. (4.19)).

Bsp.: 1) Ist $A \in \mathbb{R}^{n \times n}$ symmetrisch, d.h. $A = A^T$, so ist A diagonalisierbar, vgl. den Beweis von (6.11) (Hauptachsentransformation).

2) Ist $V, \langle , \rangle$ euklidischer Vektorraum und ist $L \in \mathrm{End}(V)$ selbstadjungiert ($\Leftrightarrow \forall v, w \in V$: $\langle L(v), w \rangle = \langle v, L(w) \rangle$), so ist L diagonalisierbar, vgl. (6.11).

Diagonalisierbarkeit einer Matrix $A \in K^{n \times n}$ oder eines $L \in \mathrm{End}(V)$ ist die beste Lösung unseres Problems, die wir uns erhoffen können. Wir werden uns zunächst mit diesem einfachen (aber häufigen) Fall beschäftigen.

(9.2) Fakt. Ist $L \in \mathrm{End}(V)$ und sind $v_1, \ldots, v_k$ EVen zu <u>verschiedenen</u> EWen $\lambda_1, \ldots, \lambda_k$ von L, so sind die $v_1, \ldots, v_k$ linear unabhängig.

Bew.: Durch Induktion nach k.

Induktionsanfang: $k = 1$. Da v_1 EV ist, gilt $v_1 \neq 0$, d.h. v_1 ist linear unabhängig.

Induktionsschritt: Wir können voraussetzen, daß $v_1, \ldots, v_{k-1}$ linear unabhängig sind. Sei $\sum_{i=1}^{k} a_i v_i = 0$. Dann gilt

$$0 = L\left(\sum_{i=1}^{k} a_i v_i\right) = \sum_{i=1}^{k} a_i L(v_i) = \sum_{i=1}^{k} a_i \lambda_i v_i$$

und

$$0 = \lambda_k \left(\sum_{i=1}^{k} a_i v_i\right) = \sum_{i=1}^{k} a_i \lambda_k v_i.$$

Subtraktion der vorangehenden Gleichungen ergibt

$$\sum_{i=1}^{k-1} a_i(\lambda_i - \lambda_k) v_i = 0.$$

Da $v_1, \ldots, v_{k-1}$ linear unabhängig sind, folgt

$$a_i(\lambda_i - \lambda_k) = 0$$

für $1 \leq i \leq k-1$ und wegen $\lambda_i \neq \lambda_k$ auch $a_i = 0$ für $1 \leq i \leq k-1$. Damit reduziert sich $\sum_{i=1}^{k} a_i v_i = 0$ auf $a_k v_k = 0$. Wegen $v_k \neq 0$ folgt auch $a_k = 0$.

(9.3) Folgerung. Seien $\dim V = n$, $L \in \mathrm{End}(V)$. Besitzt L n verschiedene EWe, so ist L diagonalisierbar.

Bem.: L besitzt genau dann n verschiedene EWe, wenn P_L n verschiedene Nullstellen besitzt.

Bew.: Seien $v_1, \ldots, v_n$ EVen zu den verschiedenen EWen von L. Dann zeigt (9.2), daß $\mathcal{G} = (v_1, \ldots, v_n)$ eine Basis aus EVen ist.

(9.4) Def.: Ist λ EW von $L \in \mathrm{End}(V)$, so heißt

$$E(\lambda) = \ker(L - \lambda \mathrm{id}_V) = \{v \in V \mid L(v) = \lambda v\}$$

der Eigenraum von L zum EW λ.

Bem.: $E(\lambda)$ ist Untervektorraum von V und $E(\lambda) \setminus \{0\}$ ist genau die Menge der EVen von L zum EW λ.

(9.5) Satz (Diagonalisierbarkeitskriterium). *Sei $L \in \mathrm{End}(V)$. L ist genau dann diagonalisierbar, wenn*

$$\dim V = \sum_{i=1}^{k} \dim E(\lambda_i)$$

gilt, wobei $\lambda_1, \ldots, \lambda_k$ die verschiedenen EWe von L bezeichnen.

Bew.: Setze $d_i := \dim E(\lambda_i)$. Aus (9.2) folgt, daß $E(\lambda_1) \oplus \ldots \oplus E(\lambda_k)$ eine direkte Summe ist, also $\sum\limits_{i=1}^{k} d_i \leq n$.

„$\Rightarrow$" Ist $\mathcal{G} = (v_1, \ldots, v_n)$ Basis aus EVen, so gilt offenbar auch $\sum d_i \geq n$, also $\sum d_i = n$.

„$\Leftarrow$" Gilt $\sum\limits_{i=1}^{k} d_i = n$, so $\dim(E(\lambda_1) \oplus \ldots \oplus E(\lambda_k)) = n$, also $E(\lambda_1) \oplus \ldots \oplus E(\lambda_k) = V$. Also ist die (disjunkte!) Vereinigung von Basen der Eigenräume $E(\lambda_i), 1 \leq i \leq k$, eine Basis von V.

Bem.: Wenn man (9.3) oder (9.5) auf einen konkreten Endomorphismus L anwenden will, muß man seine EWe, d.h. die Nullstellen von P_L bestimmen. Das wird oft nicht explizit, sondern nur approximativ (mit Hilfe der Numerik) möglich sein. Kennt man die EWe, so ist die Bestimmung der Eigenräume leicht. Sie sind die Lösungsmengen von homogenen linearen Gleichungssystemen $\{v \in V \mid L(v) - \lambda v = 0\}$, λ EW von L.

Es ist wichtig zu wissen, daß etwa im Fall des Körpers $K = \mathbb{C}$, „die meisten" $A \in \mathbb{C}^{n \times n}$ (bzw. die meisten Endomorphismen) diagonalisierbar sind. Auf die Frage, was „die meisten" mathematisch bedeuten soll, gibt es (mindestens) zwei (verschiedene) Antworten.

Gegeben eine Eigenschaft E, die Punkte $x \in \mathbb{R}^n$ haben können oder nicht.

1) Definition I: Die meisten $x \in \mathbb{R}^n$ haben die Eigenschaft E, falls die Menge $\{x \in \mathbb{R}^n \mid x \text{ hat Eigenschaft } E\}$ eine offene und dichte Teilmenge des $\mathbb{R}^n$ enthält.
2) Definition II: Die meisten $x \in \mathbb{R}^n$ haben die Eigenschaft E, falls die Menge $\{x \in \mathbb{R}^n \mid x \text{ hat nicht Eigenschaft } E\}$ (Lebesgue-)Maß 0 hat.

Dabei heißt eine Teilmenge A des $\mathbb{R}^n$ vom Maß 0, falls es für jedes $\epsilon > 0$ abzählbar viele Würfel $W_i \subseteq \mathbb{R}^n$ gibt, so daß $A \subseteq \bigcup\limits_{i \in \mathbb{N}} W_i$ und $\sum\limits_{i \in \mathbb{N}} \mathrm{vol}_n(W_i) \leq \epsilon$ gilt. Definition I benutzt topologische, Definition II maßtheoretische Begriffe.

