VI. Eigenwerte und Eigenvektoren

Wir wollen in diesem Abschnitt lineare Abbildungen $\varphi\colon V \to V$ genauer untersuchen, wobei V einen endlich-dimensionalen Vektorraum bezeichnet.

VI.1. Diagonalisierbarkeit. Wir beginnen mit der Definition einer Klasse von Endomorphismen, die besonders leicht zu verstehen sind:

VI.1.1. DEFINITION (Diagonalisierbarkeit). Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum. Eine lineare Abbildung $\varphi\colon V \to V$ wird *diagonalisierbar* genannt, falls eine geordnete Basis B von V existiert, sodass

$$[\varphi]_{BB} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

wobei $n = \dim(V)$ und $\lambda_1, \dots, \lambda_n \in \mathbb{K}$.

Eine lineare Abbildung $\varphi\colon V \to V$ ist also genau dann diagonalisierbar, wenn eine geordnete Basis $B = (b_1, \dots, b_n)$ existiert, für die

$$\varphi(b_i) = \lambda_i b_i, \qquad i = 1, \dots, n$$

gilt, wobei $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ und $n = \dim(V)$.

VI.1.2. DEFINITION (Eigenwerte und Eigenvektoren). Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und $\varphi\colon V \to V$ linear. Ein Skalar $\lambda \in \mathbb{K}$ wird *Eigenwert* von φ genannt, falls ein Vektor $v \in V$ existiert, sodass $v \neq 0$ und

$$\varphi(v) = \lambda v.$$

In diesem Fall wird v als *Eigenvektor* zum Eigenwert λ bezeichnet, und

$$E_\lambda := \{v \in V \mid \varphi(v) = \lambda v\} = \ker(\varphi - \lambda\,\mathrm{id}_V)$$

wird der *Eigenraum* zum Eigenwert λ genannt. Unter der *geometrischen Vielfachheit* des Eigenwertes λ verstehen wir die Dimension des Eigenraums E_λ. Die Menge aller Eigenwerte wird *Spektrum* von φ genannt und mit $\sigma(\varphi)$ bezeichnet.

Eine lineare Abbildung $\varphi\colon V \to V$ ist also genau dann diagonalisierbar, wenn eine Basis von V existiert, die aus Eigenvektoren von φ besteht. Solche Basen werden manchmal auch als *Eigenbasen* bezeichnet. Zur Bestimmung der Eigenwerte haben wir:

VI.1.3. PROPOSITION. *Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und $\varphi\colon V \to V$ linear. Ein Skalar $\lambda \in \mathbb{K}$ ist genau dann Eigenwert von φ, wenn*

$$\det(\varphi - \lambda\,\mathrm{id}_V) = 0.$$

Beweis. Für $\lambda \in \mathbb{K}$ gilt:

$$\begin{aligned}\lambda \text{ ist Eigenwert von } \varphi &\Leftrightarrow \exists v \in V : \varphi(v) = \lambda v,\ v \neq 0\\ &\Leftrightarrow \ker(\varphi - \lambda \operatorname{id}_V) \neq \{0\}\\ &\Leftrightarrow \varphi - \lambda \operatorname{id}_V \colon V \to V \text{ ist nicht invertierbar}\\ &\Leftrightarrow \det(\varphi - \lambda \operatorname{id}_V) = 0\end{aligned}$$

Dabei haben wir die beiden Korollare IV.2.11 und V.4.2 verwendet. □

VI.1.4. Bemerkung. Eine lineare Abbildung $\varphi\colon V \to V$ ist genau dann invertierbar, wenn 0 nicht Eigenwert von φ ist. Beide Bedingungen an φ sind nämlich zu $\ker(\varphi) = \{0\}$ äquivalent, vgl. Korollar IV.2.11.

VI.1.5. Bemerkung. Sei $\psi\colon V \xrightarrow{\cong} W$ ein linearer Isomorphismus zwischen endlich-dimensionalen $\mathbb{K}$-Vektorräumen, und $\varphi\colon V \to V$ linear. Dann hat die lineare Abbildung $\psi \circ \varphi \circ \psi^{-1}\colon W \to W$ dieselben Eigenwerte wie φ, d.h.

$$\sigma(\psi \circ \varphi \circ \psi^{-1}) = \sigma(\varphi).$$

Ein Vektor $v \in V$ ist genau dann Eigenvektor von φ zum Eigenwert λ, wenn $\psi(v)$ Eigenvektor von $\psi \circ \varphi \circ \psi^{-1}$ zum Eigenwert λ ist, d.h. bildet $b_1, \dots, b_n$ eine Basis von V, die aus Eigenvektoren von φ besteht, dann bildet $\psi(b_1), \dots, \psi(b_n)$ eine Basis von W, die aus Eigenvektoren von $\psi \circ \varphi \circ \psi^{-1}$ besteht. Insbesondere ist φ genau dann diagonalisierbar, wenn $\psi \circ \varphi \circ \psi^{-1}$ diagonalisierbar ist.

VI.1.6. Bemerkung. Sei $\varphi\colon V \to V$ diagonalisierbar und $B = (b_1, \dots, b_n)$ eine Basis aus Eigenvektoren,

$$[\varphi]_{BB} = \begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}.$$

Ist $\sigma \in \mathfrak{S}_n$ eine Permutation und bezeichnet $\tilde{B} := (b_{\sigma(1)}, \dots, b_{\sigma(n)})$ die mit σ umgeordnete Basis, dann gilt

$$[\varphi]_{\tilde{B}\tilde{B}} = \begin{pmatrix}\lambda_{\sigma(1)} & & \\ & \ddots & \\ & & \lambda_{\sigma(n)}\end{pmatrix}.$$

Die Diagonaleinträge können daher durch Übergang zu einer geeignete Eigenbasis in beliebige Reihenfolge gebracht werden, vgl. auch Aufgabe 31.

VI.1.7. Beispiel. Seien W und W' komplementäre Teilräume eines endlich-dimensionalen Vektorraums V, d.h. $V = W \oplus W'$. Es bezeichne $\pi\colon V \to V$ die Projektion auf W längs W'. Wir wollen uns davon überzeugen, dass π diagonalisierbar ist. Seien dazu $b_1, \dots, b_k$ eine Basis von W und $b_{k+1}, \dots, b_n$ eine Basis von W'. Dann bildet $B = (b_1, \dots, b_n)$ eine Basis von W und es gilt

$$[\pi]_{BB} = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}.$$

Der Projektor π ist daher diagonalisierbar. Die einzigen möglichen Eigenwerte sind 1 und 0, denn $\det(\pi - \lambda \operatorname{id}_V) = \det \left(\begin{smallmatrix} (1-\lambda)I_k & 0 \\ 0 & -\lambda I_{n-k} \end{smallmatrix}\right) = (1-\lambda)^k(-\lambda)^{n-k}$, siehe Proposition VI.1.3. Gilt $W \neq \{0\} \neq W'$ so treten beide Eigenwerte tatsächlich auf, und die zugehörigen Eigenräume sind $E_1 = W$ sowie $E_0 = W'$.

