

Ergänzungen zur Linearen Algebra 2

A.1. Motivation zur Eigenwert-Theorie

Nachdem die *Lineare Algebra 1* die Grundzüge der Theorie der Vektorräume dargelegt hat, wird in der *Linearen Algebra 2* das Studium der linearen Abbildungen zum Hauptmotiv erhoben. Wir wollen zunächst durch einige Beispiele an diesen Themenkomplex heranführen.

A.1.1. Darstellungsmatrizen und Basiswechsel. Sei K ein Körper und $f: V \rightarrow W$ eine lineare Abbildung zwischen K -Vektorräumen V und W . Aus der *Linearen Algebra 1* wissen wir, dass K -Vektorräume bis auf Isomorphie durch ihre *Dimension* klassifiziert sind. Wir gehen im Folgenden davon aus, dass V und W endlich-dimensional sind und schreiben $n = \dim V$, $m = \dim W$. Dann ist $V \cong K^n$ und $W \cong K^m$. Konkreter sei $\mathcal{B} = (b_1, \dots, b_n)$ eine geordnete Basis von V und $\mathcal{C} = (c_1, \dots, c_m)$ eine Basis von W . Dann handelt es sich bei der Abbildung

$$\varphi_{\mathcal{B}}: K^n \rightarrow V, \quad (\lambda_1, \dots, \lambda_n) \mapsto \lambda_1 b_1 + \dots + \lambda_n b_n,$$

um einen Isomorphismus und dasselbe gilt für die in analoger Weise definierte Abbildung $\varphi_{\mathcal{C}}: K^m \rightarrow W$. Wir erhalten ein Diagramm

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \uparrow \varphi_{\mathcal{B}} & & \varphi_{\mathcal{C}} \uparrow \\ K^n & & K^m. \end{array}$$

Durch die Verkettung $\ell = \varphi_{\mathcal{C}}^{-1} \circ f \circ \varphi_{\mathcal{B}}$ ergänzt sich dieses zu dem kommutativen Diagramm

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \uparrow \varphi_{\mathcal{B}} & & \varphi_{\mathcal{C}} \uparrow \\ K^n & \xrightarrow{\ell} & K^m. \end{array}$$

Nun wissen wir, dass sich jede lineare Abbildung $\ell: K^n \rightarrow K^m$ in der Form $\ell = \ell_A$, $\ell_A(v) = Av$ ($v \in V$), mit einer eindeutig bestimmten Matrix $A \in K^{m \times n}$ schreiben lässt. Die Matrix A bezeichnen wir in diesem Kontext bekanntlich als die **Darstellungsmatrix** von f bezüglich der Basen \mathcal{B} , \mathcal{C} und schreiben hierfür auch $A = M_{\mathcal{B}, \mathcal{C}}(f)$.

Wie bereits eingangs erwähnt, liegt die Hauptinteresse der linearen Abbildung f . In diesem Zusammenhang sehen wir die Matrix A vornehmlich als ein *Hilfsmittel*

zum Studium von f an. Da unser Interesse an A somit nur in seiner Dienlichkeit für den Erkenntnisgewinn über f begründet liegt, drängt sich die Frage auf, welche anderen Matrizen uns noch für diesen Zweck zur Verfügung stehen. Wir wählen somit weitere geordnete Basen \mathcal{B}' und \mathcal{C}' von V bzw. W und wollen den Zusammenhang zwischen $A = M_{\mathcal{B}, \mathcal{C}}(f)$ und $B = M_{\mathcal{B}', \mathcal{C}'}(f)$ untersuchen. Gemäß Dimensionstheorie gilt $\#\mathcal{B}' = \#\mathcal{B} = n$ und $\#\mathcal{C}' = \#\mathcal{C} = m$. Wir erhalten ein kommutatives Diagramm

$$\begin{array}{ccc}
 K^n & \xrightarrow{\ell_B} & K^m \\
 \searrow \varphi_{\mathcal{B}'} & & \swarrow \varphi_{\mathcal{C}'} \\
 & V \xrightarrow{f} W & \\
 \nearrow \varphi_{\mathcal{B}} & & \nwarrow \varphi_{\mathcal{C}} \\
 K^n & \xrightarrow{\ell_A} & K^m .
 \end{array}$$

Wir ergänzen dieses Diagramm nun um die Abbildungen $\varphi_{\mathcal{B}}^{-1} \circ \varphi_{\mathcal{B}'}$ und $\varphi_{\mathcal{C}'}^{-1} \circ \varphi_{\mathcal{C}}$, welche sich wiederum in der Form ℓ_T bzw. ℓ_S mit geeigneten Matrizen $T \in K^{n \times n}$, $S \in K^{m \times m}$ schreiben:

$$\begin{array}{ccc}
 K^n & \xrightarrow{\ell_B} & K^m \\
 \searrow \varphi_{\mathcal{B}'} & & \swarrow \varphi_{\mathcal{C}'} \\
 \ell_T \downarrow & V \xrightarrow{f} W & \uparrow \ell_S \\
 \nearrow \varphi_{\mathcal{B}} & & \nwarrow \varphi_{\mathcal{C}} \\
 K^n & \xrightarrow{\ell_A} & K^m .
 \end{array}$$

Da die Abbildungen ℓ_T und ℓ_S Isomorphismen sind, handelt es sich bei T und S um invertierbare Matrizen. Da im obigen Diagramm alle beteiligten Dreiecke und Trapeze kommutieren, kommutiert auch das äußere Rechteck und wir haben $\ell_B = \ell_S \circ \ell_A \circ \ell_T$. Bekanntlich folgt hieraus

$$B = SAT.$$

Umgekehrt kann man sich überlegen, dass zu gegebenem $A = M_{\mathcal{B}, \mathcal{C}}(f)$ aus der Matrixgleichung $B = SAT$ mit $S \in GL_m(K)$, $T \in GL_n(K)$ bereits folgt, dass B Darstellungsmatrix von f bezüglich geeignet gewählter Basen von V und W ist. (Übung: Man überlege sich, wie diese Basen aus \mathcal{B} und \mathcal{C} mit Hilfe von T und S zu berechnen sind.)

A.1.2. Äquivalenz von Matrizen und Rang-Normalform. Unsere Überlegungen führen also auf die folgende Definition: Zwei Matrizen $A, B \in K^{m \times n}$ heißen **äquivalent**, falls es invertierbare Matrizen $S \in GL_m(K)$, $T \in GL_n(K)$ mit $B = SAT$ gibt. Man prüft leicht nach, dass es sich dabei um eine Äquivalenzrelation handelt.

Satz A.1 (Rang-Normalform). Zwei Matrizen in $K^{m \times n}$ sind genau dann äquivalent zueinander, wenn diese denselben Rang besitzen. Insbesondere werden die Äquivalenzklassen bezüglich der Äquivalenzrelation der Matrixäquivalenz durch die Matrizen

$$(A.1) \quad \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}, \quad \text{mit } r = 0, 1, \dots, \max\{m, n\},$$

(r Einsen auf der Diagonalen, sonst Nullen) repräsentiert.

Beweisskizze. Da Links- bzw. Rechtsmultiplikation mit invertierbaren Matrizen den Rang unverändert lässt, haben je zwei äquivalente Matrizen denselben Rang. Ferner haben die oben notierten Matrizen jeweils Rang r . Zum Beweis des Satzes genügt es daher einzusehen, dass jede Matrix $A \in K^{m \times n}$ zu einer der oben angegebenen Matrizen äquivalent ist.

Wir begnügen uns damit, die zugehörigen Argumente lediglich grob zu skizzieren. Wir modifizieren den aus der *Linearen Algebra 1* [3] bekannten Algorithmus 4.4.10 zur Berechnung der Inversen einer Matrix. Dieser startet (im Fall $n = m$) mit der erweiterten Matrix

$$(A \mid E_m) \in K^{m \times (m+m)}$$

(E_n ist die $n \times n$ -Einheitsmatrix) und überführt diese durch sukzessive Zeilenumformungen in die Form

$$(A' \mid B),$$

mit einer Matrix A' in reduzierter Zeilenstufenform. Ist nun $A' = E_m$, so gilt $B = A^{-1}$ und ist $A' \neq E_n$, so ist A nicht invertierbar.

Wir kehren nun zu der Situation mit $A \in K^{m \times n}$ für nicht notwendigerweise identische m und n zurück. Wir betrachten

$$\left(\begin{array}{c|c} E_n & \\ \hline A & E_m \end{array} \right) \in K^{(n+m) \times (n+m)},$$

und bringen diese Matrix durch Zeilenumformungen im unteren Zeilenblock und Spaltenumformungen im linken Spaltenblock auf die Form

$$\left(\begin{array}{c|c} T & \\ \hline A' & S \end{array} \right),$$

wobei A' von der Form (A.1) sei. Dies ist immer möglich, denn wie aus Algorithmus 4.4.10 bekannt, genügen zunächst Zeilenumformungen, um A auf reduzierte Zeilenstufenform zu bringen. Im Anschluss lässt sich durch Abziehen geeigneter Vielfacher der Spalten voneinander (in geeigneter Reihenfolge) erreichen, dass alle Einträge rechts der Stufen (deren Anfang gemäß Reduzierung stets auf 1 normiert

ist) zu Null werden. Zum Schluss erhält man durch Spaltenvertauschung hieraus tatsächlich eine Matrix der Form (A.1).

Im Zuge der Besprechung von Algorithmus 4.4.10 hatte man sich überlegt, dass elementare Zeilenumformungen durch Linksmultiplikation mit geeigneten Elementarmatrizen beschrieben werden können. Die Matrix S entsteht aus E_m demnach durch „Aufsammeln“ der zur Reduzierung von A mittels Zeilenumformungen benutzten Elementarmatrizen. Ganz analog werden Spaltenumformungen durch Rechtsmultiplikation mit Elementarmatrizen bewirkt. Die Matrix T entsteht somit als Produkt der bei der Reduzierung von A mittels Spaltenumformungen benutzten Elementarmatrizen. Da Elementarmatrizen invertierbar sind, erweisen sich T und S als Produkte von invertierbaren Matrizen selbst als invertierbar. Ferner überlegt man sich, dass $A' = SAT$ gilt. Das beendet den Beweis des Satzes. \square

A.1.3. Ähnlichkeit von Matrizen. Satz A.1 zeigt, dass Äquivalenz von Matrizen durch eine einzige numerische Invariante, nämlich den Rang *vollständig klassifiziert* wird. Im Folgenden betrachten wir zwei Beispiele, die zeigen, dass Äquivalenz von Matrizen für gewisse Untersuchungen zu grob ist.

Beispiel A.2 (Fibonacci-Zahlen). Die *Fibonacci-Folge* ist eine bekannte rekursiv definierte Folge $(F_n)_{n=0,1,\dots}$ mit den Startwerten $F_0 := 0$, $F_1 := 1$ und dem Bildungsgesetz

$$F_{n+1} = F_n + F_{n-1} \quad (n = 1, 2, \dots).$$

Die ersten paar Glieder dieser Folge lauten

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 114, \dots$$

und wir stellen uns die Aufgabe, deren Wachstum genauer zu verstehen.

Hierfür schreiben wir für $n \geq 1$

$$\begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix} = \begin{pmatrix} F_n + F_{n-1} \\ F_n \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} F_n \\ F_{n-1} \end{pmatrix}.$$

Mittels vollständiger Induktion folgert man hieraus mühelos

$$\begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^n \begin{pmatrix} F_1 \\ F_0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^n \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Schreiben wir A für die Matrix $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, so erhalten wir

$$F_n = \text{zweiter Eintrag der ersten Spalte von } A^n.$$

Es stellt sich daher die Aufgabe, die Potenzen A^n besser zu verstehen.

Ganz im Sinne unserer vorangegangenen Diskussion, sind wir zunächst versucht, die Matrix A durch eine „geeigneterere“ Matrix B zu ersetzen. Wir machen also den

Ansatz $B = SAT$ mit invertierbaren Matrizen $S, T \in \text{GL}_2(\mathbb{R})$ und einer Matrix $B \in \mathbb{R}^{2 \times 2}$. Dann ist

$$A^n = (S^{-1}BT^{-1})^n = \underbrace{(S^{-1}BT^{-1})(S^{-1}BT^{-1}) \cdots (S^{-1}BT^{-1})}_{n \text{ geklammerte Faktoren}}$$

was sich auch anhand des folgenden kommutativen Diagramms veranschaulichen lässt:

$$\begin{array}{ccccccccccc} \mathbb{R}^2 & \xrightarrow{\ell_A} & \mathbb{R}^2 & \xrightarrow{\text{id}} & \mathbb{R}^2 & \xrightarrow{\ell_A} & \mathbb{R}^2 & \xrightarrow{\text{id}} & \dots & \xrightarrow{\ell_A} & \mathbb{R}^2 \\ \ell_{T^{-1}} \downarrow \wr & & \ell_{S^{-1}} \uparrow \wr & & \ell_{T^{-1}} \downarrow \wr & & \ell_{S^{-1}} \uparrow \wr & & \vdots & & \ell_{S^{-1}} \uparrow \wr \\ \mathbb{R}^2 & \xrightarrow{\ell_B} & \mathbb{R}^2 & \xrightarrow{\ell_{T^{-1}S^{-1}}} & \mathbb{R}^2 & \xrightarrow{\ell_B} & \mathbb{R}^2 & \xrightarrow{\ell_{T^{-1}S^{-1}}} & \dots & \xrightarrow{\ell_B} & \mathbb{R}^2 \end{array}$$

Das Auftreten der „Zwischenschritte“ $T^{-1}S^{-1}$ (bzw. $\ell_{T^{-1}S^{-1}}$) sorgt hier für technische Schwierigkeiten und macht die Vereinfachung, die wir uns für unser Problem durch den Übergang von A zu B erhofft hatten, wieder zunichte. Um diese Schwierigkeiten auszuräumen betrachten wir eine speziellere Art von Matrixäquivalenz. Wir beschränken uns hierfür auf $B = SAT$ mit $S = T^{-1}$. Dann ist nämlich TS die Einheitsmatrix und $\ell_{TS} = \text{id}$. Das obige kommutative Diagramm kollabiert dann zu

$$\begin{array}{ccccccccccc} \mathbb{R}^2 & \xrightarrow{\ell_A} & \mathbb{R}^2 & \xrightarrow{\ell_A} & \mathbb{R}^2 & \xrightarrow{\ell_A} & \dots & \xrightarrow{\ell_A} & \mathbb{R}^2 \\ \ell_{T^{-1}} \downarrow \wr & & \ell_{T^{-1}} \downarrow \wr & & \ell_{T^{-1}} \downarrow \wr & & \vdots & & \ell_{T^{-1}} \downarrow \wr \\ \mathbb{R}^2 & \xrightarrow{\ell_B} & \mathbb{R}^2 & \xrightarrow{\ell_B} & \mathbb{R}^2 & \xrightarrow{\ell_B} & \dots & \xrightarrow{\ell_B} & \mathbb{R}^2 \end{array}$$

und wir gewinnen die Gleichung

$$A^n = TB^nT^{-1}.$$

Die Hoffnung besteht nun darin, dass wir A so wählen können, dass sich die Berechnung von A^n als hinreichend einfach erweist.