Bsp.: 1) Ist E die Eigenschaft eines $x \in \mathbb{R}$, irrational zu sein, so haben im Sinn von Def. I <u>nicht</u> die meisten $x \in \mathbb{R}$ Eigenschaft E, d.h. es gilt nicht, daß die meisten $x \in \mathbb{R}$ irrational sind: Die Menge $\mathbb{R} \setminus \mathbb{Q}$ enthält überhaupt keine nichtleere offene Teilmenge von $\mathbb{R}$. Im Sinn von Definition II sind dagegen die meisten $x \in \mathbb{R}$ irrational: Ist $\epsilon > 0$ gegeben, so wählen wir eine Folge $(r_i)_{i \in \mathbb{N}}$, die alle rationalen Zahlen durchläuft, und setzen $W_i = [r_i - \frac{\epsilon}{2^{i+2}}, r_i + \frac{\epsilon}{2^{i+2}}]$. Dann gilt $\mathbb{Q} \subseteq \bigcup\limits_{i \in \mathbb{N}} W_i$ und $\sum\limits_{i \in \mathbb{N}} \mathrm{vol}_1(W_i) = \sum\limits_{i \in \mathbb{N}} \frac{\epsilon}{2^{i+1}} = \epsilon$.

2) Es ist nicht schwer, mit einer Konstruktion ähnlich der Konstruktion der Cantormenge eine Teilmenge A von $\mathbb{R}$ zu finden, so daß $\mathbb{R} \setminus A$ offen und dicht ist, aber A nicht Maß 0 hat. Dann sind im Sinn von Def. I die meisten $x \in \mathbb{R}$ nicht Elemente von A, während das im Sinn von Def. II nicht gilt.

Identifiziert man $\mathbb{C}^{n\times n}$ mit $\mathbb{R}^{(2n)\times(2n)} \simeq \mathbb{R}^{4n^2}$, so lassen sich diese Begriffe auch auf Matrizen $A \in \mathbb{C}^{n\times n}$ anwenden. Folgender Satz gilt für beide Definitionen von „die meisten".

(9.6) Satz (ohne Beweis)

(i) *Die meisten* $p \in \mathbb{C}[x]$ *mit* $\operatorname{grad} p \leq n$ *haben* n *verschiedene Nullstellen.*
(ii) *Die meisten* $A \in \mathbb{C}^{n\times n}$ *sind diagonalisierbar.*

In (i) wird $V := \{p \in \mathbb{C}[x] \mid \operatorname{grad} p \leq n\}$ mit $\mathbb{C}^{n+1} \simeq \mathbb{R}^{2(n+1)}$ identifiziert.

Beweisidee für (i): Wir wollen zeigen, daß die Menge $\{p \in V \mid p$ hat nicht n verschiedene Nullstellen$\}$ Maß 0 hat. Wir betrachten die Abbildung

$$F : \mathbb{C}^{n+1} \to V,\ F(a, \lambda_1, \ldots, \lambda_n) = a \prod_{i=1}^{n} (x - \lambda_i).$$

Der „Fundamentalsatz der Algebra" impliziert (mit (8.10) und (8.13)), daß jedes $p \in \mathbb{C}[x] \setminus \{0\}$ zerfällt. Das zeigt, daß

$$F(\mathbb{C}^{n+1}) = \{p \in V \mid \operatorname{grad} p = n\}$$

gilt, d.h. $F(\mathbb{C}^{n+1}) = V \setminus H$, wobei $H = \{p \in V \mid \operatorname{grad} p < n\}$ eine Hyperebene ist und Maß 0 hat. Die Menge

$$B := \{(a, \lambda_1, \ldots, \lambda_n) \in \mathbb{C}^{n+1} \mid \exists 1 \leq i < j \leq n : \lambda_i = \lambda_j\}$$

ist eine Vereinigung von $\frac{n(n-1)}{2}$ Hyperebenen in $\mathbb{C}^{n+1}$ und hat ebenfalls Maß 0 (in $\mathbb{C}^{n+1} \simeq \mathbb{R}^{2(n+1)}$). Wir benötigen nun den leicht zu beweisenden Satz, daß C^1-Abbildungen $\mathbb{R}^n \to \mathbb{R}^n$ Mengen vom Maß 0 auf Mengen vom Maß 0 abbilden. Wegen $\dim_{\mathbb{R}} \mathbb{C}^{n+1} = 2(n+1) = \dim_{\mathbb{R}} V$ und da F C^1 (sogar polynomial) ist, folgt, daß $F(B) \subseteq V$ Maß 0 hat. Die Mengen aller $p \in V$, die nicht n verschiedene Nullstellen haben, ist aber gerade $H \cup F(B)$, und da sowohl H als auch $F(B)$ Maß 0 haben, folgt das auch für $H \cup F(B)$. Behauptung (ii) läßt sich unter Verwendung von (i) mit ähnlichen Methoden zeigen.

Ziel dieses Exkurses war, zu demonstrieren, wie Ideen und Techniken, die in der Analysis I und II vermittelt werden, hier in der Linearen Algebra nützlich und nötig sind.

Nach (9.6) sind im Fall $K = \mathbb{C}$ Endomorphismen L bzw. quadratische Matrizen A in der Regel diagonalisierbar.

Im Rest dieses Kapitels befassen wir uns mit den nicht diagonalisierbaren Ausnahmefällen. Zwar können wir für solche L keine Basis aus EVen finden, aber doch eine Basis, in der die Matrix von L eine besonders einfache Gestalt – die Jordansche Normalform – annimmt. Die Überlegungen dazu sind weder kurz noch einfach, und Sie können die hier gewählte Darstellung auch im Buch R. Walter: Einführung in die lineare Algebra, Vieweg 1982, nachlesen.

Im folgenden sei V ein K-Vektorraum, $\dim V = n$, und $L \in \operatorname{End}(V)$.

(9.7) Def.:

(i) Ein Untervektorraum U von V heißt L-invariant, falls $L(U) \subseteq U$ gilt.
(ii) Ein L-invarianter Untervektorraum heißt L-reduzibel, falls L-invariante Untervektorräume U_1, U_2 existieren, so daß $U = U_1 \oplus U_2$ gilt.
(iii) Ein L-invarianter Untervektorraum U heißt L-irreduzibel, falls U nicht L-reduzibel ist.

Unser Ziel ist, eine Zerlegung von V in L-irreduzible Unterräume U_i zu finden, für die jedes $L|U_i$ von einer einfachen Bauart ist.

(9.8) Def.: Sei λ EW von L. Dann heißt

$$E'(\lambda) = \{v \in V \mid \exists k \in \mathbb{N} : (L - \lambda \operatorname{id}_V)^k(v) = 0\}$$

der Hauptraum (oder verallgemeinerte Eigenraum) zum EW λ.

Bem.: Es gilt $E(\lambda) = \ker(L - \lambda \operatorname{id}_V) \subseteq E'(\lambda) = \bigcup_{k \in \mathbb{N}} \ker(L - \lambda \operatorname{id}_V)^k$.

Fakt: $E'(\lambda)$ ist L-invarianter Untervektorraum.

Wir zeigen: $v \in E'(\lambda) \Rightarrow L(v) \in E'(\lambda)$. Ist $k \in \mathbb{N}$ so, daß $(L - \lambda \operatorname{id}_V)^k(v) = 0$ gilt, so folgt

$$(L - \lambda \operatorname{id}_V)^k(L(v)) = (L \circ (L - \lambda \operatorname{id}_V)^k)(v) = 0,$$

d.h. $L(v) \in E'(\lambda)$.

(9.9) Bem.: λ ist der einzige EW von $L|E'(\lambda)$.