Analog können wir die Spiegelung $\sigma \colon V \to V$ an W längs W' analysieren. In diesem Fall gilt

$$[\sigma]_{BB} = \begin{pmatrix} I_k & 0 \\ 0 & -I_{n-k} \end{pmatrix},$$

d.h. auch σ ist diagonalisierbar. Die einzigen möglichen Eigenwerte sind 1 und -1. Gilt $W \neq \{0\} \neq W'$ und $2 \neq 0 \in \mathbb{K}$ dann treten beide Eigenwerte auf, und die zugehörigen Eigenräume sind $E_1 = W$ sowie $E_{-1} = W'$.

Unter den Eigenwerten und Eigenvektoren einer quadratischen Matrix $A \in M_{n\times n}(\mathbb{K})$ verstehen wir die Eigenwerte und Eigenvektoren der damit assozierten linearen Abbildung $\mathbb{K}^n \to \mathbb{K}^n$, $x \mapsto Ax$. Ein Skalar $\lambda \in \mathbb{K}$ ist also genau dann Eigenwert von A, falls $0 \neq x \in \mathbb{K}^n$ existiert, sodass $Ax = \lambda x$, und jedes solche x ist Eigenvektor von A zum Eigenwert λ. Der Eigenraum zum Eigenwert λ ist gerade $E_\lambda = \{x \in \mathbb{K}^n : Ax = \lambda x\}$. Nach Proposition VI.1.3 ist λ genau dann Eigenwert von A, wenn $\det(A - \lambda I_n) = 0$ gilt. Die Matrix A wird diagonalisierbar genannt, falls die damit assozierte lineare Abbildung $\mathbb{K}^n \to \mathbb{K}^n$, $x \mapsto Ax$, diagonalisierbar ist. Dies ist genau dann der Fall, wenn eine Basis aus Eigenvektoren $b_1, \dots, b_n$ existiert, d.h. $Ab_i = \lambda_i b_i$. Bezeichnet $E = (e_1, \dots, e_n)$ die Standardbasis von $\mathbb{K}^n$, dann ist die Matrix $S := T_{EB} = (b_1|\cdots|b_n)$ invertierbar und es gilt

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}, \tag{VI.1}$$

denn $[\psi]_{BB} = T_{BE}[\psi]_{EE}T_{EB} = T_{EB}^{-1}AT_{EB} = S^{-1}AS$. Ist umgekehrt S eine invertierbare Matrix für die (VI.1) gilt, dann bilden die Spalten von S eine Basis aus Eigenvektoren und die Diagonaleinträge λ_i sind genau die Eigenwerte von A. Die Matrix A ist also genau dann diagonalisierbar, wenn eine invertierbare Matrix $S \in \mathrm{GL}_n(\mathbb{K})$ existiert, sodass $S^{-1}AS$ eine Diagonalmatrix ist.

VI.1.8. Definition (Ähnlichkeit von Matrizen). Zwei quadratische Matrizen $A, A' \in M_{n\times n}(\mathbb{K})$ werden *ähnlich* genannt, falls eine invertierbare Matrix $S \in \mathrm{GL}_n(\mathbb{K})$ existiert, für die $A' = SAS^{-1}$ gilt. Wir schreiben in diesem Fall $A' \sim A$.

VI.1.9. Lemma. *Ähnlichkeit definiert eine Äquivalenzrelation auf $M_{n\times n}(\mathbb{K})$.*

Beweis. Die Relation ist reflexsiv, denn $A = I_n A I_n^{-1}$. Die Relation ist symmetrisch, denn aus $A' = SAS^{-1}$ folgt $A = S^{-1}A'(S^{-1})^{-1}$. Die Relation ist auch transitiv, denn aus $A' = SAS^{-1}$ und $A'' = TA'T^{-1}$ folgt $A'' = (TS)A(TS)^{-1}$. Beachte hier, dass I_n, S^{-1} und TS invertierbar sind, falls dies für S und T gilt. □

Mit dieser Terminologie ist eine quadratische Matrix A also genau dann diagonalisierbar, wenn sie ähnlich zu einer Diagonalmatrix ist. Sind $A \sim A'$ zwei ähnliche Matrizen, dann haben A und A' die gleichen Eigenwerte, und A ist genau dann diagonalisierbar, wenn A' diagonalisierbar ist. Dies folgt aus Bemerkung VI.1.5. Ist $\varphi\colon V \to V$ linear, und sind B, C zwei Basen von V, dann gilt $[\varphi]_{BB} \sim [\varphi]_{CC}$, denn $[\varphi]_{CC} = (T_{BC})^{-1}[\varphi]_{BB}T_{BC}$. Auch hat φ dieselben Eigenwerte wie $[\varphi]_{BB}$, es gilt daher

$$\sigma(\varphi) = \sigma([\varphi]_{BB}), \tag{VI.2}$$

für jede Basis B von V. Schließlich ist die lineare Abbildung φ genau dann diagonalisierbar, wenn die Matrix $[\varphi]_{BB}$ diagonalisierbar ist, für eine (und dann jede) Basis B von V.

VI.1.10. Beispiel. Wir wollen Eigenwerte und Eigenvektoren der Matrix

$$A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \in M_{2\times 2}(\mathbb{R})$$

bestimmen. Aus Proposition VI.1.3 erhalten wir folgende Gleichung für die Eigenwerte:

$$0 = \det(A - \lambda I_2) = \det\begin{pmatrix} 3-\lambda & 1 \\ 1 & 3-\lambda \end{pmatrix} = (3-\lambda)^2 - 1 = (\lambda-2)(\lambda-4).$$

Die Matrix A hat daher genau zwei Eigenwerte, $\lambda_1 = 2$ und $\lambda_2 = 4$. Für die zugehörigen Eigenräume erhalten wir

$$E_{\lambda_1} = \ker(A - 2I_n) = \ker\left(\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right) = \langle\left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)\rangle$$
$$E_{\lambda_2} = \ker(A - 4I_n) = \ker\left(\begin{smallmatrix} -1 & 1 \\ 1 & -1 \end{smallmatrix}\right) = \langle\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)\rangle$$

Die beiden Eigenvektoren $\left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$ und $\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ bilden eine Basis von $\mathbb{R}^2$, d.h. die Matrix A ist diagonalisierbar. Fassen wir diese Eigenvektoren als Spalten einer Matrix S auf, dann gilt also:

$$\underbrace{\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}^{-1}}_{S^{-1}} \underbrace{\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}}_{S} = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$$

VI.1.11. Beispiel. Wir wollen Eigenwerte und Eigenvektoren der Matrix

$$A = \begin{pmatrix} 1 & -2 & -2 \\ 2 & 6 & 4 \\ -1 & -2 & 0 \end{pmatrix}$$

bestimmen. Da

$$\begin{aligned}\det(A-\lambda I_3)&=\det\begin{pmatrix}1-\lambda & -2 & -2\\ 2 & 6-\lambda & 4\\ -1 & -2 & -\lambda\end{pmatrix}\\ &=(1-\lambda)(6-\lambda)(-\lambda)+8+8-2(6-\lambda)+8(1-\lambda)-4\lambda\\ &=-\lambda^3+7\lambda^2-16\lambda+12=(2-\lambda)^2(3-\lambda)\end{aligned}$$

hat A genau zwei Eigenwerte, $\lambda_1=2$ und $\lambda_2=3$, siehe Proposition VI.1.3. Für die zugehörigen Eigenräume erhalten wir:

$$E_{\lambda_1}=\ker(A-\lambda_1 I_3)=\ker\begin{pmatrix}-1 & -2 & -2\\ 2 & 4 & 4\\ -1 & -2 & -2\end{pmatrix}=\left\langle\begin{pmatrix}2\\-1\\0\end{pmatrix},\begin{pmatrix}2\\0\\-1\end{pmatrix}\right\rangle$$

$$E_{\lambda_2}=\ker(A-\lambda_2 I_3)=\ker\begin{pmatrix}-2 & -2 & -2\\ 2 & 3 & 4\\ -1 & -2 & -3\end{pmatrix}=\left\langle\begin{pmatrix}1\\-2\\1\end{pmatrix}\right\rangle$$