Wir bezeichnen mit $\varphi_+ > \varphi_-$ die beiden Lösungen der Gleichung $X^2 = X + 1$. (Mit der aus der Schule bekannten Lösungsformel für quadratische Gleichungen erhalten wir $\varphi_{\pm} = (1 \pm \sqrt{5})/2$.) Dann ist

$$A \begin{pmatrix} \varphi_{\pm} \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \varphi_{\pm} \\ 1 \end{pmatrix} = \begin{pmatrix} \varphi_{\pm} + 1 \\ \varphi_{\pm} \end{pmatrix} = \begin{pmatrix} \varphi_{\pm}^2 \\ \varphi_{\pm} \end{pmatrix} = \varphi_{\pm} \begin{pmatrix} \varphi_{\pm} \\ 1 \end{pmatrix},$$

wobei wir für das vorletzte Gleichheitszeichen die Gleichung $\varphi_{\pm} + 1 = \varphi_{\pm}^2$ benutzt haben. Die beiden Vektoren

$$v_+ = \begin{pmatrix} \varphi_+ \\ 1 \end{pmatrix} \quad \text{und} \quad v_- = \begin{pmatrix} \varphi_- \\ 1 \end{pmatrix}$$

erweisen sich als \mathbb{R} -linear unabhängig, denn aus einer hypothetischen Gleichung $v_+ = \lambda v_-$ mit $\lambda \in \mathbb{R}$ erhalte man mit Hinblick auf die zweite Komponente der Vektoren $\lambda = 1$, was aber wegen $\varphi_+ \neq \varphi_-$ zu einem Widerspruch führt.

Wir schreiben nun T für die Matrix mit den Spalten v_+ und v_- . Diese ist wegen der linearen Unabhängigkeit von v_+ und v_- invertierbar. Ferner sieht man leicht

$$T^{-1} = \frac{1}{\varphi_+ - \varphi_-} \begin{pmatrix} 1 & -\varphi_- \\ -1 & \varphi_+ \end{pmatrix},$$

sowie

$$B = T^{-1}AT = \begin{pmatrix} \varphi_+ & 0 \\ 0 & \varphi_- \end{pmatrix}$$

und also

$$B^n = \begin{pmatrix} \varphi_+^n & 0 \\ 0 & \varphi_-^n \end{pmatrix}.$$

Damit erhalten wir nun

$$\begin{aligned} F_n &= \text{zweiter Eintrag der ersten Spalte von } TB^nT^{-1} \\ &= \text{zweiter Eintrag von } TB^n \frac{1}{\varphi_+ - \varphi_-} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ &= \frac{1}{\varphi_+ - \varphi_-} \times \text{zweiter Eintrag von } \begin{pmatrix} \varphi_+ & \varphi_- \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \varphi_+^n & 0 \\ 0 & \varphi_-^n \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ &= \frac{1}{\varphi_+ - \varphi_-} \times \text{zweiter Eintrag von } \begin{pmatrix} \varphi_+ & \varphi_- \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \varphi_+^n \\ -\varphi_-^n \end{pmatrix}. \end{aligned}$$

Schließlich erhalten wir die sogenannte **Binetsche Formel**

$$(A.2) \quad F_n = \frac{\varphi_+^n - \varphi_-^n}{\varphi_+ - \varphi_-}.$$

Die in **Beispiel A.2** aufgeworfene Frage nach dem Wachstum der Fibonacci-Folge $(F_n)_n$ ist durch die Binetsche Formel (A.2) beantwortet. Der Schlüssel zum Erfolg bestand dabei in der Transformation $B = T^{-1}AT$.

Das obige Vorgehen lässt natürlich einige Fragen offen: Woher kommt die mysteriöse Gleichung $X^2 = X + 1$, deren Lösungen φ_{\pm} auf die Vektoren v_{\pm} führen und wie wäre man zu Vektoren v_{\pm} gelangt, ohne diese vorher zu kennen?

Sei K ein Körper. Zwei quadratische Matrizen $A, B \in K^{n \times n}$ nennen wir **ähnlich**, falls es eine invertierbare Matrix $T \in \text{GL}_n(K)$ mit $B = T^{-1}AT$ gibt.

Wir betrachten nun ein weiteres Beispiel, welches sich mittels Ähnlichkeitsuntersuchungen lösen lässt.

Beispiel A.3 (Lineare Differentialgleichungen, I). Wir betrachten den Raum $V = \mathcal{C}^{\infty}(\mathbb{R}, \mathbb{C})$ aller unendlich oft differenzierbaren Funktionen $\mathbb{R} \rightarrow \mathbb{C}$. Wir wollen einige spezielle Differentialgleichungen lösen. Als Prototyp betrachten wir die Gleichung $f' = f$. Gesucht ist hier eine Funktion $f \in V$, die (auf ganz \mathbb{R}) ihre eigene Ableitung ist. Bekanntermaßen handelt es sich bei der Exponentialfunktion $\exp: \mathbb{R} \rightarrow \mathbb{C}$ um eine Lösung. Darüber hinaus erhalten wir für jedes Skalar $\lambda \in \mathbb{C}$ auch die weitere Lösung $\lambda \exp$. Wir behaupten, dass dies sämtliche Lösungen liefert. Sei nämlich $f \in V$

mit $f' = f$ gegeben, so betrachten wir (Trick!) die Funktion $g: t \mapsto f(t) \exp(-t)$. Für diese gilt wegen der Produktregel

$$\begin{aligned} g'(t) &= f'(t) \exp(-t) + f(t) \exp(-t)(-1) \\ &= f(t) \exp(-t) - f(t) \exp(-t) = 0. \end{aligned}$$

Mit dem *Mittelwertsatz der Differentialrechnung* folgt hieraus, dass g konstant ist: In der Tat, ist $t > 0$, so liefert der besagte Mittelwertsatz die Existenz einer Zwischenstelle $\tau \in (0, t)$ mit

$$\frac{g(t) - g(0)}{t - 0} = g'(\tau) = 0.$$

Also ist $g(t) = g(0)$ für alle $t > 0$. Für $t < 0$ schließt man analog. Insgesamt gilt also $g(t) = g(0)$ für alle $t \in \mathbb{R}$. Daraus ergibt sich nun $f(t) = g(t) \exp(t) = \lambda \exp(t)$ mit $\lambda = g(0)$ für alle $t \in \mathbb{R}$, also $f = \lambda \exp$.

Beispiel A.4 (Lineare Differentialgleichungen, II). Wir betrachten erneut den Raum $V = \mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$ und studieren die Differentialgleichung

$$(A.3) \quad f'' + af' + bf = 0$$

für Konstanten $a, b \in \mathbb{C}$. Um die doppelte (jedenfalls formal) Ableitung zu entfernen setzen wir $g = f'$. Dann ist

$$\begin{aligned} (A.3) &\iff \begin{cases} g' = -ag - bf \\ f' = g \end{cases} \\ &\iff \begin{pmatrix} g' \\ f' \end{pmatrix} = \begin{pmatrix} -a & -b \\ 1 & 0 \end{pmatrix} \begin{pmatrix} g \\ f \end{pmatrix}. \end{aligned}$$

Sei nun $T \in \text{GL}_2(\mathbb{C})$ eine invertierbare Matrix. Wir definieren $u, v \in V$ durch die Beziehung

$$(A.4) \quad T \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} g \\ f \end{pmatrix}.$$

Aus der Linearität des Ableitens folgt

$$T \begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} g' \\ f' \end{pmatrix}.$$

Wir erhalten

$$(A.3) \iff T \begin{pmatrix} u' \\ v' \end{pmatrix} = \begin{pmatrix} -a & -b \\ 1 & 0 \end{pmatrix} T \begin{pmatrix} u \\ v \end{pmatrix} \iff \begin{pmatrix} u' \\ v' \end{pmatrix} = T^{-1} \begin{pmatrix} -a & -b \\ 1 & 0 \end{pmatrix} T \begin{pmatrix} u \\ v \end{pmatrix}.$$

Wir studieren nun den Fall, bei dem T so gewählt werden kann, dass das obige Matrixprodukt Diagonalgestalt hat, also:

$$T^{-1} \begin{pmatrix} -a & -b \\ 1 & 0 \end{pmatrix} T = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}.$$

Dann haben wir die beiden *entkoppelten* Differentialgleichungen $u' = \lambda u$ und $v' = \mu v$ zu lösen. Die zugehörige Lösungsgesamtheit ist uns bereits aus **Beispiel A.3** vollständig bekannt: $u(t) = \rho \exp(\lambda t)$, $v(t) = \varrho \exp(\mu t)$ für geeignete Skalare $\rho, \varrho \in \mathbb{C}$ und alle $t \in \mathbb{R}$. Die ursprünglich gesuchte Funktion f (und ihre Ableitung $g = f'$) erhalten wir dann aus **(A.4)**.

Leserinnen und Leser, welche die hier besprochene Lösungsmethode erproben wollen, mögen die Ergebnisse aus **Beispiel A.2** zur Lösung der Differentialgleichung $f'' - f' - f = 0$ verwenden.

A.2. Polynome

A.2.1. Hintergründige Absicht. Wir betrachten eine quadratische Matrix $A \in K^{n \times n}$ über einem Körper K . In [4, Korollar 6.1.11] wurde festgestellt, dass die Eigenwerte $\lambda \in K$ von A genau die λ mit $\det(\lambda E_n - A) = 0$ sind.

Beispiel A.5. Wir arbeiten über einem beliebigen Körper K . Es sei $\lambda \in K$.

$$(1) \text{ Für } A = \begin{pmatrix} 0 & & \\ & 0 & \\ & & 1 \end{pmatrix} \text{ ist } \det(\lambda E_3 - A) = \det \begin{pmatrix} \lambda & & \\ & \lambda & \\ & & \lambda - 1 \end{pmatrix} = \lambda^2(\lambda - 1).$$

Die Eigenwerte von A sind genau 0 und 1. Die zugehörigen Eigenräume sind

$$\text{Eig}(A, 0) = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle \quad \text{und} \quad \text{Eig}(A, 1) = \left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle.$$

$$(2) \text{ Für } B = \begin{pmatrix} 0 & & \\ & 1 & \\ & & 1 \end{pmatrix} \text{ ist } \det(\lambda E_3 - B) = \det \begin{pmatrix} \lambda & & \\ & \lambda - 1 & \\ & & \lambda - 1 \end{pmatrix} = \lambda(\lambda - 1)^2.$$

Die Eigenwerte von A sind genau 0 und 1. Die zugehörigen Eigenräume sind

$$\text{Eig}(B, 0) = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle \quad \text{und} \quad \text{Eig}(B, 1) = \left\langle \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle.$$

Man beachte, dass die „polynomialen Ausdrücke“ $\lambda^2(\lambda - 1)$ und $\lambda(\lambda - 1)^2$ in **Beispiel A.5** für dieselben $\lambda \in K$ (nämlich genau 0 und 1) den Wert 0 liefern. Dennoch sind die Eigenräume von A und B (jeweils zum selben Eigenwert) grundlegend verschieden, wie man anhand deren Dimensionen klar erkennt.

Tatsächlich liegt in den obigen polynomialen Ausdrücken noch etwas mehr Information über die Eigenräume der betrachteten Matrizen versteckt. Diese Information reicht zwar ab Dimension zwei auch nicht aus, um eine abschließende Antwort auf die Frage nach den Eigenräumen zu liefern, ist zum Ausbau der Theorie aber dennoch wichtig.

Bemerkung A.6. Wählt man in **Beispiel A.5** den Körper $K = \mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$ mit den zwei Elementen 0 und 1, so liefert die Abbildung

$$K \rightarrow K, \quad \lambda \mapsto \det(\lambda E_3 - A) = \lambda^2(\lambda - 1)$$

die Nullabbildung (denn setzt man 0 für λ ein, so annulliert der Faktor λ^2 das Produkt $\lambda^2(\lambda - 1)$ und setzt man 1 für λ ein, so wird das Produkt vom Faktor $\lambda - 1$ annulliert). Gleichwohl liefert aber auch

$$K \rightarrow K, \quad \lambda \mapsto \det(\lambda E_3 - B) = \lambda(\lambda - 1)^2$$

die Nullabbildung (via analoger Argumentation). In den Abbildungen $K \rightarrow K, \lambda \mapsto \det(\dots)$, findet sich im Allgemeinen also auch keine weitere Information über die Größe der Eigenräume abgesehen von deren etwaiger Nichttrivialität.

A.2.2. Polynome in intuitiver Vorstellung. Motiviert durch die vorangegangenen Überlegungen sind wir jetzt bemüht, einen theoretischen Rahmen dafür zu schaffen in welchem wir etwa den Ausdruck $\lambda^2(\lambda - 1)$ und $\lambda(\lambda - 1)^2$ unterscheiden können und auch allgemeine Ergebnisse über derartige Ausdrücke gewinnen können.

Speziell handelt es sich bei diesen Ausdrücken um das, was man schon aus der Schule als „Polynome“ kennt. **Bemerkung A.6** lehrt uns aber, dass wir uns unter λ eher etwas allgemeineres als ein Element des zugrunde liegenden Körpers vorstellen sollten. Speziell stellen wir uns unter „Polynomen über K “ formale Ausdrücke der Form

$$a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0,$$

mit $n \in \mathbb{N}_0$, sogenannten „Koeffizienten“ $a_n, \dots, a_0 \in K$ und einer „Unbestimmten“ X vor. Hier übernimmt X die Rolle von λ , aber wir wechseln an dieser Stelle die Bezeichner, um anzudeuten, dass wir uns X nicht als ein Element von K vorstellen. Die Bezeichnung X ist hier natürlich nicht in Stein gemeißelt und darf auch durch andere Bezeichnungen wie x, T, t oder beliebige andere verständliche Symbole ersetzt werden. In diesen Notizen schreiben wir allerdings stets X für die Unbestimmte, sofern nichts Gegenteiliges explizit gesagt wird.