Bew.:

(i) λ ist *EW* von $L|E'(\lambda)$: Da λ EW von L ist, existiert $0 \neq v \in V$ mit $L(v) = \lambda v$, also $(L - \lambda \operatorname{id}_V)(v) = 0$. Daraus folgt $v \in E'(\lambda)$, und das zeigt, daß λ EW von $L|E'(\lambda)$ ist.
(ii) Annahme: $\mu \neq \lambda$ ist ein weiterer EW von $L|E'(\lambda)$. Dann existiert $0 \neq w \in E'(\lambda)$ mit $L(w) = \mu w$. Daraus folgt $(L - \lambda \operatorname{id}_V)(w) = (\mu - \lambda)v$ und daraus

$$(L - \lambda \operatorname{id}_V)^k(w) = (\mu - \lambda)^k w \neq 0$$

für alle $k \in \mathbb{N}$, im Widerspruch zu $w \in E'(\lambda)$.

(9.10) Lemma. Es existiert eine kleinste Zahl $f(\lambda) \in \mathbb{N}$, genannt der Index von λ, so daß $\ker(L - \lambda \operatorname{id}_V)^{f(\lambda)} = E'(\lambda)$ gilt. Es gilt $V = E'(\lambda) \oplus \operatorname{im}(L - \lambda \operatorname{id}_V)^{f(\lambda)}$.

Bem.: Wegen $L \circ (L - \lambda \operatorname{id}_V)^{f(\lambda)} = (L - \lambda \operatorname{id}_V)^{f(\lambda)} \circ L$ ist im $(L - \lambda \operatorname{id}_V)^{f(\lambda)}$ L-invarianter Unterraum von V.

Bew.: Sei $v_1, \ldots, v_l$ Basis von $E'(\lambda)$. Dann existiert für jedes $i \in \{1, \ldots, l\}$ ein kleinstes $k_i \in \mathbb{N}$ mit $(L - \lambda \operatorname{id}_V)^{k_i}(v_i) = 0$. Sei $f(\lambda) := \max_{1 \leq i \leq l} k_i$. Dann gilt $(L - \lambda \operatorname{id}_V)^{f(\lambda)}(v_i) = 0$ für

alle $i \in \{1, \ldots, l\}$, also

$$E'(\lambda) \subseteq \ker(L - \lambda \operatorname{id}_V)^{f(\lambda)}.$$

Nach Definition von $E'(\lambda)$ folgt daraus $E'(\lambda) = \ker(L - \lambda \operatorname{id}_V)^{f(\lambda)}$, und $f(\lambda)$ ist offenbar die kleinste Zahl mit dieser Eigenschaft.

Schlielich zeigen wir, daß

$$\ker(L - \lambda \operatorname{id}_V)^{f(\lambda)} \oplus \operatorname{im}(L - \lambda \operatorname{id}_V)^{f(\lambda)} = V$$

gilt. Wegen des Dimensionssatzes (4.8) gilt

$$\dim \ker(L - \lambda \operatorname{id}_V)^{f(\lambda)} + \operatorname{im}(L - \lambda \operatorname{id}_V)^{f(\lambda)} = \dim V\,,$$

so daß sich die Behauptung auf

$$\ker(L - \lambda \operatorname{id}_V)^{f(\lambda)} \cap \operatorname{im}(L - \lambda \operatorname{id}_V)^{f(\lambda)} = \{0\}$$

reduziert. Sei also $v \in \ker(L - \lambda \operatorname{id}_V)^{f(\lambda)} \cap \operatorname{im}(L - \lambda \operatorname{id}_V)^{f(\lambda)}$. Dann existiert ein $w \in V$, so dass

$$v = (L - \lambda \operatorname{id}_V)^{f(\lambda)}(w)$$

gilt. Wegen $v \in \ker(L - \lambda \operatorname{id}_V)^{f(\lambda)}$, folgt $\ker(L - \lambda \operatorname{id}_V)^{2f(\lambda)}(w) = 0$, also $w \in E'(\lambda)$. Aus $E'(\lambda) = \ker(L - \lambda \operatorname{id}_V)^{f(\lambda)}$ folgt nun

$$v = (L - \lambda \operatorname{id}_V)^{f(\lambda)}(w) = 0\,.$$

(9.11) Lemma. Seien $\lambda_1, \ldots, \lambda_s$ verschiedene EWe von L. Dann ist die Summe der $E'(\lambda_i)$ für $1 \leq i \leq s$ direkt: $E'(\lambda_1) \oplus \ldots \oplus E'(\lambda_s)$.

Bew.: Wir zeigen durch Induktion nach s, daß für alle $i \in \{1, \ldots, l\}$ gilt:

$$E'(\lambda_i) \cap \operatorname{span}\left(\bigcup_{\substack{j=1 \\ j \neq i}}^{s} E'(\lambda_j)\right) = \{0\}.$$

Für $s = 1$ ist das offensichtlich richtig. Im Induktionsschritt können wir nach Umnumerierung der λ_j annehmen, daß $i = s$ gilt. Sei $v \in E'(\lambda_s)$, $v_j \in E'(\lambda_j)$ für $1 \leq j < s$, und es gelte

$$v = \sum_{j=1}^{s-1} v_j.$$

Wir müssen beweisen, daß $v = 0$ gilt. Nach Definition von $f(\lambda_s)$ gilt:

$$(*) \qquad 0 = (L - \lambda_s \operatorname{id}_V)^{f(\lambda_s)}(v) = \sum_{j=1}^{s-1} (L - \lambda_s \operatorname{id}_V)^{f(\lambda_s)}(v_j).$$

Da $E'(\lambda_j)$ L-invariant – und damit auch $(L - \lambda_s \operatorname{id}_V)$-invariant – ist, gilt $(L - \lambda_s \operatorname{id}_V)^{f(\lambda_s)}(v_j) \in E'(\lambda_j)$ für $1 \leq j < s$. Aus der Induktionsvoraussetzung und $(*)$ folgt nun $(L - \lambda_s \operatorname{id}_V)^{f(\lambda_s)}(v_j) = 0$ für $1 \leq j < s$ und daraus $v_j \in E'(\lambda_s) \cap E'(\lambda_j)$. Wir zeigen, daß

daraus $v_j = 0$ folgt: Sonst sei $k \in \mathbb{N}$ minimal mit $(L - \lambda_s \operatorname{id}_V)^k(v_j) = 0$. Wäre $v_j \neq 0$, so $k \geq 1$, und wir können $w_j := (L - \lambda_s \operatorname{id}_V)^{k-1}(v_j)$ betrachten. Dann gilt $0 \neq w_j \in E'(\lambda_j)$ und $(L - \lambda_s \operatorname{id}_V)(w_j) = 0$, im Widerspruch zu (9.9). Das zeigt, daß $v_j = 0$ für $1 \leq j < s$, und daraus folgt $v = 0$, wie behauptet.

(9.12) Satz (1. Zerlegungssatz).*Das charakteristische Polynom $P_L \in K[x]$ von L zerfalle. Es seien $\lambda_1, \ldots, \lambda_s$ die verschiedenen EWe von L. Dann gilt*

$$V = E'(\lambda_1) \oplus \ldots \oplus E'(\lambda_s).$$

Bew.: Nach (9.11) wissen wir schon, daß die Summe direkt ist. Wir beweisen durch Induktion nach $n = \dim V$, daß sie ganz V ist. Ist $n = 1$, so gilt $s = 1$ und $V = E(\lambda_1) = E'(\lambda_1)$.