Der Eigenwert $\lambda_1=2$ hat daher geometrische Vielfachheit 2 und $\lambda_2=3$ hat geometrische Vielfachheit 1. Da $\left(\begin{smallmatrix}2\\-1\\0\end{smallmatrix}\right),\left(\begin{smallmatrix}2\\0\\-1\end{smallmatrix}\right),\left(\begin{smallmatrix}1\\-2\\1\end{smallmatrix}\right)$ eine Basis aus Eigenvektoren bildet ist A diagonalisierbar und es gilt:

$$\underbrace{\begin{pmatrix}2 & 2 & 1\\ -1 & 0 & -2\\ 0 & -1 & 1\end{pmatrix}^{-1}}_{S^{-1}}\underbrace{\begin{pmatrix}1 & -2 & -2\\ 2 & 6 & 4\\ -1 & -2 & 0\end{pmatrix}}_{A}\underbrace{\begin{pmatrix}2 & 2 & 1\\ -1 & 0 & -2\\ 0 & -1 & 1\end{pmatrix}}_{S}=\begin{pmatrix}2 & & \\ & 2 & \\ & & 3\end{pmatrix},$$

wobei die Spalten der Matrix S gerade die oben konstruierten Eigenvektoren sind.

VI.1.12. Beispiel. Die Eigenwerte einer Dreiecksmatrix

$$A=\begin{pmatrix}\lambda_1 & * & \cdots & *\\ 0 & \lambda_2 & \ddots & \vdots\\ \vdots & \ddots & \ddots & *\\ 0 & \cdots & 0 & \lambda_n\end{pmatrix}$$

sind genau ihre Diagonaleinträge, denn nach Satz V.1.3(n) gilt

$$\det(A-\lambda I_n)=(\lambda_1-\lambda)(\lambda_2-\lambda)\cdots(\lambda_n-\lambda),$$

und dies verschwindet genau dann, wenn λ mit einem der Diagonaleinträge λ_i übereinstimmt. Beachte, dass selbst eine Dreiecksmatrix nicht diagonalisierbar sein muss. Etwa hat die Matrix

$$A=\begin{pmatrix}\lambda & 1\\ 0 & \lambda\end{pmatrix}$$

nur einen Eigenwert, nämlich λ, und da der Eigenraum

$$E_\lambda = \ker(A - \lambda I_2) = \ker\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right) = \langle \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) \rangle$$

1-dimensional ist, existiert also keine Basis von $\mathbb{K}^2$, die aus Eigenvektoren von A besteht.

VI.1.13. Beispiel. Für $\alpha \in \mathbb{R} \setminus \pi\mathbb{Z}$ hat die reelle Matrix

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \in M_{2\times 2}(\mathbb{R})$$

keinen einzigen reellen Eigenwert, denn

$$0 = \det(A - \lambda I_2) = \det\begin{pmatrix} \cos\alpha - \lambda & -\sin\alpha \\ \sin\alpha & \cos\alpha - \lambda \end{pmatrix} = (\cos\alpha - \lambda)^2 + \sin^2\alpha$$

hat zwar die beiden Lösungen $\lambda = \cos\alpha \pm \mathbf{i}\sin\alpha$, diese sind aber nicht reell. Insbesondere ist die Matrix A über dem Körper $\mathbb{K} = \mathbb{R}$ nicht diagonalisierbar. Fassen wir A jedoch als komplexe Matrix auf,

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \in M_{2\times 2}(\mathbb{C}),$$

dann hat sie zwei verschiedene Eigenwerte,

$$\lambda_+ = \cos\alpha + \mathbf{i}\sin\alpha \quad \text{und} \quad \lambda_- = \cos\alpha - \mathbf{i}\sin\alpha,$$

mit entsprechenden Eigenräumen:

$$E_+ = \ker(A - \lambda_+ I_2) = \ker\left(\begin{smallmatrix} -\mathbf{i}\sin\alpha & -\sin\alpha \\ \sin\alpha & -\mathbf{i}\sin\alpha \end{smallmatrix}\right) = \langle \left(\begin{smallmatrix} 1 \\ -\mathbf{i} \end{smallmatrix}\right) \rangle$$
$$E_- = \ker(A - \lambda_- I_2) = \ker\left(\begin{smallmatrix} \mathbf{i}\sin\alpha & -\sin\alpha \\ \sin\alpha & \mathbf{i}\sin\alpha \end{smallmatrix}\right) = \langle \left(\begin{smallmatrix} 1 \\ \mathbf{i} \end{smallmatrix}\right) \rangle$$

Die beiden Vektoren $\left(\begin{smallmatrix} 1 \\ -\mathbf{i} \end{smallmatrix}\right)$ und $\left(\begin{smallmatrix} 1 \\ \mathbf{i} \end{smallmatrix}\right)$ bilden eine Basis von $\mathbb{C}^2$, die aus Eigenvektoren besteht, d.h. A ist über dem Körper $\mathbb{K} = \mathbb{C}$ diagonalisierbar:

$$\underbrace{\begin{pmatrix} 1 & 1 \\ -\mathbf{i} & \mathbf{i} \end{pmatrix}^{-1}}_{S^{-1}} \underbrace{\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} 1 & 1 \\ -\mathbf{i} & \mathbf{i} \end{pmatrix}}_{S} = \begin{pmatrix} \cos\alpha + \mathbf{i}\sin\alpha & 0 \\ 0 & \cos\alpha - \mathbf{i}\sin\alpha \end{pmatrix}$$

VI.1.14. Lemma. *Sei V ein endlich-dimensionaler Vektorraum über $\mathbb{K}$ und $\varphi\colon V \to V$ linear. Weiters seien $\lambda_1, \dots, \lambda_k$ paarweise verschiedene Eigenwerte von φ und $v_1, \dots, v_k \in V$ entsprechende Eigenvektoren, d.h. $v_i \neq 0$ und $\varphi(v_i) = \lambda_i v_i$, für $i = 1, \dots, k$. Dann sind die Vektoren $v_1, \dots, v_k$ linear unabhängig in V.*

Beweis. Wir führen den Beweis durch Induktion nach k. Der Fall $k = 1$ ist trivial, denn nach Voraussetzung gilt $v_i \neq 0$. Für den Induktionsschritt dürfen wir daher $v_1, \dots, v_{k-1}$ als linear unabhängig voraussetzen. Seien nun $\mu_1, \dots, \mu_k \in \mathbb{K}$, sodass

$$\mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_{k-1} v_{k-1} + \mu_k v_k = 0. \tag{VI.3}$$

Wenden wir auf diese Gleichung φ an, erhalten wir wegen $\varphi(v_i) = \lambda_i v_i$,

$$\mu_1\lambda_1 v_1 + \mu_2\lambda_2 v_2 + \cdots + \mu_{k-1}\lambda_{k-1}v_{k-1} + \mu_k\lambda_k v_k = 0.$$