Beispiel. Polynome über \mathbb{R} sind etwa $X^5 - 7X^2 + 8$, oder $X^2 - 3X$.

Wir werden uns in Zukunft auch mehrfach genötigt sehen, mit Polynomen über K „rechnen“ zu müssen. Polynomen über K sollte man also addieren und multiplizieren können. Dabei ist es natürlich wünschenswert, dass die üblichen Rechenregeln, wie beispielsweise Kommutativität bei Addition und Multiplikation, Distributivgesetze und dergleichen gelten. Kurzum, man wünscht sich, dass die Menge aller Polynome über K die Struktur eines Ringes trägt.

A.2.3. Formale Definition von Polynomen. Nun formalisieren wir unsere oben geäußerten Wünsche. Das, was wir als Polynom

$$(A.5) \quad f = a_n X^n + a_{n-1} X^{n-1} + \dots + a_1 X + a_0,$$

(über einem Körper K) verstehen wollen, modellieren wir nun durch die Abbildung $f: \mathbb{N}_0 \rightarrow K$, welche jedem $j \in \mathbb{N}_0$ den j -ten Koeffizienten a_j zuordnet. Dabei verständigen wir uns darauf, dass die Koeffizienten a_{n+1}, a_{n+2}, \dots , welche augenscheinlich nicht in (A.5) auftreten, als Null (in K) zu verstehen sind.

Beispiel. Das Polynom $5X^3 - 2X + 7$ modellieren wir formal als die Abbildung $\mathbb{N}_0 \rightarrow K$

$$0 \mapsto 7, \quad 1 \mapsto 2, \quad 2 \mapsto 0, \quad 3 \mapsto 5, \quad 4 \mapsto 0, \quad 5 \mapsto 0, \quad \dots$$

Wir betrachten nun die Menge der Polynome

$$(A.6) \quad K[X] = \{\text{Abbildungen } f: \mathbb{N}_0 \rightarrow K \text{ mit } \exists n \in \mathbb{N}_0 \forall j \geq n: f[j] = 0\},$$

(sprich: „ K **adjungiert** X “) wobei wir hier $f[j]$ statt $f(j)$ für den Wert der Abbildung $f: \mathbb{N}_0 \rightarrow K$ bei $j \in \mathbb{N}_0$ geschrieben haben. (Das tun wir, da wir später auch $f(x)$ für $x \in K$ definieren werden und damit etwas anderes meinen. — Siehe **Bemerkung A.11.**) Das „ X “ in „ $K[X]$ “ sollte zu diesem Zeitpunkt lediglich als Notation angesehen werden, bekommt aber gleich noch eine echte Bedeutung, die uns erlaubt X auch als Element von $K[X]$ anzusehen (siehe (A.10)).

Die Elemente von $K[X]$ bezeichnen wir als **Polynome**. Für ein Polynom $f \in K[X]$ und $j \in \mathbb{N}_0$ bezeichnen wir $f[j]$ als den j -ten **Koeffizienten** von f .

Die Menge $K[X]$ statten wir nun mit einer Addition und Multiplikation, sodass wir einen Ring bekommen (siehe **Lemma A.10**). Für $f, g \in K[X]$ definieren wir

$$(A.7) \quad \begin{aligned} f + g &= \text{die Abbildung } \begin{cases} \mathbb{N}_0 \rightarrow K, \\ m \mapsto f[m] + g[m], \end{cases} \\ f \cdot g &= \text{die Abbildung } \begin{cases} \mathbb{N}_0 \rightarrow K, \\ m \mapsto \sum_{j=0}^m f[m-j]g[j]. \end{cases} \end{aligned}$$

Lemma A.7. Für $f, g \in K[X]$ liefert es sich bei den in (A.7) definierten Abbildungen $f + g: \mathbb{N}_0 \rightarrow K$ und $f \cdot g: \mathbb{N}_0 \rightarrow K$ tatsächlich um Polynome über K im Sinne von (A.6).

Beweis. Es gilt einzusehen, dass $f + g$ und $f \cdot g$ (als Abbildungen von \mathbb{N}_0 nach K) für alle hinreichend großen Argumente den Wert $0 \in K$ liefern. Hier wählen wir $n, r \in \mathbb{N}_0$ derart, dass $f[i] = 0$ für alle $i > n$ und $g[j] = 0$ für alle $j > r$ gilt. Wir behaupten, dass dann $(f \cdot g)[m] = 0$ für alle $m > n+r$ gilt. In der Tat ist für $m > n+r$ ja definitionsgemäß

$$(f \cdot g)[m] = \sum_{j=0}^m f[m-j]g[j].$$

Für einen beliebigen Summanden $f[m-j]g[j]$ in der obigen Summe gibt es zwei Möglichkeiten: $j \leq r$ (und somit $f[m-j] = 0$ wegen $m-j \geq m-r > n+r-r = n$), oder $j > r$ (und somit $g[j] = 0$). In beiden Fällen ist einer der Faktoren im Produkt $f[m-j]g[j]$ gleich Null und somit das Produkt selbst gleich Null. Wir sehen also, dass die $(f \cdot g)[m]$ definierende Summe bereits summandenweise gleich Null ist.

Nun zur Summe $f + g$. Wir wählen n und r wie oben und behaupten, dass $(f + g)[m] = 0$ für alle $m > \max\{n, r\}$ gilt. In der Tat ist $(f + g)[m] = f[m] + g[m] = 0$, da beide Summanden gleich Null sind; der Erste wegen $m > n$ und der Zweite wegen $m > r$.

Als *Zusatz* für spätere Referenzierung im Rahmen des Beweises zu **Proposition A.9** beachten wir noch Folgendes: Haben n und r zusätzlich zu ihrer oben geforderten Eigenschaft noch die Eigenschaft, dass $f[n]$ und $g[r]$ *beide* nicht gleich 0 sind, so erhalten wir

$$(A.8) \quad (f \cdot g)[n + r] = f[n]g[r] \neq 0. \quad \square$$

Jedem $\lambda \in K$ ordnen wir die Abbildung

$$(A.9) \quad \mathbb{N}_0 \rightarrow K, \quad j \mapsto \begin{cases} \lambda & \text{falls } j = 0, \\ 0 & \text{falls } j \geq 1, \end{cases}$$

zu und schreiben für diese ebenfalls λ . Bei dieser Abbildung λ handelt es sich offensichtlich um ein Element von $K[X]$. Die Elemente von $K[X]$, die wir so bekommen, nennen wir *konstante* Polynome.

Man prüft leicht nach, dass die in (A.9) erklärte Zuordnung von $\lambda \in K$ zu dem Polynom $\lambda \in K[X]$ mit der Körperstruktur auf K verträglich ist:

Lemma A.8. *Erst in K addieren oder multiplizieren und dann gemäß (A.9) zuordnen, oder alternativ erst gemäß (A.9) zuordnen und dann mittels der zugehörigen Verknüpfung in (A.7) addieren respektive multiplizieren liefert dasselbe Ergebnis.*

Beweis. Wir führen dies exemplarisch für die Multiplikation durch. (Die Addition ist einfacher.) Zur klareren Unterscheidung benutzen wir in der folgenden Rechnung hoffentlich selbsterklärend beschriftete Indizes. Für $\lambda, \mu \in K$ in gilt

$$\begin{aligned} & \lambda_{\text{als Polynom}} \cdot_{K[X]} \mu_{\text{als Polynom}} \\ &= \text{die Abbildung} \begin{cases} \mathbb{N}_0 \rightarrow K, \\ m \mapsto \sum_{j=0}^m (\lambda_{\text{als Polynom}}[m-j] \cdot_K \mu_{\text{als Polynom}}[j]). \end{cases} \end{aligned}$$

Für $m = 0$ besteht die obige Summe nur aus einem Summanden, nämlich

$$(\lambda_{\text{als Polynom}}[0] \cdot_K \mu_{\text{als Polynom}}[0]) = \lambda \cdot_K \mu.$$

Für $m \geq 0$ ist in jedem auftretenden Summanden der Summe entweder $m - j \geq 1$ oder $j \geq 1$. Mit Hinblick auf (A.9) ist dann also für jeden Summanden stets einer der Faktoren gleich 0 (in K) und somit auch das Produkt beider Faktoren. Dann sind also alle Summanden gleich 0 und somit auch der Wert der Summe. Wir erhalten

also

$$\begin{aligned} \lambda_{\text{als Polynom}} \cdot_{K[X]} \mu_{\text{als Polynom}} &= \text{die Abbildung } \begin{cases} \mathbb{N}_0 \rightarrow K, \\ m \mapsto \begin{cases} \lambda \cdot_K \mu & \text{falls } m = 0, \\ 0 & \text{falls } m \geq 1, \end{cases} \end{cases} \\ &= (\lambda \cdot_K \mu)_{\text{als Polynom}}. \end{aligned} \quad \square$$

Das Nullelement $\lambda = 0$ von K fassen wir nun gemäß der eben erklärten Art und Weise als Polynom auf (siehe (A.9)) und schreiben, wie schon erwähnt, einfach $0 \in K[X]$ dafür und nennen dieses das **Nullpolynom**.

Wir definieren nun den **Grad** $\deg f$ eines jeden Polynoms $f \in K[X]$. Ist $f = 0$ das Nullpolynom, so definieren wir $\deg 0 := -\infty$. Für $f \neq 0$ definieren wir

$$\deg f := \max\{n \in \mathbb{N} : f[n] \neq 0\}.$$

Die Fallunterscheidung ist nötig, da für $f = 0$ die Menge $\{n \in \mathbb{N} : f[n] \neq 0\}$ leer ist und also kein Maximum besitzt. Für alle anderen Polynome f ist die fragliche Menge natürlich nicht leer, aber auch beschränkt dank unserer Definition von Polynomen (siehe (A.6)). Das zur Definition des Grades $\deg f$ (mit $f \neq 0$) hergenommene Maximum existiert also.

Beispiel. Fasst man $\lambda \in K$ gemäß (A.9) als Polynom auf, so gilt $\deg \lambda = -\infty$, falls $\lambda = 0$ ist, und $\deg \lambda = 0$ sonst.

Proposition A.9 (Gradformel). Für $f, g \in K[X]$ gilt die sogenannte **Gradformel**:

$$\deg(f \cdot g) = \deg(f) + \deg(g).$$

Dabei ist, falls einer der Terme gleich $-\infty$ ist, die Formel im Sinne von $-\infty = (-\infty) + k = k + (-\infty) = (-\infty) + (-\infty)$ (für $k \in \mathbb{N}_0$) zu verstehen. Zudem gilt die **Gradabschätzung**

$$\deg(f + g) \leq \max\{\deg(f), \deg(g)\}.$$

Beweis. Wir benutzen die Bezeichnungen n und r aus dem Beweis von Lemma A.7. Ist $f, g \neq 0$, so kann man $n = \deg f$ und $r = \deg g$ wählen. Der Beweis von Lemma A.7 liefert dann $(f \cdot g)[m] = 0$ für alle $m > n + r = \deg(f) + \deg(g)$. Andererseits liefert der Zusatz aus dem Beweis von Lemma A.7, dass $(f \cdot g)[m + r]$ ungleich 0 ist. Daraus folgt schon die Gradformel $\deg(f \cdot g) = \deg(f) + \deg(g)$.

Ist hingegen eines der Polynome f oder g (oder beide) gleich Null, so sieht man mit Blick auf (A.7) direkt $f \cdot g = 0$ und die Gradformel gilt dann im Sinne der in der Proposition beschriebenen Arithmetik mit $-\infty$.

Die Gradabschätzung folgt aus der Besprechung von $f + g$ im Beweis von Lemma A.7. □

Nun geben wir auch dem bisher noch mysteriösen „ X “ in der Schreibweise „ $K[X]$ “ eine konkrete Bedeutung: Unter dem Polynom X (über K) verstehen wir das Polynom

$$(A.10) \quad \mathbb{N}_0 \rightarrow K, \quad j \mapsto \begin{cases} 1 & \text{falls } j = 1, \\ 0 & \text{falls } j \neq 1. \end{cases}$$

Das folgende Resultat ist eine ausgebreitete Fassung von [4, Lemma 6.2.7].

Lemma A.10.

(1) Für jeden Körper K bildet $K[X]$ mit den oben definierten zweistelligen Verknüpfungen $+$ und \cdot einen kommutativen Ring.

Das Nullelement bezüglich dieser Ringstruktur ist das oben definierte Nullpolynom. Das Einselement ist das Polynom 1, welches man gemäß (A.9) aus dem Einselement 1 von K erhält.

(2) Für jedes Polynom $f \in K[X]$ existieren $n \in \mathbb{N}_0$ und $a_0, a_1, \dots, a_n \in K$ mit

$$(A.11) \quad f = \sum_{j=0}^n a_j X^j.$$

Darüber hinaus ist diese Darstellung im folgenden Sinne eindeutig: Es gilt $a_j = f[j]$ für $j = 0, 1, \dots, n$ und ist f nicht das Nullpolynom, so folgt $n \geq \deg f$.¹

Beweis. Wie im Beweis von Lemma A.8 fügen wir wieder an einigen Stellen hoffentlich selbsterklärende Indizes ein, welche zur besseren Unterscheidung der oben definierten Konzepte gedacht sind.

Der Nachweis von (1) ist eine **arbeitsintensive Strafarbeit**. Den Leserinnen und Lesern sei versichert, dass der diese Aufgabe ausarbeitende Dozent bei der Ausarbeitung derselben genau so wenig Freude hatte, wie es Freude macht, diese Zeilen zu lesen. Wir notieren mit Blick auf [3, Definition 2.2.1], was hier zu zeigen ist:

(1.a) $(K[X], +)$ ist eine abelsche Gruppe, deren neutrales Element das Nullpolynom 0 ist:

(1.a.i) Die Addition $+$ auf $K[X]$ ist assoziativ: Für alle $f, g, h \in K[X]$ gilt $(f + g) + h = f + (g + h)$.

(1.a.ii) Das Nullpolynom 0 ist neutral bezüglich der Addition auf $K[X]$: Für alle $f \in K[X]$ gilt $0 + f = f = f + 0$.