Im Induktionsschritt verwenden wir, daß nach (9.10) ein L-invarianter Unterraum U (nämlich $\operatorname{im}(L - \lambda_1 \operatorname{id}_V)^{f(\lambda_1)}$) existiert, für den

$$(*) \qquad V = E'(\lambda_1) \oplus U$$

gilt. Wegen $\dim E'(\lambda_1) \geq 1$, gilt $\dim U = \dim V - \dim E'(\lambda_1) < n$. Wir wollen die Induktionsvoraussetzung auf $\overline{L} := L|U \in \operatorname{End}(U)$ anwenden und müssen dazu wissen, daß auch $P_{\overline{L}}$ zerfällt. Aus $(*)$ folgt, daß

$$P_L = P_{L|E'(\lambda_1)} \cdot P_{\overline{L}}$$

gilt, vgl. Blatt 3, Aufgabe 3. Nach (8.16) zerfällt auch $P_{\overline{L}}$, und (9.9) impliziert, daß $\lambda_2, \ldots, \lambda_s$ die verschiedenen Nullstellen von $P_{\overline{L}}$ sind. Die Induktionsvoraussetzung ergibt nun, daß

$$U = E'_{\overline{L}}(\lambda_2) + \ldots + E'_{\overline{L}}(\lambda_s)$$

gilt, wobei $E'_{\overline{L}}(\lambda_j) \subseteq E'(\lambda_j)$ den Hauptraum von $\overline{L}$ zum EW λ_j von $\overline{L}$ bezeichnet. Daraus folgt mit $(*)$

$$V = E'(\lambda_1) + E'(\lambda_2) + \ldots + E'(\lambda_s),$$

wie behauptet.

Unser nächstes Ziel ist die Untersuchung von $L_i := L|E'(\lambda_i) \in \operatorname{End}(E'(\lambda_i))$ für $1 \leq i \leq s$. Für $T_i := L_i - \lambda_i \operatorname{id}_{E'(\lambda_i)}$ und $f_i := f(\lambda_i)$ gilt (nach (9.10)):

$$(T_i)^{f_i} = 0.$$

(9.13) Def.: Ein $T \in \operatorname{End}(V)$ heißt <u>nilpotent</u>, falls es ein $k \in \mathbb{N}_{>0}$ gibt, so daß $T^k = 0 \in \operatorname{End}(V)$ gilt. Das kleinste solche k heißt dann der Nilpotenzgrad $g = g(T)$ von T.

(9.14) Lemma. Sei $0 \neq T \in \operatorname{End}(V)$ nilpotent. Ist $v \in V \backslash \ker(T^{g-1})$, so sind $v, T(v), \ldots, T^{g-1}(v)$ linear unabhängig und spannen einen g-dimensionalen T-invarianten Unterraum von V auf. Speziell gilt $g \leq \dim V$.

Bew.: Gilt $\sum_{i=0}^{g-1} a_i T^i(v) = 0$, so auch

$$0 = T^{g-1}\left(\sum_{i=0}^{g-1} a_i T^i(v)\right) = \sum_{i=0}^{g-1} a_i T^{g-1+i}(v).$$

Nach Definition von g gilt $T^{g-1+i}(v) = 0$ für $i \geq 1$. Also reduziert sich die vorangehende Gleichung auf $a_0 T^{g-1}(v) = 0$. Nach Voraussetzung gilt $T^{g-1}(v) \neq 0$, also $a_0 = 0$. Wendet man T^{g-2} auf $\sum_{i=0}^{g-1} a_i T^i(v) = 0$ an, so erhält man $a_0 T^{g-2}(v) + a_1 T^{g-1}(v) = 0$. Wegen $a_0 = 0$ folgt wie vorher: $a_1 = 0$. Induktiv erhalten wir durch Anwenden von $T^{g-2}, \ldots, T^0 = \mathrm{id}_V$ auf $\sum_{i=0}^{g-1} a_i T^i(v) = 0$, daß auch $a_2 = 0, \ldots, a_{g-1} = 0$ gilt. Das beweist, daß $v, T(v), \ldots, T^{g-1}(v)$ linear unabhängig sind. Die T-Invarianz von $\mathrm{span}\{v, \ldots, T^{g-1}(v)\}$ folgt aus der Tatsache, daß die Basis $\{v, \ldots, T^{g-1}(v)\}$ von $\mathrm{span}\{v, \ldots, T^{g-1}(v)\}$ durch T auf die Teilmenge $\{T(v), \ldots, T^{g-1}(v), 0\}$ von $\mathrm{span}\{v, \ldots, T^{g-1}(v)\}$ abgebildet wird.

Nilpotente Endomorphismen T mit Nilpotenzgrad $g(T) = \dim V = n$ sind leicht zu verstehen:

(9.15) Lemma. Sei $T \in \mathrm{End}(V)$ nilpotent, $g(T) = n = \dim V$. Dann ist V T-irreduzibel, und es existiert eine Basis $\mathcal{G} = (v, T(v), \ldots, T^{n-1}(v))$ von V, so daß

$$\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(T) = \begin{pmatrix} 0 & \ldots & \ldots & \ldots & 0 \\ 1 & \ddots & & & \vdots \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & 1 & 0 \end{pmatrix}$$

gilt.

Bew.: Zum Beweis der Irreduzibilität von V nehmen wir an, daß $V = U_1 \oplus U_2$ mit zwei T-invarianten Unterräumen $U_1 \neq 0$, $U_2 \neq 0$ gilt. Dann sind $T_i := T|U_i \in \mathrm{End}(U_i)$ für $i = 1, 2$ nilpotent und wegen (9.14) gilt $g(T_i) \leq \dim U_i < n$. Daraus folgt

$$g(T) = \max\{g(T_1), g(T_2)\} < n,$$

im Widerspruch zur Voraussetzung $g(T) = n$. Die restliche Behauptung von (9.15) folgt direkt aus (9.14).

Bem.: Haben wir also für einen - jetzt nicht als nilpotent vorausgesetzten - Endomorphismus $L \in \mathrm{End}(V)$ einen EW λ_i, so daß

$$T_i := (L - \lambda_i \,\mathrm{id}_V) \mid E'(\lambda_i)$$

den Nilpotenzgrad $f_i(= g(T_i)) = \dim E'(\lambda_i)$ hat, so können wir mit (9.15) eine Basis $\mathcal{G}_i$ von $E'(\lambda_i)$ finden, so daß

$$\mathrm{Mat}_{\mathcal{G}_i}^{\mathcal{G}_i}(T|E'(\lambda_i)) = \begin{pmatrix} \lambda_i & \dots & \dots & \dots & 0 \\ 1 & \ddots & & & \vdots \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & \lambda_i \end{pmatrix}$$

gilt.

Nun betrachten wir den Fall eines nilpotenten $T \in \mathrm{End}(V)$ mit $g(T) < n$.

(9.16) Lemma (2. Zerlegungssatz). Sei $T \in \mathrm{End}(V)$ nilpotent und $g = g(T) < n = \dim V$. Dann ist V T-reduzibel.

Bew.: Wegen $T^{g-1} \neq 0$ können wir ein $v \in V \setminus \ker(T^{g-1})$ wählen und dazu ein $l \in V^*$ mit $l(T^{g-1}(v)) \neq 0$. Wir definieren

$$\begin{aligned} U_1 &:= \mathrm{span}\{v, \dots, T^{g-1}(v)\} \quad \text{und} \\ U_2 &:= \ker(l) \cap \ker(l \circ T) \cap \dots \cap \ker(l \circ T^{g-1}). \end{aligned}$$

Dann ist U_1 T-invariant, und es gilt $0 < \dim U_1 = g < n$. Wir haben (9.16) bewiesen, wenn wir zeigen können, daß U_2 T-invariant ist und daß $V = U_1 \oplus U_2$ gilt.