Subtrahieren wir davon λ_k-mal die Gleichung (VI.3) folgt

$$\mu_1(\lambda_1 - \lambda_k)v_1 + \mu_2(\lambda_2 - \lambda_k)v_2 + \cdots + \mu_{k-1}(\lambda_{k-1} - \lambda_k)v_{k-1} = 0.$$

Da $v_1, \dots, v_{k-1}$ linear unabhängig sind, muss also

$$\mu_i(\lambda_i - \lambda_k) = 0, \qquad i = 1, \dots, k-1,$$

gelten. Nach Voraussetzung sind die Eigenwerte $\lambda_1, \dots, \lambda_k$ paarweise verschieden, also $\lambda_i - \lambda_k \neq 0$, für $i = 1, \dots, k-1$. Wir erhalten somit

$$\mu_1 = \mu_2 = \cdots = \mu_{k-1} = 0.$$

Mit (VI.3) folgt auch $\mu_k v_k = 0$ und da $v_k \neq 0$ schließlich $\mu_k = 0$. Dies zeigt, dass die Vektoren $v_1, \dots, v_k$ linear unabhängig sind. □

VI.1.15. Proposition. *Sei V ein n-dimensionaler Vektorraum über $\mathbb{K}$ und $\varphi\colon V \to V$ linear. Dann besitzt φ höchstens n verschiedene Eigenwerte. Hat φ genau n verschiedene Eigenwerte, so ist φ diagonalisierbar.*

Beweis. Dies folgt sofort aus Lemma VI.1.14. Sind $\lambda_1, \dots, \lambda_k$ paarweise verschiedene Eigenwerte von φ, dann existieren Eigenvektoren $0 \neq v_i \in V$ mit $\varphi(v_i) = \lambda_i v_i, = 1, \dots, k$. Nach Lemma VI.1.14 sind die Vektoren $v_1, \dots, v_k$ linear unabhängig, es muss daher $k \leq n$ gelten, d.h. φ kann höchstens n verschiedene Eigenwerte haben. Gilt $k = n$, dann bildet das linear unabhängige System $v_1, \dots, v_n$ eine Basis von V, die aus Eigenvektoren von φ besteht, d.h. φ ist diagonalisierbar. □

Nach Proposition VI.1.15 hat eine Matrix $A \in M_{n\times n}(\mathbb{K})$ also höchstens n verschiedene Eigenwerte. Sind es genau n verschiedene Eigenwerte, dann ist A diagonalisierbar. Etwa ist jede Dreiecksmatrix mit paarweise verschiedenen Diagonaleinträgen diagonalisierbar, vgl. Beispiel VI.1.12.

Proposition VI.1.15 liefert ein hinreichendes Kriterium für die Diagonalisierbarkeit linearer Abbildung und Matrizen. Für Diagonalisierbarkeit muss dieses Kriterium jedoch nicht notwendigerweise erfüllt sein. Etwa ist die identische Abbildung $\mathrm{id}_V\colon V \to V$ trivialerweise diagonalisierbar, obwohl sie nur einen Eigenwert, nämlich 1 hat. Das gleiche gilt für die Einheitsmatrix I_n.

Analog zu Proposition II.5.5 haben gilt:

VI.1.16. Proposition. *Sei V ein $\mathbb{K}$-Vektorraum und $W_1, \dots, W_k$ Teilräume von V. Dann sind folgende Aussagen äquivalent:*

(a) Es gilt $V = W_1 + \cdots + W_k$ und $W_i \cap (W_1 + \cdots + W_{i-1} + W_{i+1} + \cdots + W_k) = \{0\}$, für alle $i = 1, \dots, k$.

(b) Es ist $V = W_1 + \cdots + W_k$ und für alle $w_i \in W_i$ mit $w_1 + \cdots + w_k = 0$ muss schon $w_1 = \cdots = w_k = 0$ gelten.

(c) *Zu jedem $v \in V$ existieren eindeutig bestimmte Vektoren $w_i \in W_i$, sodass $v = w_1 + \cdots + w_k$.*
(d) *Ist U ein weiterer $\mathbb{K}$-Vektorraum und sind $\varphi_i \colon W_i \to U$ linear, dann existiert eine eindeutig bestimmte lineare Abbildung $\varphi \colon V \to U$, sodass $\varphi|_{W_i} = \varphi_i$, für alle $i = 1, \dots, k$.*
(e) *Es existieren Projektoren $\pi_i \colon V \to V$, $\pi_i \circ \pi_i = \pi_i$, $i = 1, \dots, k$, sodass $\operatorname{img}(\pi_i) = W_i$, $\pi_1 + \cdots + \pi_k = \mathrm{id}_V$ und $\pi_i \circ \pi_j = 0 = \pi_j \circ \pi_i$, für alle $i \neq j$.*

Sind diese äquivalenten Eigenschaften erfüllt, und ist B_i eine Basis von W_i, $i = 1, \dots, k$, dann bildet $B_1 \cup \cdots \cup B_k$ eine Basis von V und es gilt $B_i \cap B_j = \emptyset$, für alle $i \neq j$. Im endlich-dimensionalen Fall haben wir daher die Dimensionsformel

$$\dim(V) = \dim(W_1) + \cdots + \dim(W_k).$$

Beweis. Ad (a)⇒(b): Schreiben wir die Gleichung $w_1 + \cdots + w_k = 0$ in der Form $-w_i = w_1 + \cdots + w_{i-1} + w_{i+1} + \cdots + w_k$, so sehen wir, dass

$$-w_i \in W_i \cap \bigl(W_1 + \cdots + W_{i-1} + W_{i+1} + \cdots + W_k\bigr), \qquad i = 1, \dots, k.$$

Aus der zweiten Voraussetzung in (a) folgt daher $w_i = 0$, für alle i.

Ad (b)⇒(c): Wegen $V = W_1 + \cdots + W_k$ lässt sich jedes $v \in V$ in der Form $v = w_1 + \cdots + w_k$ schreiben, wobei $w_i \in W_i$. Um die Eindeutigkeit zu zeigen, sei $v = \tilde{w}_1 + \cdots + \tilde{w}_k$ eine weitere solche Darstellung, $\tilde{w}_i \in W_i$. Es gilt dann $0 = w_1' + \cdots + w_k'$, wobei $w_i' := \tilde{w}_i - w_i \in W_i$. Aus der zweiten Voraussetzung in (b) folgt $w_i' = 0$, also $w_i = \tilde{w}_i$, d.h. die Darstellung $v = w_1 + \cdots + w_k$ ist eindeutig.