(1.a.iii) Jedes Element von $K[X]$ besitzt ein additiv inverses Element: Für jedes $f \in K[X]$ gibt es ein $g \in K[X]$ mit $f + g = 0 = g + f$.

(1.a.iv) Die Addition $+$ auf $K[X]$ ist kommutativ: Für alle $f, g \in K[X]$ gilt $f + g = g + f$.

(1.b) Die Multiplikation \cdot auf $K[X]$ ist assoziativ: Für alle $f, g, h \in K[X]$ gilt $(f \cdot g) \cdot h = f \cdot (g \cdot h)$.

¹Die einzige Uneindeutigkeit in dieser Darstellung besteht also darin, dass man die Summe „nach Hinten hinaus“ um beliebig viele Terme mit Koeffizienten $a_j = 0$ verlängern kann.

(1.c) Das Polynom 1 (im Sinne von (A.9)) ist neutral bezüglich der Multiplikation auf $K[X]$: Es gilt $1 \cdot f = f = f \cdot 1$ für alle $f \in K[X]$.

(1.d) Die Addition und Multiplikation auf $K[X]$ gehorchen den Distributivgesetzen: Für alle $f, g, h \in K[X]$ gilt

$$f \cdot (g + h) = f \cdot g + f \cdot h \quad \text{und} \quad (f + g) \cdot h = f \cdot h + g \cdot h.$$

(1.e) Die Multiplikation auf $K[X]$ ist kommutativ: Für alle $f, g \in K[X]$ gilt $f \cdot g = g \cdot f$.

Die oben grau gedruckten Aussagen folgen mittels (1.a.iv) bzw. (1.e) aus deren schwarz gedrucktem Pendant. Wir dürfen bei der Abarbeitung der obigen Aussagen also auf die grau gedruckten Teile verzichten.

Im Folgenden seien stets $f, g, h \in K[X]$ und m, n und j bezeichnen stets Elemente von \mathbb{N}_0 .

(1.a.i): Bei $(f + g) + h$ handelt es sich um die Abbildung von \mathbb{N}_0 nach K mit

$$\begin{aligned} ((f + g) + h)[m] &= (f + g)[m] +_K h[m] = (f[m] +_K g[m]) +_K h[m] \\ &= f[m] +_K (g[m] +_K h[m]) = f[m] +_K (g + h)[m] \\ &= (f + (g + h))[m], \end{aligned}$$

wobei wir im Schritt von der ersten zur zweiten Zeile die Assoziativität der Addition in K verwendet haben. Also gilt $(f + g) + h = f + (g + h)$.

(1.a.ii): Es gilt

$$(0_{\text{als Polynom}} + f)[n] = 0_{\text{als Polynom}}[n] +_K f[n] = 0 +_K f[n] = f[n],$$

wobei wir im letzten Schritt die additive Neutralität des Nullelements in K benutzt haben.

(1.a.iii): Zu $f \in K[X]$ betrachten wir $g: \mathbb{N}_0 \rightarrow K$ mit $g[m] = -f[m]$ (additiv inverses Element zu $f[m] \in K$ bezüglich der Addition in K) für alle m . Bei g handelt es sich natürlich um ein Polynom, denn $f[m] = 0$ für alle hinreichend großen m und somit $g[m] = -f[m] = -0 = 0$ für eben jene m . Nun gilt

$$(f + g)[m] = f[m] +_K g[m] = f[m] +_K (-f[m]) = 0 = 0_{\text{als Polynom}}[n],$$

also $f + g = 0_{\text{als Polynom}}$.

(1.a.iv): Wir haben

$$(f + g)[m] = f[m] +_K g[m] = g[m] +_K f[m] = (g + f)[m],$$

wegen der Kommutativität der Addition in K . Also ist $f + g = g + f$.

(1.b): Zunächst einmal ist

$$\begin{aligned} ((f \cdot g) \cdot h)[m] &= \sum_{j=0}^m (f \cdot g)[m-j] h[j] \\ &= \sum_{j=0}^m \left(\sum_{i=0}^{m-j} f[(m-j)-i] \cdot_K g[i] \right) \cdot_K h[j] \end{aligned}$$

$$= \sum_{j=0}^m \sum_{i=0}^{m-j} f[m - (j + i)] \cdot_K (g[i] \cdot_K h[j]),$$

wobei wir im letzten Schritt die Distributivität der Multiplikation in K , sowie Assoziativität von sowohl Addition und Multiplikation in K benutzt haben.² Durch wiederholte Berufung auf Assoziativität und Kommutativität in K können wir die auftretenden Summanden nach gleichen Werten von $k := j + i$ gruppieren. Dabei durchläuft k unter Berücksichtigung der Summationsgrenzen die Werte $0, 1, \dots, m$. Anschließendes Ausklammern von $f[m - k]$ (unter Berufung auf Distributivität in K) liefert nun

$$\begin{aligned} ((f \cdot g) \cdot h)[m] &= \sum_{k=0}^m f[m - k] \cdot_K \left(\sum_{j=0}^k g[k - j] \cdot_K h[j] \right) \\ &= \sum_{k=0}^m f[m - k] \cdot_K (g \cdot h)[k] \\ &= (f \cdot (g \cdot h))[m]. \end{aligned}$$

Also ist $(f \cdot g) \cdot h = f \cdot (g \cdot h)$, wie behauptet.

(1.c): Wir haben

$$(1_{\text{als Polynom}} \cdot f)[m] = \sum_{j=0}^m 1_{\text{als Polynom}}[m - j] \cdot_K f[j].$$

In der obigen Summe ist für alle $j < m$ natürlich $m - j > 0$ und somit $1_{\text{als Polynom}}[m - j] = 0$. Die zugehörigen Summanden sind also alle gleich 0. Der Summand für $j = m$ ist hingegen $1_{\text{als Polynom}}[m - m] \cdot_K f[m] = 1 \cdot_K f[m] = f[m]$. Insgesamt ergibt sich somit $(1_{\text{als Polynom}} \cdot f)[m] = f[m]$ und also $1_{\text{als Polynom}} \cdot f = f$.

(1.d): Unter Anwendung der Rechengesetze in K erhalten wir

$$\begin{aligned} f \cdot (g + h)[m] &= \sum_{j=0}^m f[m - j] \cdot_K (g + h)[j] \\ &= \sum_{j=0}^m f[m - j] \cdot_K (g[j] +_K h[j]) \\ &= \sum_{j=0}^m (f[m - j] \cdot_K g[j] +_K f[m - j] \cdot_K h[j]) \\ &= \sum_{j=0}^m (f[m - j] \cdot_K g[j]) +_K \sum_{j=0}^m (f[m - j] \cdot_K h[j]) \end{aligned}$$

²Tatsächlich benutzen wir allein durch den Gebrauch der Summennotation schon stillschweigend die Assoziativität in K , da wir sonst bei jeder Summe \sum penibel zu spezifizieren hätten, in welcher Reihenfolge die Summanden zu Paaren zusammenzuklammern sind.

$$\begin{aligned}
 &= (f \cdot g)[m] +_K (f \cdot h)[m] \\
 &= ((f \cdot g) + (f \cdot h))[m].
 \end{aligned}$$

Das zeigt $f \cdot (g + h) = (f \cdot g) + (f \cdot h)$.

(1.e): Wir haben

$$(f \cdot g)[m] = \sum_{j=0}^m f[m-j] \cdot_K g[j].$$

Für $j = 0, \dots, m$ durchläuft $k := m - j$ die Werte $m, \dots, 0$. Durch Umordnung der Summanden (vermöge Kommutativität der Addition in K) erhalten wir

$$(f \cdot g)[m] = \sum_{k=0}^m f[k] \cdot_K g[m-k] = \sum_{k=0}^m g[m-k] \cdot_K f[k] = (g \cdot f)[m].$$

Also gilt $f \cdot g = g \cdot f$.

Das beendet schließlich den Beweis von (1).

Wir wenden uns nun dem Beweis von (2) zu. Wir bemerken hierzu zunächst, dass das Produkt $a_j X^j$ in der Gleichung (A.11) streng genommen von der Form „ein Element von K mal das Element $X^j \in K[X]$ “ ist. Dabei ist natürlich gemeint, dass man $a_j \in K$ gemäß der in (A.9) beschriebenen Art und Weise als Polynom aufzufassen hat und dann dieses Polynom mit dem Polynom X^j gemäß (A.7) zu multiplizieren hat. Wir erinnern an die Definition (A.10) von X und wollen nun die Potenzen X^j (als Funktionen von \mathbb{N}_0 nach K) bestimmen. Konventionsgemäß ist unter der Nullten Potenz eines Ringelements stets das neutrale Element des Rings zu verstehen. Dieses ist in $K[X]$, wie wir bereits wissen, genau $1_{\text{als Polynom}}$ im Sinne von (A.9). Also:

$$X^0[m] = \begin{cases} 1 & \text{falls } m = 0, \\ 0 & \text{falls } m \geq 1. \end{cases}$$

Wir behaupten nun, dass

$$X^j[m] = \begin{cases} 1 & \text{falls } m = j, \\ 0 & \text{falls } m \neq j, \end{cases}$$

für alle $m \in \mathbb{N}_0$ gilt. Das beweisen wir mittels Induktion über $j = 0, 1, 2, \dots$; Der Induktionsanfang ist dank der oben erklärten Konvention gesichert. Für den Induktionsschritt von j zu $j + 1$ rechnen wir: für $m \in \mathbb{N}_0$ ist

$$X^{j+1}[m] = (X \cdot X^j)[m] = \sum_{k=0}^m X[k] \cdot_K X^j[m-k].$$

Wegen der Definition (A.10) von X erweist sich diese Summe als $= X^j[m-1]$. Gemäß der Induktionsvoraussetzung wissen wir allerdings, dass X^j als Funktion von \mathbb{N}_0 nach K auf allen Argumenten den Wert 0 liefert, außer auf dem Argument j , für welches

der Wert 1 geliefert wird. Nun ist aber $m - 1$ genau dann gleich j , wenn $m = j + 1$ gilt. Also haben wir

$$X^{j+1}[m] = \begin{cases} 1 & \text{falls } m = j + 1, \\ 0 & \text{falls } m \neq j + 1, \end{cases}$$

wie benötigt.

Mit dieser Vorüberlegung sehen wir nun:

$$\left(\sum_{j=0}^n a_j X^j \right) [m] = \begin{cases} a_m & \text{falls } m \leq n, \\ 0 & \text{falls } m > n. \end{cases}$$

Mit dieser expliziten Beschreibung der rechten Seite von (A.11) ergibt sich nun leicht die in (2) behauptete Darstellbarkeit von f und die zugehörige Eindeutigkeitsaussage. Die Details dazu überlassen wir den Leserinnen und Lesern. \square

Wir erklären nun, was wir mit dem Auswerten eines Polynoms $f \in K[X]$ meinen. Sei $\lambda \in K$ gegeben. Dann definieren wir die **Auswertung** $f(\lambda)$ von f bei λ durch

$$(A.12) \quad f(\lambda) := \sum_j f[j] \lambda^j,$$

wobei $j \in \mathbb{N}_0$ die (höchstens endlich vielen) Zahlen durchlaufe, für welche $f[j] \neq 0$ gilt. Falls es gar keine solchen j gibt (also f das Nullpolynom ist), so interpretieren wir den Wert der Summe als 0.

Beispiel. Wir betrachten das Polynom $f = 5X^3 - 6X + 2 \in \mathbb{R}[X]$ und werten dieses bei -1 aus:

$$f(-1) = 5(-1)^3 - 6(-1) + 2(-1)^0 = -5 + 6 + 2 = 3.$$

Auswertung bei der Kreiszahl $\pi \in \mathbb{R}$ liefert dafür

$$f(\pi) = 5\pi^3 - 6\pi + 2\pi^0 = 5\pi^3 - 6\pi + 2.$$

(Diesen Ausdruck könnten wir allenfalls noch durch numerische Approximation eingrenzen, aber einfacher exakt ausdrücken können wir diesen Wert hier nicht.)

Bemerkung A.11. Die in (A.12) definierte Auswertung $f(\lambda)$ vom Polynom $f \in K[X]$ bei $\lambda \in K$ ist der Grund, weshalb wir seit Beginn der Einführung von Polynomen als Abbildungen von \mathbb{N}_0 nach K stets $f[m]$ für den Wert von $f: \mathbb{N}_0 \rightarrow K$ bei $m \in \mathbb{N}_0$ geschrieben haben. Man könnte sich darauf berufen, dass wegen der Unterscheidung zwischen \mathbb{N}_0 und K zwar wenig Verwechslungsgefahr zwischen $f[m]$ und $f(\lambda)$ besteht, doch eine solche Verwechslungsgefahr schleicht sich sofort ein, wenn man Körper wie \mathbb{Q} , \mathbb{R} , oder \mathbb{C} betrachtet, welche ja \mathbb{N}_0 enthalten. (Ganz unabhängig schreiben wir auch konventionsgemäß 2 für $1 + 1$ auch, wenn wir in einem beliebigen Körper arbeiten. So ist etwa im Körper $K = \mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$ natürlich $1_K + 1_K = 0_K$, also „ $2 = 1 + 1 = 0$ “ und in dieser Schreibweise ist ohne zusätzlichen Kontext nicht ersichtlich, ob wir die ganze Zahl 2, oder $1_K + 1_K$ meinen.)