T-Invarianz von U_2: Ist $w \in U_2$, so gilt $l(w) = 0, \dots, l(T^{g-1}(w)) = 0$, also speziell $l(T(w)) = 0, \dots, l(T^{g-2}(T(w)) = 0$ und außerdem wegen $T^g(w) = 0$: $l(T^g(w)) = l(T^{g-1}(T(w)) = 0$. Zusammen zeigen diese Gleichungen, daß $T(w) \in U_2$ gilt. Also ist U_2 T-invariant.

Zeige $U_1 \cap U_2 = \{0\}$: Sei $w \in U_1 \cap U_2$. Da $w \in U_1$ ist, existieren $a_0, \dots, a_{g-1} \in K$ mit $w = \sum_{i=0}^{g-1} a_i T^i(v)$. Da $w \in U_2$ ist, gilt $l(T^{g-1}(w))) \dots = l(w) = 0$.

Wegen $T^{g-1}(w) = a_0 T^{g-1}(v)$ gilt $0 = l(T^{g-1}(w)) = a_0 l(T^{g-1}(v))$, und wegen $l(T^{g-1}(v)) \neq 0$ folgt daraus $a_0 = 0$. Wie im Beweis von (9.14) folgt induktiv $a_1 = 0, \dots, a_{g-1} = 0$, also $w = 0$. Das beweist $U_1 \cap U_2 = \{0\}$. Es liegt also eine direkte Summe $U_1 \oplus U_2$ vor, und wir haben

$$\dim U_1 \oplus U_2 = \dim U_1 + \dim U_2.$$

Es gilt $\dim U_1 = g$ und aus (3.23) folgt $\dim U_2 \geq n - g$, also

$$\dim U_1 \oplus U_2 \geq g + (n - g) = n = \dim V.$$

Da $U_1 \oplus U_2$ Untervektorraum von V ist, folgt daraus $V = U_1 \oplus U_2$, wie behauptet.

Zusammen besagen (9.15) und (9.16) gerade, daß für ein nilpotentes $0 \neq T \in \mathrm{End}(V)$ gilt:

$$V \ T\text{-irreduzibel} \Leftrightarrow g(T) = \dim V.$$

Auf jedem T-irreduziblen Unterraum U von V ist $T|U$ also von der einfachen, in (9.15) beschriebenen Bauart. Daß eine Zerlegung in irreduzible Unterräume möglich ist, folgt direkt aus der Definition von irreduzibel:

(9.17) Lemma. Sei $L \in \mathrm{End}(V)$, $V \neq \{0\}$. Dann existieren $k \in \mathbb{N}_{>0}$ und L-irreduzible Unterräume $U_1 \neq \{0\}, \ldots, U_k \neq \{0\}$, so daß

$$V = U_1 \oplus \ldots \oplus U_k$$

gilt.

Bew.: Durch Induktion nach $\dim V = n$. Der Induktionsschritt besteht in der einfachen Bemerkung, daß V entweder L-irreduzibel ist oder eine Zerlegung $V = U_1 \oplus U_2$ mit L-invarianten Unterräumen $U_1 \neq 0$ und $U_2 \neq 0$ gestattet. Wegen $\dim U_1 < n$, $\dim U_2 < n$ ist auf diese die Induktionsvoraussetzung anwendbar.

Wir beweisen nun zunächst den Satz über die Jordansche Normalform in der Formulierung für Endomorphismen und werden das später in eine Aussage über Matrizen übersetzen.

(9.18) Satz (Jordan-Zerlegung). *Sei $L \in \mathrm{End}(V)$ und das charakteristische Polynom von L zerfalle. Sind $\lambda_1, \ldots, \lambda_s$ die verschiedenen EWe von L, so gilt*

$$L = E'(\lambda_1) \oplus \ldots \oplus E'(\lambda_s).$$

Jeder der Haupträume $E'(\lambda_t), 1 \leq t \leq s$, zerfällt in L-irreduzible Unterräume

$$E'(\lambda_t) = U_t^1 \oplus \ldots \oplus U_t^{k_t}$$

und für jedes $i \in \{1, \ldots, k_t\}$ ist $(L - \lambda_t \mathrm{id}_V)|U_t^i =: S_t^i \in \mathrm{End}(U_t^i)$ nilpotent vom Nilpotenzgrad $g(S_t^i) = \dim U_t^i$.

Bew.: Die erste Aussage ist gerade der 1. Zerlegungssatz (9.12). Wir wenden dann (9.17) auf $L|E'(\lambda_t)$ an und erhalten eine Zerlegung von $E'(\lambda_t)$ in L-irreduziblen Unterräumen, genannt $U_t^1, \ldots, U_t^{k_t}$. Nach (9.10) wissen wir, daß $(L - \lambda_t \mathrm{id}_V)|U_t^i =: S_t^i$ für jedes $t \in \{1, \ldots, s\}$ und jedes $i \in \{1, \ldots, k_t\}$ nilpotent ist. Da U_t^i L-irreduzibel ist, ist U_t^i auch S_t^i-irreduzibel. Aus dem 2. Zerlegungssatz (9.16) folgt nun, daß der Nilpotenzgrad $g(S_t^i)$ von S_t^i gleich der Dimension von U_t^i ist, d.h. daß S_t^i von dem einfachen, durch (9.15) beschriebenen Typ ist.

Nach (9.15) existiert zu jedem der L-irreduziblen Unterräume U_t^i, $1 \leq t \leq s$, $1 \leq i \leq k_t$ eine Basis $\mathcal{G}_t^i$, so daß

$$\mathrm{Mat}_{\mathcal{G}_t^i}^{\mathcal{G}_t^i}(S_t^i) = \begin{pmatrix} 0 & \ldots & \ldots & \ldots & 0 \\ 1 & \ddots & & & \vdots \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & 1 & 0 \end{pmatrix}$$

gilt. Wegen $L|U_t^i = S_t^i + \lambda_t \operatorname{id}_{U_t^i}$ folgt

$$(*) \qquad \operatorname{Mat}_{\mathcal{G}_t^i}^{\mathcal{G}_t^i}(L|U_t^i) = \begin{pmatrix} \lambda_t & \dots & \dots & \dots & 0 \\ 1 & \ddots & & & \vdots \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & \lambda_t \end{pmatrix} =: J_t^i.$$

Damit erhalten wir

(9.19) Satz (Jordansche Normalform). *Sei $L \in \operatorname{End}(V)$ und das charakteristische Polynom P_L von L zerfalle. Die verschiedenen Nullstellen von P_L seien $\lambda_1, \dots, \lambda_s$. Dann existiert eine Basis $\mathcal{G}$ von V, so daß*

$$(**) \qquad \operatorname{Mat}_{\mathcal{G}}^{\mathcal{G}}(L) = \begin{pmatrix} J_1^1 & & & & 0 \\ & \ddots & & & \\ & & \ddots & & \\ & & & \ddots & \\ 0 & & & & J_s^{k_s} \end{pmatrix}$$

gilt, wobei die in der Diagonale stehenden „Jordanblöcke“ J_t^i $(\dim U_t^i) \times (\dim U_t^i)$-Matrizen vom Typ $(*)$ sind.