Ad (c)⇒(d): Da sich jedes $v \in V$ auf eindeutige Weise in der Form $v = w_1 + \cdots + w_k$ schreiben lässt, stellt

$$\varphi \colon V \to U, \quad \varphi(v) := \varphi_1(w_1) + \cdots + \varphi_k(w_k),$$

eine wohldefiniert Abbildung dar. Es ist noch zu zeigen, dass φ linear ist. Für $\lambda \in \mathbb{K}$ gilt $\lambda v = \lambda w_1 + \cdots + \lambda w_k$ und wegen der Linearität von φ_i daher

$$\varphi(\lambda v) = \varphi_1(\lambda w_1) + \cdots + \varphi_k(\lambda w_k) = \lambda\bigl(\varphi_1(w_1) + \cdots + \varphi_k(w_k)\bigr) = \lambda\varphi(v).$$

Ist $\tilde{v} = \tilde{w}_1 + \cdots + \tilde{w}_k$, $\tilde{w}_i \in W_i$, dann gilt $v + \tilde{v} = (w_1 + \tilde{w}_1) + \cdots + (w_k + \tilde{w}_k)$ mit $w_i + \tilde{w}_i \in W_i$ und wegen der Linearität der φ_i also

$$\begin{aligned}\varphi(v + \tilde{v}) &= \varphi_1(w_1 + \tilde{w}_1) + \cdots + \varphi_k(w_k + \tilde{w}_k)\\ &= \bigl(\varphi_1(w_1) + \cdots + \varphi_k(w_k)\bigr) + \bigl(\varphi_1(\tilde{w}_1) + \cdots + \varphi_k(\tilde{w}_k)\bigr) = \varphi(v) + \varphi(\tilde{v}).\end{aligned}$$

Damit ist die Linearität von φ gezeigt. Nach Konstruktion gilt $\varphi|_{W_i} = \varphi_i$.

Ad (d)⇒(e): Nach Voraussetzung existieren lineare Abbildung $\pi_i \colon V \to W_i$, $i = 1, \dots, k$, sodass $\pi_i|_{W_i} = \mathrm{id}_{W_i}$ und $\pi_i|_{W_j} = 0$ für alle $j \neq i$. Fassen wir diese als Abbildungen $\pi_i \colon V \to V$ auf, dann gilt offensichtlich $\operatorname{img}(\pi_i) = W_i$. Auch sind dies Projektoren, denn nach Konstruktion gilt $(\pi_i \circ \pi_i)|_{W_j} = \pi_i|_{W_j}$, für alle j, aus der Eindeutigkeitsaussage in (d) folgt daher $\pi_i \circ \pi_i = \pi_i$. Weiters gilt $(\pi_1 + \cdots + \pi_k)|_{W_i} = \mathrm{id}_V|_{W_i}$, für alle i, und wegen der Eindeutigkeitsaussage in

(d) daher $\pi_1 + \cdots + \pi_k = \mathrm{id}_V$. Analog lässt sich für $i \neq j$ die Relation $\pi_i \circ \pi_j = 0$ zeigen, denn $(\pi_i \circ \pi_j)|_{W_l} = 0$, für alle $l = 1, \dots, k$.

Ad (e)⇒(a): Es ist $V = W_1 + \cdots + W_k$, denn für jedes $v \in V$ haben wir $v = \mathrm{id}_V(v) = (\pi_1 + \cdots + \pi_k)(v) = \pi_1(v) + \cdots + \pi_k(v)$, wobei $\pi_i(v) \in \mathrm{img}(\pi_i) = W_i$. Aus $W_j = \mathrm{img}(\pi_j)$ und $\pi_i \circ \pi_j = 0$ erhalten wir $\pi_i(W_j) = \{0\}$, für alle $i \neq j$. Folglich bildet π_i den gesamten Teilraum $W_1 + \cdots + W_{i-1} + W_{i+1} + \cdots + W_k$ auf 0 ab. Für $w \in W_i \cap \bigl(W_1 + \cdots + W_{i-1} + W_{i+1} + \cdots + W_k\bigr)$ gilt daher $w = \pi_i(w) = 0$, also $W_i \cap \bigl(W_1 + \cdots + W_{i-1} + W_{i+1} + \cdots + W_k\bigr) = \{0\}$. Damit ist die Äquivalenz der fünf Eigenschaften gezeigt.

Sei nun B_i eine Basis von W_i, $i = 1, \dots, k$. Für $i \neq j$ gilt $B_i \cap B_j = \emptyset$, denn $B_i \cap B_j \subseteq W_i \cap W_j \subseteq W_i \cap \bigl(W_1 + \cdots + W_{i-1} + W_{i+1} + \cdots + W_k\bigr) = \{0\}$. Da $V = W_1 + \cdots + W_k$, ist $B := B_1 \cup \cdots \cup B_k$ ein Erzeugendensystem von V. Sei nun $\sum_{b \in B} \lambda_b b = 0$, wobei fast alle $\lambda_b \in \mathbb{K}$ verschwinden. Dann gilt $\sum_{i=1}^k \sum_{b \in B_i} \lambda_b b = 0$, und wegen (b) daher $\sum_{b \in B_i} \lambda_b b = 0$, für jedes $i = 1, \dots, k$. Aufgrund der linearen Unabhängigkeit von B_i erhalten wir $\lambda_b = 0$, für alle $b \in B$. Dies zeigt, dass B linear unabhängig, also eine Basis von V ist. □

VI.1.17. Definition (Direkte Summe). Sind die äquivalenten Eigenschaften in Proposition VI.1.16 erfüllt, dann schreiben wir

$$V = \bigoplus_{i=1}^k W_i \qquad \text{oder} \qquad V = W_1 \oplus \cdots \oplus W_k, \tag{VI.4}$$

und sagen V ist die *direkte Summe* der Teilräume $W_1, \dots, W_k$. Offensichtlich ist dies eine Verallgemeinerung der in Definition II.5.6 eingeführten direkten Summe zweier Teilräume. Die Projektoren aus Proposition VI.1.16(e) werden als die mit der Zerlegung (VI.4) *assoziierten Projektionen* oder als die damit assoziierte *Zerlegung der Eins* bezeichnet.

VI.1.18. Satz (Diagonalisierbarkeit). *Sei V ein n-dimensionaler Vektorraum über $\mathbb{K}$, und $\varphi\colon V \to V$ linear. Weiters bezeichnen $\lambda_1, \dots, \lambda_k$ alle (verschiedenen) Eigenwerte und $E_{\lambda_i} = \ker(\varphi - \lambda_i \,\mathrm{id}_V)$ die zugehörigen Eigenräume von φ. Dann sind folgende Aussagen äquivalent:*

(a) φ ist diagonalisierbar.
(b) Es gilt $\dim(E_{\lambda_1}) + \cdots + \dim(E_{\lambda_k}) = n$, d.h. die Summe der geometrischen Vielfachheiten aller Eigenwerte stimmt mit $\dim(V)$ überein.
(c) Es gilt $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$.

Ist φ diagonalisierbar, und bezeichnen $\pi_i\colon V \to V$ die mit der Zerlegung $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$ assoziierten Projektoren, dann gilt (Spektralzerlegung von φ)

$$\varphi = \lambda_1 \pi_1 + \cdots + \lambda_k \pi_k. \tag{VI.5}$$

Weiters existiert eine Basis B von V, sodass

$$[\varphi]_{BB} = \begin{pmatrix} \lambda_1 I_{m_1} & & \\ & \ddots & \\ & & \lambda_k I_{m_k} \end{pmatrix}, \tag{VI.6}$$

wobei $m_i = \dim(E_{\lambda_i})$ die geometrischen Vielfachheiten bezeichnen.