Wir geben ein Beispiel: Für $f = 3X^2 - 5 \in \mathbb{R}[X]$ ist

$$f[2] = 3 \neq 7 = 3 \cdot 2^2 - 5 = f(2),$$

aber (vielleicht verwirrender Weise?)

$$f[0] = -5 = 3 \cdot 0^2 - 5 = f(0).$$

Lemma A.12. *Das Auswerten von Polynomen ist verträglich mit der Addition und Multiplikation von Polynomen: Für alle $f, g \in K[X]$ und alle $\lambda \in K$ gilt*

$$(f + g)(\lambda) = f(\lambda) +_K g(\lambda)$$

und

$$(f \cdot g)(\lambda) = f(\lambda) \cdot_K g(\lambda).$$

Beweis. Wir behandeln nur die Multiplikation und beschränken uns auf den Fall, dass keines der beiden Polynome das Nullpolynom ist. (Das erlaubt uns Summationsindizes bis zum Polynomgrad zu erstrecken, was wegen $\deg 0 = -\infty$ sonst Probleme verursachen würde.) Dank Distributivität ist

$$\begin{aligned} f(\lambda) \cdot_K g(\lambda) &= \left(\sum_{m=0}^{\deg f} f[m] \lambda^m \right) \left(\sum_{k=0}^{\deg g} g[k] \lambda^k \right) = \sum_{m=0}^{\deg f} f[m] \lambda^m \sum_{k=0}^{\deg g} g[k] \lambda^k \\ &= \sum_{m=0}^{\deg f} \sum_{k=0}^{\deg g} f[m] g[k] \lambda^{m+k}. \end{aligned}$$

Nun schreiben wir $j := m + k$ und bemerken, dass j die Werte 0 bis $\deg f + \deg g = \deg(f \cdot g)$ annimmt. Durch Umgruppieren der Summanden erhalten wir

$$f(\lambda) \cdot_K g(\lambda) = \sum_{j=0}^{\deg(f \cdot g)} \sum_k f[j-k] g[k] \lambda^j$$

wobei die innere Summation über k von 0 bis $\min\{j, \deg g\}$ läuft. Allerdings können wir diese Summation auch problemlos bis j laufen lassen, da $g[k]$ für alle $k > \deg g$ ohnehin gleich Null ist. Also haben wir

$$f(\lambda) \cdot_K g(\lambda) = \sum_{j=0}^{\deg(f \cdot g)} \sum_{k=0}^j f[j-k] g[k] \lambda^j = \sum_{j=0}^{\deg(f \cdot g)} (f \cdot g)[j] \lambda^j = (f \cdot g)(\lambda).$$

Das war zu zeigen. □

Beispiel A.13. **Lemma A.12** vereinfacht manche Überlegungen im Bezug auf Polynomauswertung. Wollen wir beispielsweise das Polynom $f = (X - 5)^2 \in \mathbb{R}[X]$ bei 5 auswerten, so müssen wir nicht zunächst f ausmultiplizieren, $f = X^2 - 10X + 25$, und anschließend

$$f(5) = 5^2 - 10 \cdot 5 + 25 = 25 - 50 + 25 = 0$$

berechnen, sondern können direkt durch

$$f(5) = (5 - 5)^2 = 0^2 = 0$$

zum gewünschten Ergebnis gelangen.

Sei nun $f \in K[X]$ ein Polynom. Wie oben besprochen, können wir f an jedem $\lambda \in K$ auswerten (siehe (A.12)). Dadurch erhalten wir eine Abbildung

$$(A.13) \quad \tilde{f}: \begin{cases} K \longrightarrow K, \\ \lambda \longmapsto f(\lambda) = \text{Auswertung von } f \text{ bei } \lambda \text{ im Sinne von (A.12)}. \end{cases}$$

Wir bezeichnen diese Abbildung mit \tilde{f} (siehe auch [4, § 6.2.2]).

Die oben gewählte Bezeichnung \tilde{f} führt bei Auswertung bei $\lambda \in K$ auf die merkwürdig anmutende Gleichung $\tilde{f}(\lambda) = f(\lambda)$. Weshalb schreiben wir \tilde{f} und nicht einfach nur f ? Dann würden wir aber wieder genau in die Falle tappen, die wir in **Bemerkung A.6** schon identifiziert haben:

Bemerkung A.14 (**Bemerkung A.6** in neuem Gewand). Sei $K = \mathbb{F}_2$ der Körper mit zwei Elementen. Die Polynome $f = X^2 \cdot (X - 1) \in K[X]$ und $g = X \cdot (X - 1) \in K[X]$ erfüllen jeweils $f(0) = 0 = g(0)$ und $f(1) = 0 = g(1)$. Somit gilt für die aus f und g gemäß (A.13) gewonnenen Abbildungen $\tilde{f}, \tilde{g}: K \rightarrow K$ tatsächlich $\tilde{f} = \tilde{g}$. Allerdings gilt $f \neq g$.

Wir betrachten nun die Abbildung

$$(A.14) \quad \Phi: \begin{cases} K[X] \longrightarrow \text{Abb}(K, K) = \{\text{Abbildungen von } K \text{ nach } K\}, \\ f \longmapsto \tilde{f} \text{ aus (A.13)}. \end{cases}$$

Wir haben in **Bemerkung A.14** gesehen, dass diese Abbildung nicht injektiv zu sein braucht. Das Bild von Φ bezeichnen wir mit $P(K)$ und nennen seine Elemente $\tilde{f}: K \rightarrow K$ **polynomiale Abbildungen** über K . Polynomiale Abbildungen (über \mathbb{R}) sind nun das, was man in der Schule als „Polynome“ kennt. Aus **Lemma A.10 (2)** folgt nämlich, dass eine Abbildung $\varphi: K \rightarrow K$ genau dann eine polynomiale Abbildung ist, wenn es $n \in \mathbb{N}_0$ und $a_0, a_1, \dots, a_n \in K$ gibt derart, dass die Abbildung φ sich als

$$\lambda \mapsto a_n \lambda^n + \dots + a_1 \lambda + a_0$$

schreiben lässt.

Beispiele.

- (1) Die Abbildung $\mathbb{R} \rightarrow \mathbb{R}, \lambda \mapsto \lambda^8 - 12\lambda^3 + 5\lambda - 10$, ist eine polynomiale Abbildung.
- (2) Die Abbildung $\mathbb{R} \rightarrow \mathbb{R}, \lambda \mapsto \sin(\lambda)$, ist *keine* polynomiale Abbildung. Um das einzusehen fehlt uns allerdings noch etwas Theorie. Es wird sich nämlich noch später zeigen, dass eine polynomiale Abbildung nur endlich oft den Wert 0 annehmen kann (siehe [4, Korollar 6.2.12]). Da die Sinusfunktion unendlich viele Nullstellen besitzt ($0, \pm\pi, \pm2\pi, \pm3\pi, \dots$), kann diese also nicht polynomial sein.

Die Menge $\text{Abb}(K, K)$ der Abbildungen $K \rightarrow K$ (oder allgemeiner: Abbildungen $M \rightarrow K$ für eine beliebige Menge $M \neq \emptyset$) wird zusammen mit punktweise definierter Addition und Multiplikation zu einem Ring. Unsere Abbildung Φ aus (A.14) erweist sich als verträglich mit der Addition/Multiplikation auf $K[X]$ und dem Ring $\text{Abb}(K, K)$ im Sinne der Gültigkeit von

$$\Phi(f + g) = \Phi(f) +_{\text{Abb}(K, K)} \Phi(g) \quad \text{und} \quad \Phi(f \cdot g) = \Phi(f) \cdot_{\text{Abb}(K, K)} \Phi(g)$$

für alle $f, g \in K[X]$. Zum Beweis dafür muss man lediglich die vorliegenden Definitionen entpacken und sich auf Lemma A.12 berufen. Die Details dazu führen wir hier aber nicht aus.

Wir beenden an dieser Stelle unsere Expedition in den Dschungel der formalen Rechenschlachten und kehren gedanklich zum Skript von Hans Franzen zurück; Etwa auf der Höhe von [4, Satz 6.2.8].

A.3. Jordan-Normalform

Im Folgenden sei $f: V \rightarrow V$ stets ein Endomorphismus eines endlich-dimensionalen K -Vektorraums. (Wir befinden uns gedanklich in [3, §§ 8.4–8.5].)

In Anhang A.1 hatten wir anhand einiger Beispiele gesehen, dass es für einige Untersuchungen günstig sein kann f zu diagonalisieren. Die Problemstellung, einen Endomorphismus zu diagonalisieren, wurde dann in [4, Kapitel 6] zum Hauptgegenstand unseres Interesses erhoben. Durch speziellere Formen von Diagonalisierbarkeit gelang damit auch eine linear-algebraische Beschreibung von euklidischen Isometrien (siehe Aufgabe 9.2 und [4, Satz 7.6.14]) und die exakte Beschreibung von Lösungsmengen quadratischer Gleichungen in mehreren Variablen (siehe [4, Beispiel 7.7.32]).

Wir wissen, dass Diagonalisierbarkeit von f äquivalent zur Existenz einer Basis aus Eigenvektoren ist ([4, Lemma 6.4.3]) und die Summe der Eigenräume von f eine direkte Summe ist ([4, Lemma 6.4.8]). Diagonalisierbarkeit von f ist also äquivalent zu einer Zerlegung

$$(A.15) \quad V = \bigoplus_{\lambda} \text{Eig}(f, \lambda),$$

worin λ die Eigenwerte von f durchläuft. Leider kann es jedoch passieren, dass die Summe auf der rechten Seite „zu klein“ ausfällt und man f somit nicht diagonalisieren kann.

Beispiel A.15. Wir betrachten $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ mit

$$f(v) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} v.$$

Man sieht, dass f nur den Eigenwert 0 hat, aber der zugehörige Eigenraum $\text{Eig}(f, 0) = \text{span}\{(1, 0)\}$ ist lediglich eindimensional. Also ist f nicht diagonalisierbar.

A.3.1. Hauptraumzerlegung. Wir wollen uns nun einen Ersatz für Diagonalisierbarkeit für nicht diagonalisierbare Endomorphismen beschaffen. Geringfügig konkreter gesprochen lautet unser Ziel wie folgt:

Für einen Endomorphismus $f: V \rightarrow V$ eines endlich-dimensionalen K -Vektorraums V , finde eine Basis \mathcal{B} von V derart, dass die Darstellungsmatrix $M_{\mathcal{B}, \mathcal{B}}(f)$ eine „besonders einfache, möglichst diagonale Form“ annimmt.

Tatsächlich zielen wir auf eine obere Dreiecksgestalt mit möglichst wenig besetzten Einträgen abseits der Diagonalen ab. Das charakteristische Polynom einer Matrix in oberer Dreiecksform ist als Produkt $\prod_{\lambda} (X - \lambda)$ gegeben, wobei λ die Diagonaleinträge durchläuft. Wenn für unser f das obige Ziel mit einer oberen Dreiecksmatrix $M_{\mathcal{B}, \mathcal{B}}(f)$ realisiert werden kann, so muss das charakteristische Polynom $\chi_f = \chi_{M_{\mathcal{B}, \mathcal{B}}(f)}$ also notwendiger Weise in Linearfaktoren zerfallen.

Diese Annahme treffen wir im Folgenden auch: χ_f möge in Linearfaktoren zerfallen. Unter dieser zusätzlichen Voraussetzung besteht die für unsere nachfolgenden Überlegungen fundamentale **Hauptraumzerlegung** (siehe auch [3, Satz 8.4.6]):

Satz A.16 (Hauptraumzerlegung). *Sei $f: V \rightarrow V$ ein Endomorphismus endlich-dimensionalen K -Vektorraums mit zerfallendem charakteristischem Polynom. Dann hat man die folgende direkte Summenzerlegung von V :*

$$V = \bigoplus_{\lambda} \text{Hau}(f, \lambda).$$

Hier ist³

$$(A.16) \quad \text{Hau}(f, \lambda) = \bigcup_{i \in \mathbb{N}} \ker((\lambda \text{id}_V - f)^i)$$

der sogenannte **Hauptraum** zum Eigenwert λ von f .

Die Hauptraumzerlegung wird uns als erster Ersatz für (A.15) dienen. Der Beweis von Satz A.16 erfordert etwas Arbeit und wir verschieben diesen auf einen späteren Zeitpunkt (siehe Anhang A.4).

Bemerkung. Bekanntlich ist die Vereinigung $U \cup W$ von Unterräumen eines Vektorraums nur sehr selten wieder ein Unterraum, nämlich genau dann einer der beiden Unterräume im jeweils anderen enthalten ist ($U \subseteq W$ oder $W \subseteq U$). Im Falle der Haupträume ist das aber klar: Für $v \in V$ mit $(\lambda \text{id}_V - f)^i v = 0$ gilt auch

$$(\lambda \text{id}_V - f)^{i+1} v = (\lambda \text{id}_V - f) \circ (\lambda \text{id}_V - f)^i v = 0$$

und also

$$\ker((\lambda \text{id}_V - f)^i) \subseteq \ker((\lambda \text{id}_V - f)^{i+1}).$$

Damit sieht man leicht, dass es sich bei $\text{Hau}(f, \lambda)$ tatsächlich um einen Unterraum von V handelt.

³Potenzen $(\lambda \text{id}_V - f)^i$ des Endomorphismus $\lambda \text{id}_V - f$ sind als Verkettung $(\lambda \text{id}_V - f) \circ \dots \circ (\lambda \text{id}_V - f)$ [i Terme] zu verstehen.

Beispiel A.17. Für $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ mit

$$f(v) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} v$$

(aus **Beispiel A.15**) gilt $f^2 = f \circ f = \text{Nullabbildung}$. Daher ist $\text{Hau}(f, 0) = \mathbb{R}^2$.

Lemma A.18. Der Hauptraum $\text{Hau}(f, \lambda)$ von f zum Eigenwert λ ist f -invariant, d.h. es gilt

$$f(\text{Hau}(f, \lambda)) \subseteq \text{Hau}(f, \lambda).$$

Beweis. Für beliebiges $v \in V$ gilt

$$\begin{aligned} ((\lambda \text{id}_V - f) \circ f)(v) &= (\lambda \text{id}_V - f)(f(v)) = (\lambda f(v) - f(f(v))) \\ &= f(\lambda v - f(v)) = f((\lambda \text{id}_V - f)(v)) = (f \circ (\lambda \text{id}_V - f))(v). \end{aligned}$$

Also ist $(\lambda \text{id}_V - f) \circ f = f \circ (\lambda \text{id}_V - f)$ und induktiv folgt

$$(\lambda \text{id}_V - f)^i \circ f = f \circ (\lambda \text{id}_V - f)^i$$

für alle $i \in \mathbb{N}$. Sei nun speziell $v \in \text{Hau}(f, \lambda)$. Dann gibt es ein $i \in \mathbb{N}$ mit $v \in \ker((\lambda \text{id}_V - f)^i)$. Die obige Gleichung zeigt jedoch, dass $f(v)$ auch in $\ker((\lambda \text{id}_V - f)^i)$ liegt. Also ist $f(\text{Hau}(f, \lambda)) \subseteq \text{Hau}(f, \lambda)$. \square

Wählt man nun eine Basis \mathcal{B}_λ von jeden Hauptraum $\text{Hau}(f, \lambda)$, so ist

$$\mathcal{B} = (\mathcal{B}_\lambda : \lambda \text{ Eigenwert von } f)$$

dank **Satz A.16** eine Basis von V . Die Darstellungsmatrix $M_{\mathcal{B}, \mathcal{B}}(f)$ von f bezüglich \mathcal{B} hat wegen **Lemma A.18** Blockdiagonalgestalt,

$$M_{\mathcal{B}, \mathcal{B}}(f) = \begin{pmatrix} \boxed{*} & & \dots & 0 \\ & \boxed{*} & & \vdots \\ \vdots & & \ddots & \\ 0 & \dots & & \boxed{*} \end{pmatrix},$$

wobei pro Eigenwert λ von f ein quadratischer Block der Größe $\dim \text{Hau}(f, \lambda) = \#\mathcal{B}$ auftritt.