Interpretiert für Matrizen $A \in K^{n \times n}$ besagt (9.19):

(9.20) Folgerung. Sei $A \in K^{n \times n}$ und das charakteristische Polynom von A zerfalle. Dann existiert ein $B \in \mathrm{GL}_n(K)$, so daß $B^{-1}AB$ die Form $(**)$ hat.

Bem.:

1) Ist $K = \mathbb{C}$, so ist die Voraussetzung „das charakteristische Polynom zerfalle“ nach dem Fundamentalsatz der Algebra stets erfüllt. Ist K ein Unterkörper von $\mathbb{C}$, z.B. $K = \mathbb{Q}$ oder $K = \mathbb{R}$, und zerfällt das charakteristische Polynom in $K[x]$ nicht, so können wir uns durch „Komplexifizierung“ der Situation helfen, d.h. ist $L \in \operatorname{End}(\mathbb{R}^n)$, so betrachten wir die $\mathbb{C}$-lineare Fortsetzung $\tilde{L} \in \operatorname{End}(\mathbb{C}^n)$ von L, vgl. (6.11) und (6.25). In der Algebra konstruiert man zu jedem Körper K einen K enthaltenden Körper $\tilde{K}$, in dem jedes Polynom zerfällt. Also kann man die Voraussetzung „das charakteristische Polynom zerfalle“ auch für beliebige Körper durch ein der Komplexifizierung ähnliches Verfahren erzwingen.
2) Wie bereits gesagt ist das Hauptproblem bei der Berechnung der Jordanschen Normalform – die Berechnung der EWe $\lambda_1, \dots, \lambda_s$ – i.a. nicht explizit lösbar. Kennt man $\lambda_1, \dots, \lambda_s$, so ist die Jordansche Normalform und eine zugehörige Basis durch Lösen von linearen Gleichungssystemen berechenbar. Dazu noch einige Hinweise: Zu berechnen sind für jedes $t \in \{1, \dots, s\}$ die Zahl k_t der L-irreduziblen Unterräume

$U_t^1, \ldots, U_t^{k_t}$, in die $E'(\lambda_t)$ zerfällt, und die Dimensionen der U_t^i. Offenbar gilt

$$\sum_{i=1}^{k_t} \dim U_t^i = \dim E'(\lambda_t) \text{ und } \sum_{t=1}^{s} \dim E'(\lambda_t) = \dim V.$$

Berechnet man P_L mit der Jordanschen Normalform, so sieht man, daß $\dim E'(\lambda_t)$ gerade die Vielfachheit $k(P_L, \lambda_t)$ der Nullstelle λ_t ist. $\dim E'(\lambda_t)$kann nach (9.8), (9.10) durch Lösen homogener linearer Gleichungen bestimmt werden. Die Zerlegung von $\dim E'(\lambda_t)$ in L-irreduzible Unterräume kann mit dem im Beweis von (9.16) (2. Zerlegungssatz) verwendeten Verfahren explizit durchgeführt werden.

3) Während eine Basis $\mathcal{G}$, in der $\mathrm{Mat}_{\mathcal{G}}^{\mathcal{G}}(L)$ Jordansche Normalform hat, nicht eindeutig durch L bestimmt ist, ist die Jordansche Normalform $(**)$ in (9.19) bis auf die Reihenfolge der Blöcke J_t^i eindeutig durch L bestimmt. Der Beweis dieser Tatsache wurde in der Vorlesung nicht durchgeführt. Er ist nicht schwierig und beruht darauf, daß man für eine nilpotente lineare Abbildung T die Anzahl der irreduziblen Unterräume einer gegebenen Dimension sukzessive aus den Zahlen $\dim(\ker T^j)$ ausrechnen kann, vgl. dazu z.B. R. Walter, Einführung in die lineare Algebra, S. 245, Satz A.

<u>Beispiel zur Berechnung der Jordanschen Normalform.</u>

Wir berechnen die Jordansche Normalform von

$$A = \begin{pmatrix} 3 & 2 & -3 \\ 4 & 10 & -12 \\ 3 & 6 & -7 \end{pmatrix}$$

$$\begin{aligned} \det(A - \lambda E_3) &= (3-\lambda)(\lambda^2 - 3\lambda + 2) + 8(\lambda - 2) - 9(\lambda - 2) \\ &= (3-\lambda)(\lambda-2)(\lambda-1) - (\lambda-2) = -(\lambda-2)(\lambda^2 - 4\lambda + 4) = -(\lambda-2)^3 \end{aligned}$$

Also zerfällt P_A und $\lambda = 2$ ist der einzige EW von A. Nach (9.10) ist $A - 2E_3 =: T$ nilpotent und wegen $A - 2E_3 \neq 0$ gilt für den Nilpotenzgrad g von $A - 2E_3$ nach (9.14)

$$2 \leq g \leq 3.$$

Wir berechnen $(A - 2E_3)^2 = 0$, also $g = 2$. Da der Nilpotenzgrad unabhängig von der Wahl der Basis ist (es gilt $(B^{-1}AB)^j = B^{-1}A^jB$ für jedes $B \in \mathrm{GL}_3(\mathbb{R})$) hat auch die Jordansche Normalform von $A - 2E_3$ Nilpotenzgrad 2. Unter den nilpotenten Jordanschen Normalformen

$$\begin{pmatrix} \boxed{0} & 0 & 0 \\ 0 & \boxed{0} & 0 \\ 0 & 0 & \boxed{0} \end{pmatrix}, \quad \begin{pmatrix} \boxed{\begin{matrix} 0 & 0 \\ 1 & 0 \end{matrix}} & \begin{matrix} 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & 0 \end{matrix} & \boxed{0} \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

hat aber nur $\left(\begin{smallmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right)$ Nilpotenzgrad 2, also ist

$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + 2E_3 = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

die Jordansche Normalform von A. Eine Matrix $B \in \mathrm{GL}_3(\mathbb{R})$, für die $B^{-1}AB = \left(\begin{smallmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 2 \end{smallmatrix}\right)$ gilt, kann man wie folgt bestimmen. Wir wählen zunächst einen beliebigen Vektor $v \notin \ker(T)$ für $T := A - 2E_3$, vgl. (9.14). Nach (9.15) ist dann $\mathcal{G}_1 := (v, Tv)$ Basis eines T-irreduziblen, T-invarianten Unterraums, genannt U_1, und $\mathrm{Mat}_{\mathcal{G}_1}^{\mathcal{G}_1}(A|U_1) = \left(\begin{smallmatrix} 2 & 0 \\ 1 & 2 \end{smallmatrix}\right)$. Wir können z.B. $v = \left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right)$ wählen. Dann ist $Tv = \left(\begin{smallmatrix} 1 \\ 4 \\ 3 \end{smallmatrix}\right)$. Für Be_3 müssen wir einen von Tv unabhängigen EV von A wählen, z.B. $Be_3 = \left(\begin{smallmatrix} 3 \\ 0 \\ 1 \end{smallmatrix}\right)$. Die Matrix B ist dann durch $Be_1 = v = \left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right)$, $Be_2 = Tv = \left(\begin{smallmatrix} 1 \\ 4 \\ 3 \end{smallmatrix}\right)$ und $Be_3 = v = \left(\begin{smallmatrix} 3 \\ 0 \\ 1 \end{smallmatrix}\right)$ gegeben, d.h. $B = \left(\begin{smallmatrix} 1 & 1 & 3 \\ 0 & 4 & 0 \\ 0 & 3 & 1 \end{smallmatrix}\right)$. Man rechnet nach, daß $B^{-1} = \left(\begin{smallmatrix} 1 & 2 & -3 \\ 0 & \frac{1}{4} & 0 \\ 0 & -\frac{3}{4} & 1 \end{smallmatrix}\right)$ und $B^{-1}AB = \left(\begin{smallmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 2 \end{smallmatrix}\right)$ gilt.