Beweis. Es bezeichne $E = E_{\lambda_1} + \cdots + E_{\lambda_k}$ die Summe aller Eigenräume. Dies ist eine direkte Summe, d.h. es gilt

$$E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}. \tag{VI.7}$$

Sind nämlich $v_i \in E_{\lambda_i}$ und $v_1 + \cdots + v_k = 0$, dann folgt mit Lemma VI.1.14 sofort $v_1 = \cdots = v_k = 0$, siehe auch Proposition VI.1.16(b). Insbesondere haben wir

$$\dim(E) = \dim(E_{\lambda_1}) + \cdots + \dim(E_{\lambda_k}). \tag{VI.8}$$

Ad (a)$\Rightarrow$(b): Nach Voraussetzung existiert eine Basis $b_1, \dots, b_n$ von V, die aus Eigenvektoren besteht. Da jeder dieser Basisvektoren in einem der Eigenräume E_{λ_i} liegt, enthält E eine linear unabhängige Teilmenge mit n Elementen, es ist daher $\dim(E) \geq n$. Wegen $E \subseteq V$ muss aber auch $\dim(E) \leq n$ gelten. Somit ist $\dim(E) = n$. Aus (VI.8) erhalten wir nun $\dim(E_{\lambda_1}) + \cdots + \dim(E_{\lambda_k}) = n$.

Ad (b)$\Rightarrow$(c): Nach Voraussetzung gilt $\dim(V) = \dim(E_{\lambda_1}) + \cdots + \dim(E_{\lambda_k})$. Zusammen mit (VI.8) folgt $\dim(E) = \dim(V)$ und daher $E = V$, da ja $E \subseteq V$. Aus (VI.7) folgt somit $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$.

Ad (c)$\Rightarrow$(a): Für $i = 1, \dots, k$ sei $b_1^{(i)}, \dots, b_{m_i}^{(i)}$ eine Basis des Eigenraums E_{λ_i}, wobei $m_i = \dim(E_{\lambda_i})$. Nach Proposition VI.1.16 ist dann

$$b_1^{(1)}, \dots, b_{m_1}^{(1)}, b_1^{(2)}, \dots, b_{m_2}^{(2)}, \dots, b_1^{(k)}, \dots, b_{m_k}^{(k)}$$

eine Basis von $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$. Nach Konstruktion besteht diese Basis aus Eigenvektoren von φ, also ist φ diagonalisierbar. Die Matrixdarstellung von φ bezüglich dieser Basis hat die Form (VI.6). Die Gleichung (VI.5) folgt aus der Eindeutigkeitsaussage in Proposition VI.1.16(d), denn offensichtlich gilt $\varphi|_{E_{\lambda_i}} = \lambda_i \operatorname{id}_{E_{\lambda_i}} = (\lambda_1 \pi_1 + \cdots + \lambda_k \pi_k)|_{E_{\lambda_i}}$, für alle $i = 1, \dots, k$. □

Ist $p \in \mathbb{K}[z]$ ein Polynom, $p = \sum_i p_i z^i$, und $A \in M_{n\times n}(\mathbb{K})$ eine quadratische Matrix, dann können wir A in p einsetzen und erhalten eine Matrix $p(A) \in M_{n\times n}(\mathbb{K})$, genauer

$$p(A) := \sum_i p_i A^i = p_0 I_n + p_1 A + p_2 A^2 + p_3 A^3 + \cdots$$

Analog definieren wir für jeden Endomorphismus $\varphi\colon V \to V$,

$$p(\varphi) := \sum_i p_i \varphi^i = p_0 \operatorname{id}_V + p_1 \varphi + p_2 \varphi^2 + p_3 \varphi^3 + \cdots$$

wobei $\varphi^i = \varphi \circ \cdots \circ \varphi$ und $\varphi^0 = \operatorname{id}_V$.

VI.1.19. Lemma. *Sind $p, q \in \mathbb{K}[z]$, $\lambda \in \mathbb{K}$ und $A \in M_{n\times n}(\mathbb{K})$, dann gilt*

$$(p+q)(A) = p(A) + q(A), \quad (\lambda p)(A) = \lambda p(A) \quad \text{und} \quad (pq)(A) = p(A)q(A),$$

d.h. $\mathbb{K}[z] \to M_{n\times n}(\mathbb{K})$, $p \mapsto p(A)$, ist ein Homomorphismus von $\mathbb{K}$-Algebren. Weiters gilt $p(A^t) = p(A)^t$ und $p(SAS^{-1}) = Sp(A)S^{-1}$, für jedes $S \in \mathrm{GL}_n(\mathbb{K})$.

Ist V ein $\mathbb{K}$-Vektorraum und $\varphi \in \operatorname{end}(V)$, dann haben wir analog

$$(p+q)(\varphi) = p(\varphi) + q(\varphi), \quad (\lambda p)(\varphi) = \lambda p(\varphi) \quad \text{und} \quad (pq)(\varphi) = p(\varphi)q(\varphi),$$

d.h. $\mathbb{K}[z] \to \operatorname{end}(V)$, $p \mapsto p(\varphi)$, ist ein Homomorphismus von $\mathbb{K}$-Algebren. Weiters gilt $p(\varphi^t) = p(\varphi)^t$ und $p(\psi \circ \varphi \circ \psi^{-1}) = \psi \circ p(\varphi) \circ \psi^{-1}$, für jeden linearen Isomorphismus $\psi\colon V \xrightarrow{\cong} W$. Ist V endlich-dimensional und B eine geordnete Basis von V, dann gilt $[p(\varphi)]_{BB} = p([\varphi]_{BB})$.

Beweis. Bezeichnen $p = \sum_i p_i z^i$ und $q = \sum_i q_i z^i$ zwei Polynome, dann folgt

$$\begin{aligned}(p+q)(A) &= \big(\textstyle\sum_i p_i z^i + \sum_i q_i z^i\big)(A) = \big(\sum_i (p_i + q_i) z^i\big)(A) \\ &= \textstyle\sum_i (p_i+q_i)A^i = \sum_i p_i A^i + \sum_i q_i A^i = p(A) + q(A).\end{aligned}$$

Für $\lambda \in \mathbb{K}$ erhalten wir analog

$$(\lambda p)(A) = \big(\lambda \textstyle\sum_i p_i z^i\big)(A) = \big(\sum_i (\lambda p_i) z^i\big)(A) = \sum_i (\lambda p_i) A^i = \lambda \sum_i p_i A^i = \lambda p(A).$$

Ebenso

$$\begin{aligned}(pq)(A) &= \Big(\big(\textstyle\sum_i p_i z^i\big)\big(\sum_j q_j z^j\big)\Big)(A) \\ &= \Big(\textstyle\sum_k \big(\sum_{i+j=k} p_i q_j\big) z^k\Big)(A) = \sum_k \big(\sum_{i+j=k} p_i q_j\big) A^k \\ &= \textstyle\sum_k \sum_{i+j=k} p_i A^i q_j A^j = \big(\sum_i p_i A^i\big)\big(\sum_j q_j A^j\big) = p(A)q(A)\end{aligned}$$

und $p(A^t) = (\sum_i p_i z^i)(A^t) = \sum_i p_i (A^t)^i = \sum_i p_i (A^i)^t = \big(\sum_i p_i A^i\big)^t = p(A)^t$. Für $S \in \mathrm{GL}_n(\mathbb{K})$ gilt $(SAS^{-1})^i = (SAS^{-1})\cdots(SAS^{-1}) = SA^iS^{-1}$ und daher

$$p(SAS^{-1}) = \textstyle\sum_i p_i (SAS^{-1})^i = \sum_i p_i SA^iS^{-1} = S\big(\sum_i p_i A^k\big)S^{-1} = Sp(A)S^{-1}.$$