Im Allgemeinen können die jeweiligen Blöcke sehr dicht besetzt sein. Wir wollen dies nun noch durch sehr spezielle Wahl der Basen \mathcal{B}_λ verhindern.

Dank **Lemma A.18** lässt sich f zu einem Endomorphismus von $U_\lambda := \text{Hau}(f, \lambda)$ einschränken. Wir schreiben $f_\lambda := f|_{U_\lambda}: U_\lambda \rightarrow U_\lambda$ und zerlegen f_λ wie folgt:

$$(A.17) \quad f_\lambda = \lambda \text{id}_{U_\lambda} + \underbrace{(f_\lambda - \lambda \text{id}_{U_\lambda})}_{=: g_\lambda}.$$

Lemma A.19. g_λ ist **nilpotent**, d.h. es gibt ein $i_0 \in \mathbb{N}$ mit $g_\lambda^{i_0} = \text{Nullabbildung}$.

Beweis. Da V endlich-dimensional ist, können in der Kette

$$\ker((\lambda \text{id}_V - f)^1) \subseteq \ker((\lambda \text{id}_V - f)^2) \subseteq \ker((\lambda \text{id}_V - f)^3) \subseteq \dots$$

nur endlich viele Inklusionen strikt sein. Es gibt somit einen Exponenten $i_0 \in \mathbb{N}$ mit

$$U_\lambda = \text{Hau}(f, \lambda) = \ker((\lambda \text{id}_V - f)^{i_0}) = \ker((-g_\lambda)^{i_0}) = \ker(g_\lambda^{i_0}).$$

Also verschwindet $g_\lambda^{i_0}$ auf dem gesamten Raum U_λ und ist also die Nullabbildung. \square

Das nächste Ergebnis hat Anwendungen in der Theorie der Differentialgleichungen für die Berechnung des sogenannten Matrix-Exponentials $\exp(A)$ einer komplexen $n \times n$ -Matrix A .

Korollar A.20 (Jordan-Zerlegung). Sei $f: V \rightarrow V$ ein Endomorphismus eines endlich-dimensionalen K -Vektorraums. Das charakteristische Polynom von f zerfällt in Linearfaktoren. Dann gibt es eine Zerlegung $f = f_{\text{diag}} + f_{\text{nil}}$ mit

- (1) f_{diag} ist diagonalisierbar,
- (2) f_{nil} ist nilpotent,
- (3) $f_{\text{diag}} \circ f_{\text{nil}} = f_{\text{nil}} \circ f_{\text{diag}}$.

Beweis. Man definiere f_{diag} als $\lambda \text{id}_{U_\lambda}$ auf jedem Hauptraum $U_\lambda = \text{Hau}(f, \lambda)$ von f und f_{nil} auf U_λ als $f_\lambda - \lambda \text{id}_{U_\lambda}$ (siehe (A.17)). Nun setze man f_{diag} und f_{nil} linear auf $\bigoplus_\lambda U_\lambda = V$ fort. Die im Korollar genannten drei Eigenschaften sind jedenfalls auf jedem Hauptraum erfüllt und man sieht leicht, dass sich diese dann auf die direkte Summe der Haupträume fortsetzen. \square

A.3.2. Normalform für nilpotente Endomorphismen. Der Endomorphismus g_λ aus (A.17) ist nilpotent laut Lemma A.19. Es ist nun unser Ziel, nilpotente Endomorphismen bezüglich einer geeigneten Basis in einer gewissen Normalform darzustellen. Wir studieren diese Problematik zunächst allgemein und ändern dafür unsere Notation geringfügig. Sei $g: U \rightarrow U$ ein nilpotenter Endomorphismus auf einem endlich-dimensionalen K -Vektorraum.

Wir betrachten die aufsteigende Folge

$$\{0\} = \ker(g^0) \subseteq \ker(g^1) \subseteq \ker(g^2) \subseteq \dots$$

und schreiben abkürzend $U_i = \ker(g^i)$.

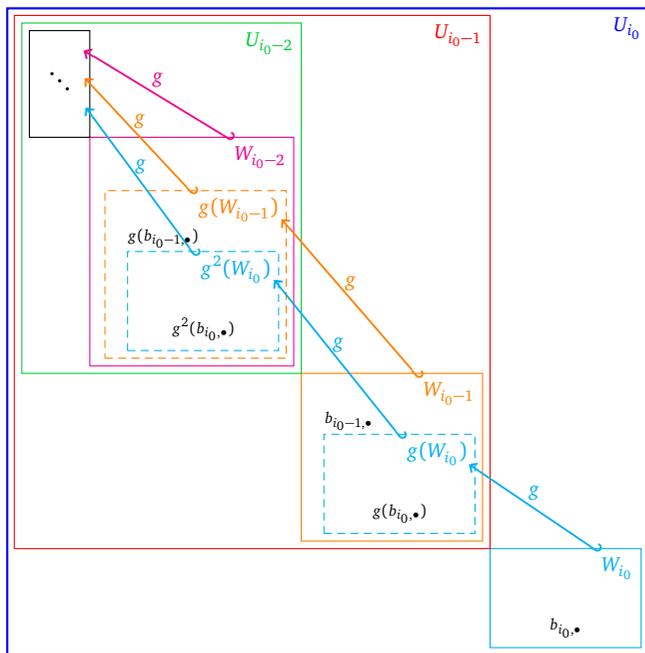
Lemma A.21. Für $i \in \mathbb{N}$ sei W ein beliebiger **Komplementärraum** zu U_i in U_{i+1} , d.h. W sei ein Unterraum mit $U_i \oplus W = U_{i+1}$.⁴ Dann gelten die folgenden Aussagen:

- (1) $g|_W: W \rightarrow U_i$ ist injektiv und
- (2) das Bild $\text{im}(g|_W) = g(W)$ hat trivialen Schnitt mit U_{i-1} .

⁴Man beachte, dass ein Unterraum im Allgemeinen viele Komplementärräume besitzen kann. Beispielsweise sind $\mathbb{R}(0, 1)$ und $\mathbb{R}(1, 1)$ beide Komplementärräume von $\mathbb{R}(1, 0)$ im \mathbb{R}^2 .

Beweis. Sei $i \in \mathbb{N}$ und $w \in \ker(g|_W)$. Angenommen $w \neq 0$. Wegen $U_i \cap W = \{0\}$ folgt dann $w \notin U_i = \ker(g^i)$, also $g^i(w) \neq 0$. Wegen $w \in \ker(g|_W)$ ist jedoch $g(w) = 0$. Das ist ein Widerspruch und beweist die Injektivität von $g|_W$.

Nun betrachten wir den Schnitt $g(W) \cap U_{i-1}$. Sei $g(w)$ ein Element davon. Dieses wird dann von g^{i-1} auf 0 abgebildet. Also ist $g^i(w) = g^{i-1}(g(w)) = 0$ und somit $w \in U_i$. Wegen $w \in U_i \cap W = \{0\}$ folgt $w = 0$ und also $g(W) \cap U_{i-1} = \{0\}$, wie behauptet. \square



Eingabe: $g \in \text{End}(V)$ mit $g^{i_0} = 0$, $i_0 \geq 1$

Ausgabe: Basis \mathcal{B} von $U = U_{i_0}$ aus (A.19).

Schritt 1: Setze $i := i_0$, $\mathcal{B} := \emptyset$, $\mathcal{B}' := \emptyset$.

➤ **Schritt 2:** Füge alle Elemente von \mathcal{B}' zu \mathcal{B} hinzu.

Schritt 3: Ergänze \mathcal{B}' zu einer Basis eines Komplements W_i von U_{i-1} in U_i .

Schritt 4: Füge die soeben ergänzten Elemente auch zu \mathcal{B} hinzu.

Schritt 5: Bezeichne die ergänzten Vektoren mit $b_{i,\bullet} = (b_{i,1}, \dots, b_{i,r_i})$ und füge diese zu \mathcal{B} hinzu. Setze $\mathcal{B}' := g(\mathcal{B}')$.

➤ **Schritt 6:** Falls $i > 1$ gilt, verringere i um 1 und gehe zu Schritt 2.

Schritt 7: Gib \mathcal{B} zurück.

Abbildung 1. Algorithmus zur Bestimmung der Basis \mathcal{B} aus (A.19) mit (A.21) (bei geeigneter Sortierung der Basiselemente). Eingabe ist ein nilpotenter Endomorphismus $g: V \rightarrow V$ auf einem endlich-dimensionalen K -Vektorraum V . Die Rechtecke in der Abbildung stellen jeweils Unterräume dar. Die Abbildung ist allerdings mit Vorsicht zu interpretieren: Beispielsweise enthalten alle gezeichneten Räume den Nullvektor, auch wenn manche von diesen in der Abbildung disjunkt zu sein scheinen.

Sei g wie oben. Wegen der Nilpotenz von g gibt es ein $i_0 \in \mathbb{N}$ mit

$$g^{i_0} = \text{Nullabbildung.}$$

Daher gilt $U_{i_0} = \ker(g^{i_0}) = U$ (der gesamte Raum, auf dem g definiert ist). Wir wählen nun einen Komplementärraum W_{i_0} von U_{i_0-1} in U_{i_0} , also

$$U_{i_0} = U_{i_0-1} \oplus W_{i_0}.$$

Wir wählen nun eine Basis⁵ $b_{i_0, \bullet} = (b_{i_0,1}, \dots, b_{i_0, r_{i_0}})$ von W_{i_0} .

Für die folgenden Überlegungen nehmen wir $i_0 > 1$ an. (Im Fall $i_0 = 1$ ignoriere man diese.) Dank **Lemma A.21** handelt es sich bei $g(b_{i_0,1}), \dots, g(b_{i_0, r_{i_0}})$ um linear unabhängige Elemente in $U_{i_0-1} \setminus U_{i_0-2}$. Wir können diese nun zu einer Basis eines Komplementärraumes W_{i_0-1} von U_{i_0-2} in U_{i_0-1} ergänzen:

$$(g(b_{i_0,1}), \dots, g(b_{i_0, r_{i_0}}), b_{i_0-1,1}, \dots, b_{i_0-1, r_{i_0-1}}) = (g(b_{i_0, \bullet}), b_{i_0-1, \bullet}).$$

(Falls die $g(b_{i_0, \bullet})$ bereits einen Komplementärraum von U_{i_0-2} in U_{i_0-1} aufspannen, so setzen wir $r_{i_0-1} = 0$ und verstehen $b_{i_0-1, \bullet}$ als eine leere Liste.)

Ist nun zusätzlich $i_0 > 2$, so liefert eine abermalige Berufung auf **Lemma A.21**, dass

$$(g^2(b_{i_0, \bullet}), g(b_{i_0-1, \bullet}))$$

eine linear unabhängige Familie in $U_{i_0-2} \setminus U_{i_0-3}$ bildet. Wir können diese nun zu einer Basis eines Komplementärraumes W_{i_0-2} von U_{i_0-3} in U_{i_0-2} ergänzen:

$$(g^2(b_{i_0, \bullet}), g(b_{i_0-1, \bullet}), b_{i_0-2, \bullet}).$$

Die Anzahl der ergänzten $b_{i_0-2, \bullet}$ bezeichnen wir mit r_{i_0-2} und lassen explizit den Fall $r_{i_0-2} = 0$ zu, falls keine Vektoren zu ergänzen waren.

Wir führen dieses Verfahren sukzessive weiter und erhalten die folgende Zerlegung von U :

$$\begin{aligned}
 U &= \underbrace{W_0}_{=U_0=\{0\}} \oplus W_1 \oplus \dots \oplus W_{i_0-2} \oplus W_{i_0-1} \oplus W_{i_0} \cdot \\
 &\quad \underbrace{\hspace{10em}}_{=U_1} \\
 &\quad \underbrace{\hspace{10em}}_{\dots} \\
 &\quad \underbrace{\hspace{10em}}_{=U_{i_0-2}} \\
 &\quad \underbrace{\hspace{10em}}_{=U_{i_0-1}} \\
 &\quad \underbrace{\hspace{10em}}_{=U_{i_0}} \quad (\text{Leserichtung: } \curvearrowleft)
 \end{aligned}$$

⁵Die Notation mit dem fetten Punkt „•“ darf hoffentlich als selbsterklärend angesehen werden. Ohne derartige abkürzenden Kunstgriffe würden die weiteren Ausführungen unüberschaubar werden.