Abgesehen von der Einsicht in die möglichen Typen von linearen Abbildungen, die die Sätze (9.18) und (9.19) vermitteln, ist die Jordansche Normalform auch von praktischem Nutzen. Es kommt z.B. oft vor, daß hohe Potenzen A^m einer Matrix $A \in K^{n\times n}$ berechnet werden sollen Für allgemeine Matrizen A ist das mit sehr großem Rechenaufwand verbunden. Kann man jedoch ein $B \in \mathrm{GL}_n(K)$ finden, so daß $A = B^{-1}JB$ gilt und J Jordansche Normalform hat, so gilt $A^m = B^{-1}J^mB$, wobei J^m sehr leicht zu berechnen ist. Eine damit zusammenhängende theoretische Erkenntnis ist der

(9.21) Satz (von Cayley-Hamilton). *Setzt man $L \in \mathrm{End}(V)$ in sein charakteristisches Polynom P_L ein, so erhält man $P_L(L) = 0 \in \mathrm{End}(V)$.*

Vorbemerkung: Wenn L diagonalisierbar ist, so ist (9.21) leicht einzusehen, vgl. Blatt 7, Aufgabe 3. Nach Satz (9.6)(ii) ist die Menge der diagonalisierbaren $A \in \mathbb{C}^{n\times n}$ dicht in $\mathbb{C}^{n\times n}$. Da die Abbildung $A \in \mathbb{C}^{n\times n} \to P_A(A) \in \mathbb{C}^{n\times n}$ stetig ist und auf der dichten Menge der diagonalisierbaren Matrizen $= 0$ ist, folgt, daß $P_A(A)$ für alle $A \in \mathbb{C}^{n\times n}$ (und damit auch für alle $A \in \mathbb{R}^{n\times n}$) gilt.

Bew.: von (9.21): Wir führen den Beweis nur für den Fall durch, daß P_L zerfällt. Wie in Bem. 1 nach (9.20) gesagt wurde, ist das keine wirkliche Einschränkung. Wir verwenden (9.18). Offenbar genügt es zu zeigen, daß für jeden EW λ_t von L gilt:

$$P_L(L) \mid E'(\lambda_t) = 0.$$

Wegen (9.14) gilt für $n_t := \dim E'(\lambda_t)$:

$$(*) \qquad (L - \lambda_t \,\mathrm{id}_V \mid E'(\lambda_t))^{n_t} = 0.$$

Andererseits enthält P_L den Faktor $(x-\lambda_t)^{n_t}$. Deshalb existiert ein $Q \in \mathrm{End}(V)$ mit

$$(**) \qquad P_L(L) = Q \circ (L - \lambda_t \,\mathrm{id}_V)^{n_t}.$$

Aus $(*)$ und $(**)$ folgt $P_L(L) \mid E'(\lambda_t) = 0$, wie behauptet.

Eine weitere wichtige Anwendung der Jordanschen Normalform betrifft das explizite Lösen von linearen Differentialgleichungssystemen mit konstanten Koeffizienten. Gegeben ist dabei eine Matrix $A \in \mathbb{C}^{n\times n}$ und gesucht sind alle n-Tupel $\begin{pmatrix} z_1(t) \\ \vdots \\ z_n(t) \end{pmatrix}$ von differenzierbaren Funktionen $z_i : \mathbb{R} \to \mathbb{C}(\simeq \mathbb{R}^2)$, für die

$$(*) \qquad \begin{pmatrix} z_1'(t) \\ \vdots \\ z_n'(t) \end{pmatrix} = A \begin{pmatrix} z_1(t) \\ \vdots \\ z_n(t) \end{pmatrix}$$

für alle $t \in \mathbb{R}$ gilt. Man führt die Funktion $z = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} : \mathbb{R} \to \mathbb{C}^n$ ein und schreibt für $(*)$ kürzer:

$$z' = Az.$$

Ist $B \in \mathrm{GL}_n(\mathbb{C})$, gilt $A = B^{-1}JB$ und ist $w : \mathbb{R} \to \mathbb{C}^n$ eine Lösung von $w' = Jw$, so ist $z := B^{-1}w$ eine Lösung von $(*)$:

$$z' = B^{-1}w' = B^{-1}Jw = B^{-1}JBz = Az.$$

Es genügt also, alle Lösungen der Gleichung

$$(**) \qquad w' = Jw$$

zu kennen. Wir betrachten zunächst den (nach (9.6) typischen) Fall, daß $J = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$ eine Diagonalmatrix ist.

Dann „entkoppeln“ sich die Differentialgleichungen $(**)$ zu:

$$w_j' = \lambda_j w_j \text{ für } 1 \le j \le n$$

mit $\lambda_j \in \mathbb{C}$. Davon sind die Lösungen explizit angebbar,

$$w_j(t) = w_j(0)e^{\lambda_j t} \text{ für } 1 \le j \le n,$$

wobei die „Anfangswerte“ $w_j(0), 1 \le j \le n$, die Lösung eindeutig bestimmen.

Folgerung: Sei $A \in \mathbb{C}^{n\times n}$ diagonalisierbar, $A = B^{-1} \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} B$, und $a \in \mathbb{C}^n$. Dann existiert genau eine Lösung $z : \mathbb{R} \to \mathbb{C}^n$ von $(*)$ mit $z(0) = 0$, nämlich $z(t) = B^{-1}w(t)$, wobei $w_j(t) = c_j e^{\lambda_j t}$ für $1 \le j \le n$ und $c = Ba$.

Bem.:

1) Offenbar konvergieren die (Diagonal-)Matrizen $\sum\limits_{k=0}^{m} \frac{(tJ)^k}{k!}$ für $m \to \infty$ gegen die Matrix $\begin{pmatrix} e^{\lambda_1 t} & & 0 \\ & \ddots & \\ & & \ddots \\ 0 & & e^{\lambda_n t} \end{pmatrix} =: \exp(tJ)$.

   Entsprechend bezeichnet man $\lim\limits_{m\to\infty} \sum\limits_{k=0}^{m} \frac{(tA)^k}{k!} =: \exp(tA)$ und wegen $A^k = B^{-1}J^kB$ gilt

$$\exp(tA) = B^{-1}\exp(tJ)B.$$

   Die Lösungen $w(t)$ von $(**)$ lassen sich damit als

$$w(t) = \exp(tJ)w(0)$$

   und die Lösungen $z(t) = B^{-1}w(t)$ von $(*)$ als

$$z(t) = \exp(tA)z(0)$$

   schreiben.