Die entsprechenden Eigenschaften von $p(\varphi)$ lassen sich völlig analog herleiten. Ist schließlich V endlich-dimensional und B eine geordnete Basis von V, dann gilt

$$[p(\varphi)]_{BB} = \big[\textstyle\sum_i p_i \varphi^i\big]_{BB} = \sum_i p_i [\varphi^i]_{BB} = \sum_i p_i ([\varphi]_{BB})^i = p([\varphi]_{BB}),$$

wobei wir die Eigenschaften in Satz IV.6.15 verwendet haben. □

VI.1.20. Beispiel. Für $p = 2 + 3z + 5z^2$ und $A = \left(\begin{smallmatrix}1 & 2\\ 3 & 4\end{smallmatrix}\right)$ ist

$$p(A) = 2I_2 + 3A + 5A^2 = 2\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} + 3\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix} + 5\begin{pmatrix}7 & 10\\ 15 & 22\end{pmatrix} = \begin{pmatrix}40 & 56\\ 84 & 124\end{pmatrix}.$$

VI.1.21. Beispiel. Ist $p \in \mathbb{K}[z]$ ein Polynom und D eine Dreiecksmatrix,

$$D = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ & \lambda_2 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n \end{pmatrix}, \quad \text{dann gilt} \quad p(D) = \begin{pmatrix} p(\lambda_1) & * & \cdots & * \\ & p(\lambda_2) & \ddots & \vdots \\ & & \ddots & * \\ & & & p(\lambda_n) \end{pmatrix},$$

siehe Aufgabe 35. Für die Spektra gilt daher $\sigma(p(D)) = p(\sigma(D))$.

VI.1.22. Bemerkung. Beachte, dass i.A. $p(\lambda A) \neq \lambda p(A)$, $p(A+B) \neq p(A) + p(B)$ und $p(AB) \neq p(A)p(B)$ gilt, wobei $\lambda \in \mathbb{K}$, $p \in \mathbb{K}[z]$ und $A, B \in M_{n\times n}(\mathbb{K})$.

VI.1.23. Proposition. *Sei V ein endlich-dimensionaler Vektorraum über $\mathbb{K}$, $\varphi\colon V \to V$ eine lineare Abbildung und $p \in \mathbb{K}[z]$ ein Polynom. Ist v Eigenvektor von φ zum Eigenwert λ, dann ist v auch Eigenwert von $p(\varphi)$ zum Eigenwert $p(\lambda)$. Für die Spektra gilt daher $p(\sigma(\varphi)) \subseteq \sigma(p(\varphi))$. Existiert eine Basis B von V, sodass $[\varphi]_{BB}$ obere Dreiecksgestalt hat, dann haben wir sogar $p(\sigma(\varphi)) = \sigma(p(\varphi))$.*

Beweis. Ist v Eigenvektor zum Eigenwert λ, d.h. $\varphi(v) = \lambda v$, dann folgt $p(\varphi)(v) = \big(\sum_i p_i \varphi^i\big)(v) = \sum_i p_i \varphi^i(v) = \sum_i p_i \lambda^i v = p(\lambda)v$, also ist v auch Eigenvektor von $p(\varphi)$ zum Eigenwert $p(\lambda)$. Insbesondere gilt $p(\sigma(\varphi)) \subseteq \sigma(p(\varphi))$. Sei nun B eine Basis von V, sodass $[\varphi]_{BB}$ obere Dreiecksgestalt hat. Aus Beispiel VI.1.21 folgt $p(\sigma([\varphi]_{BB})) = \sigma(p([\varphi]_{BB}))$ und somit

$$p(\sigma(\varphi)) = p(\sigma([\varphi]_{BB})) = \sigma(p([\varphi]_{BB})) = \sigma([p(\varphi)]_{BB}) = \sigma(p(\varphi)),$$

wobei wir auch zwei mal (VI.2) verwendet haben. $\square$

Für eine Matrix $A \in M_{n\times n}(\mathbb{K})$ und ein Polynom $p \in \mathbb{K}[z]$ besagt Proposition VI.1.23, dass jeder Eigenvektor von A zum Eigenwert λ auch ein Eigenvektor von $p(A)$ ist, und zwar zum Eigenwert $p(\lambda)$. Für die Spektra gilt daher $p(\sigma(A)) \subseteq \sigma(p(A))$. Ist A ähnlich zu einer oberen Dreiecksmatrix, dann haben wir sogar $p(\sigma(A)) = \sigma(p(A))$.

VI.1.24. Beispiel. Betrachte das Polynom $p = z^2$ und die reelle Matrix $A = \left(\begin{smallmatrix} 1 & & \\ & -1 & \\ & & 2 \end{smallmatrix}\right)$. Dann gilt $p(A) = \left(\begin{smallmatrix} 1 & & \\ & 1 & \\ & & 4 \end{smallmatrix}\right)$, also $p(\sigma(A)) = \sigma(p(A))$. Beachte jedoch, dass A drei 1-dimensionale Eigenräume besitzt, während $p(A)$ nur zwei Eigenräume hat von denen einer 2-dimensional ist.

VI.1.25. Beispiel. In Beispiel VI.1.13 haben wir gesehen, dass die reelle Matrix $A = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ über $\mathbb{K} = \mathbb{R}$ keinen Eigenwert besitzt, d.h. $\sigma(A) = \emptyset$. Setzen wir A in das Polynom $p = z^2$ ein, erhalten wir $p(A) = A^2 = \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, eine diagonalisierbare Abbildung mit Spektrum $\sigma(p(A)) = \{-1\}$. Dieses Beispiel zeigt, dass i.A. $p(\sigma(A)) \neq \sigma(p(A))$.

Wir beenden diesen Abschnitt mit einer ersten Anwendung:

VI.1.26. Beispiel (Fibonacci-Folge). Betrachte die rekursiv definierte Fibonacci-Folge: $0, 1, 1, 2, 3, 5, 8, 13, 21, 34, \dots$ Genauer:

$$x_0 = 0, \quad x_1 = 1, \quad x_n = x_{n-1} + x_{n-2}, \quad n \geq 2.$$

Wir wollen eine geschlossene Formel für x_n herleiten. Dazu betrachten wir die Vektoren $v_n := \left(\begin{smallmatrix} x_n \\ x_{n+1} \end{smallmatrix}\right) \in \mathbb{R}^2$ und die Matrix $A = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right)$. Es gilt dann

$$v_0 = \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right), \qquad v_n = Av_{n-1}, \quad n \geq 1.$$

Somit ist $v_n = A^n v_0$ und x_n also die erste Komponente von $A^n v_0$. Es genügt daher eine geschlossene Formel für A^n herzuleiten. Dies lässt sich durch Diagonalisieren von A erreichen. Lösen der quadratischen Gleichung $0 = \det(A - \lambda I_2) = \det\left(\begin{smallmatrix} -\lambda & 1 \\ 1 & 1-\lambda \end{smallmatrix}\right) = \lambda^2 - \lambda - 1$ liefert die beiden verschiedenen Eigenwerte

$$\lambda_\pm = \frac{1 \pm \sqrt{5}}{2}.$$

Die Matrix A ist also diagonalisierbar, d.h. es existiert eine invertierbare Matrix S, sodass $S^{-1}AS = D$, wobei $D = \left(\begin{smallmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{smallmatrix}\right)$. Für die Potenzen von A folgt

$$A^n = \left(SDS^{-1}\right)^n = (SDS^{-1}) \cdots (SDS^{-1}) = SD^nS^{-1} = S\left(\begin{smallmatrix} \lambda_+^n & 0 \\ 0 & \lambda_-^n \end{smallmatrix}\right)S^{-1}.$$