Die gewählten Basiselemente werden durch g gemäß des folgenden Schemas abgebildet:⁶

$$\begin{array}{ccccccccccc}
 \{0\} & \xleftarrow{g} & W_1 & \xleftarrow{g} & \dots & \xleftarrow{g} & W_{i_0-2} & \xleftarrow{g} & W_{i_0-1} & \xleftarrow{g} & W_{i_0}, \\
 0 & \xleftarrow{\quad} & g^{i_0-1}(b_{i_0,\bullet}) & \xleftarrow{\quad} & \dots & \xleftarrow{\quad} & g^2(b_{i_0,\bullet}) & \xleftarrow{\quad} & g(b_{i_0,\bullet}) & \xleftarrow{\quad} & b_{i_0,\bullet}, \\
 \text{(A.18)} & & 0 & \xleftarrow{\quad} & g^{i_0-2}(b_{i_0-1,\bullet}) & \xleftarrow{\quad} & \dots & \xleftarrow{\quad} & g(b_{i_0-1,\bullet}) & \xleftarrow{\quad} & b_{i_0-1,\bullet}, \\
 & & & & \vdots & & \ddots & & & & \\
 0 & \xleftarrow{\quad} & b_{1,\bullet} & & & & & & & &
 \end{array}$$

(Das „Vorrücken“ der Vektoren sollte hierbei offensichtlich sein. Allenfalls diskutabel ist, ob „am Ende“ der Kette links wirklich auf 0 abgebildet wird. Das gilt allerdings konstruktionsbedingt wegen $b_{i,\bullet} \in U_i = \ker(g^i)$.) Bei

$$\text{(A.19)} \quad \mathcal{B} = \left(\begin{array}{l} 1 \leq i \leq i_0, \\ g^k(b_{i_0,j}) : 1 \leq j \leq r_i, \\ 0 \leq k < i \end{array} \right)$$

handelt es sich also um eine Basis von U . Wir denken uns diese Basis derartig sortiert, dass die Blöcke

$$\mathcal{B}_{i,j}: \quad g^{i-1}(b_{i,j}), g^{i-2}(b_{i,j}), \dots, g^2(b_{i,j}), g(b_{i,j}), b_{i,j}$$

zusammen gruppiert sind und in der Reihenfolge

$$\text{(A.20)} \quad \mathcal{B}_{1,1}, \dots, \mathcal{B}_{1,r_1}, \dots, \mathcal{B}_{i_0,1}, \dots, \mathcal{B}_{i_0,r_{i_0}}$$

durchlaufen werden. In (A.19) entspricht dies, erst i zu wählen, dann j und dann k , aber k von $i-1$ aus absteigend zu durchlaufen. Dann hat $M_{\mathcal{B},\mathcal{B}}(g)$ (wegen (A.18))

⁶Der Pfeil „ $\xleftrightarrow{\quad}$ “ in $W_{i+1} \xleftrightarrow{\quad} W_i$ deutet auf eine „Inklusion“ hin und soll uns an Lemma A.21 (1) erinnern.

folgende Form:

$$(A.21) \quad M_{\mathcal{B}, \mathcal{B}}(g) = \left(\begin{array}{cccc} \boxed{J_1(0)} & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & \boxed{J_1(0)} & & \\ \dots & & & & \\ \vdots & & & & \\ 0 & \dots & & & \boxed{J_{i_0}(0)} \end{array} \right),$$

$\left. \begin{array}{c} \dots \\ \dots \\ \dots \end{array} \right\} r_1 \text{ Blöcke}$

 $\left. \begin{array}{c} \dots \\ \dots \\ \dots \end{array} \right\} r_{i_0} \text{ Blöcke}$

wobei die eingekreisten Blöcke zu den $\mathcal{B}_{i,j}$ korrespondieren und folgende Gestalt haben:

$$(A.22) \quad J_1(0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad J_2(0) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad J_3(0) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad \dots,$$

$$J_i(0) = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix} \in K^{i \times i}.$$

(Lauter Einsen auf der ersten oberen Nebendiagonalen und sonst überall Null-Einträge.)

Wir wollen uns kurz überlegen, wie sich die Anzahl r_i der Blöcke $J_i(0)$ in (A.21) aus den Dimensionen $\dim U_1, \dots, \dim U_{i_0}$ ableitet. Wir schreiben zur Abkürzung d_i für $\dim U_i$. Zunächst gilt

$$r_{i_0} = \dim W_{i_0} = d_{i_0} - d_{i_0-1}.$$

Im Fall $i_0 \geq 2$ gilt weiter

$$\begin{aligned} r_{i_0-1} &= \dim W_{i_0-1} - r_{i_0} = d_{i_0-1} - d_{i_0-2} - (d_{i_0} - d_{i_0-1}) \\ &= -d_{i_0-2} + 2d_{i_0-1} - d_{i_0}, \end{aligned}$$

und im Fall $i_0 \geq 3$

$$\begin{aligned} r_{i_0-2} &= \dim W_{i_0-2} - r_{i_0} - r_{i_0-1} = d_{i_0-2} - d_{i_0-3} - (d_{i_0} - d_{i_0-1}) - (-d_{i_0-2} + 2d_{i_0-1} - d_{i_0}) \\ &= -d_{i_0-3} + 2d_{i_0-2} - d_{i_0-1}. \end{aligned}$$

Ein Induktionsbeweis liefert

$$(A.23) \quad \begin{pmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & -1 & 2 & -1 & \\ & & & -1 & 2 & -1 \\ & & & & -1 & \mathbf{1} \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_{i_0-1} \\ d_{i_0} \end{pmatrix} = \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_{i_0-1} \\ r_{i_0} \end{pmatrix}.$$

Wir verzichten darauf, die Details dazu weiter auszuführen.

Für spätere Verwendung bemerken wir noch, dass die obige Matrix invertierbar ist. In der Tat: Sie hat Determinante 1. Das folgt leicht durch Induktion über die Dimension, indem man die letzte Spalte zu der vorherigen addiert, das mit der neuen vorletzten Spalte wiederholt und so fortfährt:

$$\begin{pmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & -1 & 2 & -1 & \\ & & & -1 & 2 & -1 \\ & & & & -1 & \mathbf{1} \end{pmatrix} \rightsquigarrow \begin{pmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & -1 & 2 & -1 & \\ & & & -1 & \mathbf{1} & -1 \\ & & & & 0 & \mathbf{1} \end{pmatrix} \rightsquigarrow \dots;$$

Es entsteht eine obere Dreiecksmatrix, deren Diagonale nur mit Einsen besetzt ist. Diese hat also Determinante 1, aber da das Addieren einer Spalte zu einer anderen die Determinante unverändert lässt, gilt dies auch für die Matrix in (A.23).

Alternativ hätten wir uns auch überlegen können, dass die Inverse der Matrix in (A.23) wie folgt gegeben ist:

$$\begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \boxed{2} & 2 & \dots & 2 \\ 1 & 2 & \boxed{3} & \dots & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \dots & \boxed{i_0} \end{pmatrix}.$$

Beispiele. In niedrigen Dimensionen $n = \dim U$ lässt sich leicht auflisten, welche Formen für $A = M_{\mathcal{B}, \mathcal{B}}(g)$ möglich sind.

(1) $n = 1$:

$$\left(\boxed{0} \right).$$

(2) $n = 2$:

$$\left(\begin{array}{c|c} \boxed{0} & 0 \\ \hline 0 & \boxed{0} \end{array} \right), \quad \left(\begin{array}{c|c} 0 & 1 \\ \hline 0 & 0 \end{array} \right).$$

(3) $n = 3$:

$$\left(\begin{array}{|c|c|c|} \hline 0 & 0 & 0 \\ \hline 0 & 0 & 0 \\ \hline 0 & 0 & 0 \\ \hline \end{array} \right), \quad \left(\begin{array}{|c|c|c|} \hline 0 & 0 & 0 \\ \hline 0 & 0 & 1 \\ \hline 0 & 0 & 0 \\ \hline \end{array} \right), \quad \left(\begin{array}{|c|c|c|} \hline 0 & 1 & 0 \\ \hline 0 & 0 & 0 \\ \hline 0 & 0 & 0 \\ \hline \end{array} \right), \quad \left(\begin{array}{|c|c|c|} \hline 0 & 1 & 0 \\ \hline 0 & 0 & 1 \\ \hline 0 & 0 & 0 \\ \hline \end{array} \right).$$

Man beachte, dass die dritte Matrix aus der zweiten durch Permutation der Blöcke gewonnen werden kann. (Streng genommen kommt die dritte Matrix auch gar nicht als $M_{\mathcal{B},\mathcal{B}}(g)$ in (A.21) vor, da wir uns ja auf die Anordnung der Basis wie in (A.20) festgelegt haben. Demnach sind die Größen i der Jordan-Blöcke $J_i(0)$ *aufsteigend* sortiert.) Tatsächlich, wenn

$$M_{\mathcal{B},\mathcal{B}}(g) = \text{zweite Matrix von oben}$$

gilt, so lässt sich durch Umordnung der Blöcke in \mathcal{B} eine neue Basis \mathcal{B}' mit

$$M_{\mathcal{B}',\mathcal{B}'}(g) = \text{dritte Matrix von oben}$$

gewinnen. Man beachte, dass diese vier Matrizen zu den vier additiven Zerlegungen

$$1 + 1 + 1 = 1 + 2 = 2 + 1 = 3$$

der Zahl $n = 3$ korrespondieren und man die Zerlegungen $1 + 2$ und $2 + 1$ gewissermaßen identifizieren kann. In (A.21) haben wir die Blöcke in unserer Basis so sortiert, dass die dritte Matrix von oben nicht als Matrix (A.21) vorkommen würde.

(4) $n = 4$: Hier hat man 8 mögliche Matrizen korrespondierend zu den folgenden Zerlegungen der Zahl $n = 4$:

$$4 = 1 + 1 + 1 + 1 = \underbrace{1 + 1 + 2}_{3 \text{ Permutationen}} = 2 + 2 = \underbrace{1 + 3}_{2 \text{ Permutationen}}.$$

Wir verzichten darauf, diese explizit hinzuschreiben.

Beispiel A.22. Wir setzen

$$A := \begin{pmatrix} 0 & 2 & 1 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{und berechnen} \quad A^2 = \begin{pmatrix} 0 & 0 & 6 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Wir betrachten den Endomorphismus

$$g: \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad v \mapsto Av.$$

Dann ist g^3 die Nullabbildung. Wir sehen sofort

$$U_1 = \ker(g) = \text{span}\{e_1\}, \quad U_2 = \ker(g^2) = \text{span}\{e_1, e_2\},$$

$$U_3 = \ker(g^3) = \mathbb{R}^3 = \text{span}\{e_1, e_2, e_3\},$$

mit den Standardbasisvektoren e_1, e_2, e_3 des \mathbb{R}^3 . Wir sehen, dass beim Schritt von U_2 zu U_3 der Vektor e_3 hinzugekommen ist. Wir setzen $b_{3,1} = e_3$ berechnen nun

$$g(b_{3,1}) = g(e_3) = (1, 3, 0).$$

Es gilt

$$U_2 \oplus \text{span}\{b_{3,1}\} = U_3 \quad \text{und} \quad U_1 \oplus \text{span}\{g(b_{3,1})\} = U_2.$$

Dementsprechend gibt es hier keine Vektoren $b_{2,\bullet}$ zu ergänzen. Wir betrachten nun weiter

$$g^2(b_{3,1}) = g(1, 3, 0) = (6, 0, 0).$$

Es gilt

$$U_0 \oplus \text{span}\{g^2(b_{3,1})\} = U_1$$

und auch hier müssen wir keine Vektoren $b_{1,\bullet}$ ergänzen. Wir erhalten die Basis

$$\mathcal{B} = (g^2(b_{3,1}), g(b_{3,1}), b_{3,1}) = \left(\begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right).$$

Nun gilt

$$M_{\mathcal{B}, \mathcal{B}}(g) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

wie wir schon oben abstrakt gerechtfertigt haben. Auf Matrizen-Ebene haben wir

$$\begin{pmatrix} 6 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} A \begin{pmatrix} 6 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Beispiel A.23. Wir betrachten den Endomorphismus

$$g: \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad v \mapsto \begin{pmatrix} 0 & 4 & 5 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} v.$$

Dann ist g^2 die Nullabbildung. Wir sehen sofort

$$U_1 = \ker(g) = \text{span} \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -5 \\ 4 \end{pmatrix} \right\}, \quad U_2 = \ker(g^2) = \mathbb{R}^3.$$

Wir sehen, dass beim Übergang von U_1 zu U_2 via g der dritte Standardeinheitsvektor $e_3 = (0, 0, 1)$ hinzugekommen ist. Wir setzen $b_{2,1} = e_2$ und berechnen nun

$$g(b_{2,1}) = g(e_3) = (5, 0, 0),$$

Es gilt $U_1 \oplus \text{span}\{b_{2,1}\} = U_2$, aber $U_0 \oplus \text{span}\{g(b_{2,1})\} \subsetneq U_1$. Um Gleichheit herzustellen, ergänzen wir noch um den Vektor $b_{1,1} = (-5, 4, 0) \in U_1 \setminus U_0$. Wir betrachten die Basen

$$\mathcal{B} = (g(b_{2,1}), b_{2,1}, b_{1,1}) = \left(\begin{pmatrix} 5 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -5 \\ 4 \end{pmatrix} \right), \quad \mathcal{B}' = (b_{1,1}, g(b_{2,1}), b_{2,1}).$$

Wir haben

$$(A.24) \quad M_{\mathcal{B},\mathcal{B}}(g) = \begin{pmatrix} \boxed{0} & \boxed{1} & 0 \\ \boxed{0} & \boxed{0} & 0 \\ 0 & 0 & \boxed{0} \end{pmatrix} \quad \text{und} \quad M_{\mathcal{B}',\mathcal{B}'}(g) = \begin{pmatrix} \boxed{0} & 0 & 0 \\ 0 & \boxed{0} & \boxed{1} \\ 0 & \boxed{0} & \boxed{0} \end{pmatrix}.$$

Beispiele. Sei $g: \mathbb{R}^6 \rightarrow \mathbb{R}^6$ ein nilpotenter Endomorphismus. Wir geben weiter unten Werte für die Folge $d_\bullet = (d_i)_{i=1,2,3,\dots}$ mit $d_i = \dim \ker U_i = \dim \ker(g^i)$ vor und wollen die Matrix $M_{\mathcal{B},\mathcal{B}}(g)$ aus (A.21), also die Anzahlen r_i der Blöcke $J_i(0)$ für $i = 1, 2, 3, \dots$, bestimmen.

- (1) Es möge $d_\bullet = (3, 5, 6, 6, 6, \dots)$ gelten. Wir können $i_0 = 3$ wählen und erhalten dann gemäß (A.23)

$$\begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} = \begin{pmatrix} 2 & -1 & \\ -1 & 2 & -1 \\ & -1 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Die Matrix $M_{\mathcal{B},\mathcal{B}}(g)$ aus (A.21) besteht somit aus je einem Block $J_1(0)$, $J_2(0)$ und $J_3(0)$, also:

$$M_{\mathcal{B},\mathcal{B}}(g) = \begin{pmatrix} \boxed{1} & & & & & \\ & \boxed{0} & \boxed{1} & & & \\ & \boxed{0} & \boxed{0} & & & \\ & & & \boxed{0} & \boxed{1} & \boxed{0} \\ & & & \boxed{0} & \boxed{0} & \boxed{1} \\ & & & \boxed{0} & \boxed{0} & \boxed{0} \end{pmatrix}.$$

- (2) Wir gehen wie in (1) wieder von $d_\bullet = (3, 5, 6, 6, 6, \dots)$ aus, wählen nun aber $i_0 = 4$. Dann liefert (A.23)

$$\begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{pmatrix} = \begin{pmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & -1 & 2 & -1 \\ & & -1 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 5 \\ 6 \\ 6 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}.$$

Das führt auf dasselbe Ergebnis wie in (1). (Das sollte uns natürlich auch nicht wundern.)