2) Ist $A \in \mathbb{R}^{n\times n}$ reell diagonalisierbar, d.h. $A = B^{-1}JB$ mit $B \in \mathrm{GL}_n(\mathbb{R})$, $J = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ & & \ddots \\ 0 & & \lambda_n \end{pmatrix}$ mit $\lambda_j \in \mathbb{R}$, so erhalten wir mit demselben Verfahren und denselben Formeln reelle Lösungen $z : \mathbb{R} \to \mathbb{R}^n$, $z = B^{-1}w$ mit $w_j(t) = e^{\lambda_j t} \in \mathbb{R}$ von $z' = Az$. In vielen Fällen wird $A \in \mathbb{R}^{n\times n}$ jedoch nur komplex diagonalisierbar sein, d.h. $A = B^{-1}JB$ mit $B \in \mathrm{GL}_n(\mathbb{C})$, $J = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ & & \ddots \\ 0 & & \lambda_n \end{pmatrix}$ mit $\lambda_j \in \mathbb{C}$. Wir erhalten komplexe Lösungen $z : \mathbb{R} \to \mathbb{C}^n$ von $z' = Az$. Schreiben wir $z(t) = x(t) + iy(t)$ mit $x, y : \mathbb{R} \to \mathbb{R}^n$, so gilt auch $x' = Ax$, $y' = Ay$, d.h. $x = \mathrm{Re}z$ und $y = \mathrm{Im}z$ sind reelle Lösungen von $(*)$. Das kann man z.B. daran sehen, daß wegen $A = \overline{A}$ mit $z : \mathbb{R} \to \mathbb{C}^n$ auch $\overline{z} : \mathbb{R} \to \mathbb{C}^n$ Lösung von $(*)$ ist:

$$\overline{z}' = \overline{Az} = \overline{A}\overline{z} = A\overline{z}.$$

   Da $(*)$ linear ist, sind dann auch $x = \frac{1}{2}(z + \overline{z})$ und $y = \frac{1}{2}(z - \overline{z})$ Lösungen von $(*)$.

Als nächstes betrachten wir den Fall, daß $A \in \mathbb{C}^{n\times n}$ eine Jordansche Normalform hat, die nur ein einziges „Jordankästchen“ hat, d.h. $A = B^{-1}JB$ mit $B \in \mathrm{GL}_n(\mathbb{C})$ und

$$J = T + \lambda E_n = \begin{pmatrix} \lambda & \dots & \dots & \dots & 0 \\ 1 & \ddots & & & \vdots \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & \lambda \end{pmatrix}, \lambda \in \mathbb{C}.$$

Mit etwas probieren stellt man fest, daß man Lösungen $w : \mathbb{R} \to \mathbb{C}^n$ von $w' = Jw$ wie folgt erzeugen kann: Man wählt ein Polynom $p(t) =: \sum_{k=0}^{n-1} \frac{c_{n-k}}{k!} t^k$, $c_{n-k} = p^{(k)}(0)$, vom Grad $n-1$ und setzt

$$w(t) := e^{\lambda t} \left(p^{(n-1)}(t), p^{(n-2)}(t), \dots, p(t)\right),$$

wobei $p^{(k)}(t)$ die k'te Ableitung von $p(t)$ bezeichnet. Es gilt dann nämlich:

$$w'(t) = \lambda w(t) + e^{\lambda t} \left(0, p^{(n-1)}(t), \dots, p'(t)\right) = Jw(t).$$

Man kann auch leicht sehen, daß jede Lösung w von $w' = Jw$ von diesem Typ ist: Setzt man $q(t) := e^{-\lambda t} w(t)$, so folgt $q' = Tq$. Differenziert man die Gleichung $q' = Tq$ weiter, so erhält man $q^{(k)} = T^k q$ und wegen $T^n = 0$: $q^{(n)} = 0$. Daraus folgt: q ist ein Polynom vom Grad $\leq n-1$. Man erhält also auf diese Art explizit alle Lösungen $w : \mathbb{R} \to \mathbb{C}^n$ von $w' = Jw$ und mit $z := B^{-1}w$ alle Lösungen von $z' = Az$. Die „Anfangswerte“ $w(0) = \left(p^{(n-1)}(0), \dots, p(0)\right) = (c_1, \dots, c_n) \in \mathbb{C}^n$ sind beliebig vorgebbar und bestimmen die Lösung $w(t)$ eindeutig, und analog für $z(t) := B^{-1}w(t)$.

Im allgemeinen hat die Jordansche Normalform J von A mehrere „Jordankästchen“ $J_1^1, \dots, J_s^{k_s}$ wie in (9.19)(∗∗). Die Lösungen von $w' = Jw$ setzen sich dann aus den Lösungen zu den einzelnen Jordankästchen $\left(w_t^{k_t}\right)' = J_t^{k_t} w_t^{k_t}$, die wir nach dem vorangehenden explizit kennen, zusammen.

Bsp.: Wir betrachten das lineare Differentialgleichungssystem

$$(*) \qquad x' = Ax$$

für Vektorfunktionen $x : \mathbb{R} \to \mathbb{R}^3$ und die Matrix $A = \begin{pmatrix} 3 & 2 & -3 \\ 4 & 10 & -12 \\ 3 & 6 & -7 \end{pmatrix}$ aus dem vorangehenden Beispiel. Explizit bedeutet $(*)$ dann:

$$\begin{aligned} x_1'(t) &= 3x_1(t) + 2x_2(t) - 3x_3(t) \\ x_2'(t) &= 4x_1(t) + 10x_2(t) - 12x_3(t) \\ x_3'(t) &= 3x_1(t) + 6x_2(t) - 7x_3(t) \end{aligned}$$

Wir haben die Jordansche Normalform J von A berechnet:

$$J = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Zu dem Jordankästchen $\left(\begin{smallmatrix} 2 & 0 \\ 1 & 2 \end{smallmatrix}\right)$ gehören die Lösungen

$$\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = e^{2t} \begin{pmatrix} c_1 \\ c_2 + c_1 t \end{pmatrix}$$

und zu dem EV e_3 von J: $y_3(t) = c_3 e^{2t}$.

Die Lösungen von $y' = Jy$ sind dann $y(t) = e^{2t} \begin{pmatrix} c_1 \\ c_2 + c_1 t \\ c_3 \end{pmatrix}$ mit beliebigen $c_1, c_2, c_3 \in \mathbb{R}$.

Die Lösungen x von $(*)$ sind $x = B^{-1}y$ mit $B^{-1} = \begin{pmatrix} 1 & 2 & -3 \\ 0 & \frac{1}{4} & 0 \\ 0 & -\frac{3}{4} & 1 \end{pmatrix}$.

Bem.:

1) Auch hier kann man zeigen, daß für alle $t \in \mathbb{R}$ die Limites $\lim\limits_{m\to\infty} \sum\limits_{k=0}^{m} \frac{(tJ)^k}{k!} =: \exp(tJ)$ und $\lim\limits_{m\to\infty} \sum\limits_{k=0}^{m} \frac{(tA)^k}{k!} =: \exp(tA)$ existieren, daß $\exp(tA) = B^{-1}\exp(tJ)B$ gilt und daß die Lösungen w von $w' = Jw$ gerade durch $w(t) = \exp(tJ)w(0)$ gegeben sind. Dann folgt, daß die Lösungen von $z' = Az$ als $z(t) = \exp(tA)z(0)$ geschrieben werden können.
2) Ist $A \in \mathbb{R}^{n\times n}$, so zerfällt P_A über $\mathbb{C}$, so daß A eine komplexe Jordansche Normalform $J \in \mathbb{C}^{n\times n}$ besitzt. Ist $A = B^{-1}JB$ mit $B \in \mathrm{GL}_n(\mathbb{C})$ und sind $z = B^{-1}w$ die komplexen Lösungen von $z' = Az$, so erhält man die reellen Lösungen von $x' = Ax$ als die Realteile (oder – gleichwertig – als die Imaginärteile) der komplexen Lösungen.