Um eine Matrix S wie oben zu bestimmen, betrachten wir die beiden Eigenräume:

$$E_+ = \ker(A - \lambda_+ I_2) = \ker\begin{pmatrix} -\frac{1+\sqrt{5}}{2} & 1 \\ 1 & \frac{1-\sqrt{5}}{2} \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ \frac{1+\sqrt{5}}{2} \end{pmatrix} \right\rangle$$

$$E_- = \ker(A - \lambda_- I_2) = \ker\begin{pmatrix} -\frac{1-\sqrt{5}}{2} & 1 \\ 1 & \frac{1+\sqrt{5}}{2} \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ \frac{1-\sqrt{5}}{2} \end{pmatrix} \right\rangle$$

Somit

$$S = \begin{pmatrix} 1 & 1 \\ \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \end{pmatrix} \quad \text{und} \quad S^{-1} = \frac{1}{\sqrt{5}}\begin{pmatrix} -\frac{1-\sqrt{5}}{2} & 1 \\ \frac{1+\sqrt{5}}{2} & -1 \end{pmatrix}.$$

Es folgt

$$\begin{aligned}
\begin{pmatrix} x_n \\ x_{n+1} \end{pmatrix} &= v_n = A^n v_0 = S\begin{pmatrix} \lambda_+^n & 0 \\ 0 & \lambda_-^n \end{pmatrix}S^{-1}\begin{pmatrix} 0 \\ 1 \end{pmatrix} \\
&= \frac{1}{\sqrt{5}}S\begin{pmatrix} \lambda_+^n & 0 \\ 0 & \lambda_-^n \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{5}}S\begin{pmatrix} \lambda_+^n \\ -\lambda_-^n \end{pmatrix} = \frac{1}{\sqrt{5}}\begin{pmatrix} \lambda_+^n - \lambda_-^n \\ \lambda_+^{n+1} - \lambda_-^{n+1} \end{pmatrix}
\end{aligned}$$

also $x_n = \frac{\lambda_+^n - \lambda_-^n}{\sqrt{5}}$ und daher

$$x_n = \frac{\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n}{\sqrt{5}},$$

die gesuchte geschlossene Formel für die n-te Fibonacci-Zahl. Mit dieser Methode lassen sich natürlich viel allgemeinere lineare Rekursionen lösen.

VI.2. Charakteristisches Polynom. Die die Eigenwerte charakterisierende Gleichung $\det(\varphi - \lambda \operatorname{id}_V) = 0$, siehe Proposition VI.1.3, ist polynomial in λ. Wir wollen dieses Polynom nun genauer untersuchen und damit u.a. eine weitere Charakterisierung der Diagonalisierbarkeit herleiten.

Unter dem *Grad* eines Polynoms $p = p_0 + p_1 z + p_2 z^2 + \cdots$ verstehen wir die größte Zahl n, für die $p_n \neq 0$ gilt. Wir schreiben dafür $\deg(p) = n$. Der Grad des Nullpolynoms wird durch $\deg(0) := -\infty$ definiert. Mit dieser Konvention gilt

$$\deg(pq) = \deg(p) + \deg(q),$$

für beliebige Polynome p und q. Beachte, dass ein Polynom genau dann konstant ist, wenn sein Grad kleiner oder gleich Null ist. Polynome vom Grad 1 sind genau die Polynome der Form $p = p_0 + p_1 z$, wobei $p_0, p_1 \in \mathbb{K}$ und $p_1 \neq 0$.

VI.2.1. Lemma (Polynomdivision). *Seien $p, q \in \mathbb{K}[z]$ Polynome und $q \neq 0$. Dann existieren eindeutig bestimmte Polynome $s, r \in \mathbb{K}[z]$, sodass*

$$p = qs + r \quad \textit{und} \quad \deg(r) < \deg(q).$$

Beweis. Wir beginnen mit der Eindeutigkeit der Darstellung. Seien dazu $s, \tilde{s}, r, \tilde{r} \in \mathbb{K}[z]$, sodass $qs + r = p = q\tilde{s} + \tilde{r}$, $\deg(r) < \deg(q)$ und $\deg(\tilde{r}) < \deg(q)$. Es folgt $q(s - \tilde{s}) = \tilde{r} - r$ und $\deg(\tilde{r} - r) < \deg(q)$. Somit $\deg(q(s - \tilde{s})) < \deg(q)$, also $s - \tilde{s} = 0$. Dies zeigt, $s = \tilde{s}$, woraus aber auch sofort $r = \tilde{r}$ folgt. Damit ist die Eindeutigkeit der Darstellung gezeigt.

Sei $q = \sum_{i=0}^m q_i z^i$, wobei $m = \deg(q) \geq 0$ und daher $q_m \neq 0$. Ist $\deg(p) < m$, dann haben $s = 0$ und $r = p$ die gewünschten Eigenschaften. O.B.d.A. sei daher $n := \deg(p) \geq m$. Wir führen den Beweis mittels Induktion nach dem Grad von p. Bezeichne dazu $p = \sum_{i=0}^n p_i z^i$, also $p_n \neq 0$. Betrachten wir das Polynom $\tilde{s} := \frac{p_n}{q_m} z^{n-m}$ dann gilt $\deg(p - q\tilde{s}) < \deg(p)$. Nach Induktionsvoraussetzung existieren daher Polynome $\bar{s}$ und r, sodass $p - q\tilde{s} = q\bar{s} + r$ und $\deg(r) < \deg(q)$. Setzen wir $s := \tilde{s} + \bar{s}$, so folgt $p = qs + r$, wie gewünscht. □

VI.2.2. Beispiel. Mit dem bekannten Algorithmus erhalten wir etwa:

$$\underbrace{x^5 - 2x^4 + 2x^3 + 9x^2 - 12x + 10}_{p} = (\underbrace{x^2 - 2x + 3}_{q})(\underbrace{x^3 - x + 7}_{s}) + \underbrace{5x - 11}_{r}.$$

Beachte, dass dieser Algorithmus auf der im Beweis von Lemma VI.2.1 verwendeten Konstruktion basiert.

VI.2.3. Lemma. *Sei $p \in \mathbb{K}[z]$ ein Polynom und $\lambda \in \mathbb{K}$ eine Nullstelle von p. Dann existiert ein eindeutig bestimmtes Polynom $q \in \mathbb{K}[z]$, sodass $p = (z - \lambda)q$.*

Beweis. Nach Lemma VI.2.1 existieren eindeutig bestimmte Polynome q und r, sodass $p = (z - \lambda)q + r$ und $\deg(r) < \deg(z - \lambda) = 1$. Somit ist r ein konstantes Polynom, d.h. $r = r_0$, wobei $r_0 \in \mathbb{K}$. Einsetzen von λ liefert

$$0 = p(\lambda) = (\lambda - \lambda)q(\lambda) + r(\lambda) = r_0,$$

also $r = 0$, und daher $p = (z - \lambda)q$. □