- (3) Es möge $d_\bullet = (3, 5, 5, 6, 6, \dots)$ gelten. Dann können wir $i_0 = 4$ wählen und erhalten

$$\begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{pmatrix} = \begin{pmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & -1 & 2 & -1 \\ & & -1 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 5 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ -1 \\ 1 \end{pmatrix}.$$

Was ist hier geschehen? Die Anzahl r_3 der $J_3(0)$ Blöcke kann zwar Null sein, aber sicher nicht negativ! In der Tat kann die oben gewählte Folge d_\bullet nicht als Folge von Dimensionen eines nilpotenten g wie oben vorkommen. Das sehen wir einerseits an der eben erhaltenen widersprüchlichen Anzahl

r_3 , aber auch aus dem Faktum, dass $U_3 = \ker(g^3)$ ja ein eindimensionales Komplement W_4 in $U_4 = \ker(g^3)$ haben würde und dieses Komplement laut **Lemma A.21 (2)** von g nach U_3 abgebildet werden würde, sich aber nur trivial mit U_2 schneidet. Damit müsste U_3 also mindestens Dimension $5 + 1 = 6 > d_3$ haben, was nicht funktioniert.

A.3.3. Jordan-Normalform: Synthese. Wir kehren nun gedanklich zu unseren Darlegungen aus **Anhang A.3.1**, und zwar speziell zu der Zerlegung **(A.17)** zurück. Wir hatten bereits gesehen, dass $M_{\mathcal{B},\mathcal{B}}(f)$ Blockdiagonalgestalt hat, wenn man die Basis \mathcal{B} aus Basen der einzelnen Haupträume $U_\lambda = \text{Hau}(f, \lambda)$ von f zusammensetzt. Unsere Überlegungen aus **Anhang A.3.2** angewendet auf $g = g_\lambda$ aus **(A.17)** liefern eine Basiswahl \mathcal{B}_λ auf U_λ derart, dass $M_{\mathcal{B}_\lambda, \mathcal{B}_\lambda}(g_\lambda)$ die Gestalt **(A.21)** besitzt. Nun kommt der springende Punkt: der Endomorphismus $\lambda \text{id}_{U_\lambda}$ hat bezüglich der Basis \mathcal{B}_λ (ja, sogar bezüglich jeder Basis) die Darstellungsmatrix $\lambda E_{\dim U_\lambda}$.

Wir wählen nun in jedem Haupträume $U_\lambda = \text{Hau}(f, \lambda)$ von f eine Basis U_λ , wie oben beschrieben, und fügen diese blockweise zu einer Basis \mathcal{B} von V zusammen. Ein erneuter Blick auf **(A.17)** liefert nun

$$(A.25) \quad M_{\mathcal{B},\mathcal{B}}(f) = \begin{pmatrix} \boxed{J_\bullet(\circ)} & & \dots & & 0 \\ & \boxed{J_\bullet(\circ)} & & & \vdots \\ & \vdots & & \ddots & \\ 0 & & \dots & & \boxed{J_\bullet(\circ)} \end{pmatrix},$$

wobei die eingezeichneten Blöcke folgende Gestalt haben (vgl. **(A.22)**):

$$J_i(\lambda) = \lambda E_i + J_i(0) = \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix} \in K^{i \times i}.$$

Man bezeichnet $J_i(\lambda)$ auch als **Jordan-Matrix** (oder **Jordan-Block**) der Größe i zum Eigenwert λ .

Von Blockdiagonalmatrizen wie **(A.25)**, deren Blöcke nur Jordan-Blöcke sind, sagt man, diese seien in **Jordan-Normalform**. Unsere bisherigen Ausführungen haben uns nun also den folgenden Satz geliefert, wenn man den Schönheitsfehler ignoriert, dass der Beweis vom Satz über die Hauptraumzerlegung, **Satz A.16**, noch aussteht:

Satz A.24 (Jordan-Normalform, Existenz). *Sei $f: V \rightarrow V$ ein Endomorphismus eines endlich-dimensionalen K -Vektorraums. Das charakteristische Polynom von f zerfalle*

in Linearfaktoren. Dann existiert eine Basis \mathcal{B} bezüglich derer die Darstellungsmatrix $M_{\mathcal{B},\mathcal{B}}(f)$ von f in Jordan-Normalform vorliegt.

Bemerkung A.25 (Jordan-Normalform für Matrizen). Man kann [Satz A.24](#) natürlich auch für Matrizen formulieren. Dann besagt der Satz, dass für jede $n \times n$ -Matrix A mit in Linearfaktoren zerfallendem charakteristischem Polynom eine invertierbare $n \times n$ -Matrix T existiert, sodass $T^{-1}AT$ in Jordan-Normalform vorliegt.

Beispiele. Wir besprechen hier in kleinen Dimensionen $n = \dim V$, welche Gestalt für $M_{\mathcal{B},\mathcal{B}}(f)$ in [\(A.25\)](#) in Frage kommt. Mit λ , μ und ρ seien paarweise verschiedene Elemente von K bezeichnet.

(1) $n = 1$:

$$\left(\boxed{\lambda} \right).$$

(2) $n = 2$:

zwei Eigenwerte,	ein Eigenwert.
$\begin{pmatrix} \boxed{\lambda} & 0 \\ 0 & \boxed{\mu} \end{pmatrix},$	$\begin{pmatrix} \boxed{\lambda} & 0 \\ 0 & \boxed{\lambda} \end{pmatrix},$
	$\begin{pmatrix} \boxed{\lambda} & 1 \\ 0 & \boxed{\lambda} \end{pmatrix}.$

(3) $n = 3$:

drei Eigenwerte,	zwei Eigenwerte,	ein Eigenwert.
$\begin{pmatrix} \boxed{\lambda} & 0 & 0 \\ 0 & \boxed{\mu} & 0 \\ 0 & 0 & \boxed{\rho} \end{pmatrix},$	$\begin{pmatrix} \boxed{\lambda} & 0 & 0 \\ 0 & \boxed{\lambda} & 0 \\ 0 & 0 & \boxed{\mu} \end{pmatrix},$	$\begin{pmatrix} \boxed{\lambda} & 0 & 0 \\ 0 & \boxed{\lambda} & 0 \\ 0 & 0 & \boxed{\lambda} \end{pmatrix},$
	$\begin{pmatrix} \boxed{\lambda} & 1 & 0 \\ 0 & \boxed{\lambda} & 0 \\ 0 & 0 & \boxed{\mu} \end{pmatrix},$	$\begin{pmatrix} \boxed{\lambda} & 0 & 0 \\ 0 & \boxed{\lambda} & 1 \\ 0 & 0 & \boxed{\lambda} \end{pmatrix},$
		$\begin{pmatrix} \boxed{\lambda} & 1 & 0 \\ 0 & \boxed{\lambda} & 1 \\ 0 & 0 & \boxed{\lambda} \end{pmatrix}.$

Zusätzlich gibt es noch zwei weitere Möglichkeiten, die durch Vertauschung der beiden Blöcke in den beiden Matrizen in der mittleren Zeile entstehen:

$$\begin{pmatrix} \boxed{\mu} & 0 & 0 \\ 0 & \boxed{\lambda} & 1 \\ 0 & 0 & \boxed{\lambda} \end{pmatrix}, \quad \begin{pmatrix} \boxed{\lambda} & 1 & 0 \\ 0 & \boxed{\lambda} & 0 \\ 0 & 0 & \boxed{\lambda} \end{pmatrix}.$$

In der Situation von [Satz A.24](#) sagen wir $M_{\mathcal{B},\mathcal{B}}(f)$ sei die/eine **Jordan-Normalform von f** . Die Frage nach dem bestimmten oder unbestimmtem Artikel klärt dabei der folgende Satz:

Satz A.26 (Jordan-Normalform, Eindeutigkeit). Sei $f: V \rightarrow V$ wie in [Satz A.24](#). Sind J und J' zwei Jordan-Normalformen von f , so sind für alle i und alle λ die Anzahl der Jordan-Blöcke $J_i(\lambda)$, die in J bzw. J' vorkommen, identisch. Die beiden Matrizen J und J' unterscheiden sich also lediglich in der Reihenfolge der Blöcke.

Beweisskizze. Man überlegt sich zunächst, dass

$$\ker((f - \lambda \text{id}_V)^i) = \ker((f_\lambda - \lambda \text{id}_{U_\lambda})^i) = \ker(g_\lambda^i)$$

für alle $i = 1, 2, \dots$ gilt (für die Notation, siehe [\(A.17\)](#)). In [\(A.23\)](#) hatten wir uns überlegt, dass man die Anzahl r_i der Jordan-Blöcke $J_i(0)$ von $g = g_\lambda$ (= Anzahl der Jordan-Blöcke $J_i(\lambda)$ von f) in unserem Existenzbeweis für die Normalform nilpotenter Matrizen aus den Dimensionen $\dim U_1, \dots, \dim U_{i_0}$ berechnen kann. Im Anschluss darauf hatten wir gesehen, dass die Matrix in [\(A.23\)](#) invertierbar ist. Das heißt, dass die Anzahlen der Jordan-Blöcke auch die Folge der Dimensionen $\dim U_1, \dim U_2, \dots$ eindeutig bestimmen. Damit sind diese Anzahlen allein durch f schon eindeutig bestimmt. \square

Wir geben noch eine Anleitung zur Bestimmung der Jordan-Normalform eines Endomorphismus:

Algorithmus A.27 (Bestimmung der Jordan-Normalform eines Endomorphismus).

- (*) **Eingabe:** Ein Endomorphismus $f: V \rightarrow V$ eines endlich-dimensionalen K -Vektorraums V .
- (†) **Ausgabe:** Jordan-Normalform von f , falls diese existiert.
- (1) Berechne das charakteristische Polynom χ_f von f .
- (2) Zerlege χ_f in Linearfaktoren. Falls χ_f nicht in Linearfaktoren zerfällt, so hat f keine Jordan-Normalform. Dann breche den Algorithmus hier ab.
- (3) Berechne die Nullstellenordnung $m_\lambda = \nu(\chi_f, \lambda)$ für jeden Eigenwert von f . (Das ist die Anzahl der Faktoren $X - \lambda$, die in der Zerlegung von χ_f vorkommen.)
- (4) Für jeden Eigenwert λ von f berechne $d_i(\lambda) = \dim \ker((f - \lambda \text{id}_V)^i)$ für $i = 1, 2, \dots$, bis diese Dimension m_λ erreicht. Das i für welches dies geschieht, bezeichne mit $i_0(\lambda)$.
- (5) Berechne $r_i(\lambda)$ aus den $d_i(\lambda)$ (für $i = 1, 2, \dots, i_0(\lambda)$) gemäß [\(A.23\)](#).
- (6) Die Jordan-Normalform von f besteht aus $r_i(\lambda)$ Blöcken $J_i(\lambda)$; i läuft hier von 1 bis $i_0(\lambda)$ und λ durchläuft die Eigenwerte von f .

A.3.4. Ähnlichkeitsuntersuchungen. Die Jordan-Normalform erlaubt es uns, Endomorphismen auf Ähnlichkeit zu untersuchen.

Korollar A.28. Seien $f, f': V \rightarrow V$ zwei Endomorphismen eines endlich-dimensionalen K -Vektorraums V . Ferner mögen die charakteristischen Polynome von f und f' jeweils zerfallen. Dann sind f und f' genau dann ähnlich, wenn f und f' dieselbe Jordan-Normalform haben.

Beweis. Angenommen f und f' sind ähnlich und $J = M_{\mathcal{B}, \mathcal{B}}(f)$ die Jordan-Normalform von f bezüglich einer geeigneten Basis \mathcal{B} von V . Ähnlichkeit bedeutet, dass es einen Isomorphismus $\phi: V \rightarrow V$ gibt, sodass folgendes Diagramm kommutiert:

$$\begin{array}{ccc} V & \xrightarrow{f'} & V \\ & \searrow \phi \sim & \nearrow \phi \sim \\ & V & \xrightarrow{f} & V \end{array}$$

Sei $n = \dim V = \#\mathcal{B}$. Wir erweitern unser Diagramm wie folgt:

$$\begin{array}{ccccc} V & \xrightarrow{f'} & & & V \\ & \searrow \phi \sim & & & \nearrow \phi \sim \\ & & V & \xrightarrow{f} & V \\ & \nearrow \varphi_{\mathcal{B}} \sim & & & \searrow \varphi_{\mathcal{B}} \sim \\ K^n & \xrightarrow{v \mapsto Jv} & & & K^n \end{array}$$

$\phi \circ \varphi_{\mathcal{B}}$ \uparrow $\varphi_{\mathcal{B}}$ \downarrow $\phi \circ \varphi_{\mathcal{B}}$

Die Abbildung $\phi \circ \varphi_{\mathcal{B}}$ ist die Basisabbildung $\varphi_{\phi(\mathcal{B})}: K^n \rightarrow V$ bezüglich der Basis $\phi(\mathcal{B})$ ($= \phi$ auf alle Elemente von \mathcal{B} angewendet). Also ist $J = M_{\phi(\mathcal{B}), \phi(\mathcal{B})}(f')$. Das zeigt, dass J Jordan-Normalform von f und f' ist.

Der Beweis, dass wenn f und f' dieselbe Jordan-Normalform J haben, f und f' dann auch ähnlich sind verläuft ähnlich. Dieser sei den Leserinnen und Lesern selbst überlassen. Als kleinen Abstrich bei deinem Argument gilt es noch zu beachten, dass hier der Fall offen gelassen wurde, dass f und f' jeweils Jordan-Normalform J und J' haben, aber die Matrizen J und J' nicht identisch, sondern nur im Sinne von [Satz A.26](#) gleich sind, also jeweils dieselbe Anzahl an Jordan-Blöcken zu allen Eigenwerten aufweisen. In diesem Fall lässt sich allerdings durch Umordnung der zugehörigen Basisvektoren für die Darstellung von f' als J' einsehen, dass J auch Jordan-Normalform von f' ist. Eine Instanz davon hatten wir schon in [\(A.24\)](#) in [Beispiel A.23](#) gesehen. Die exakte Vorgehensweise aufzuschreiben ufert nur in einer weiteren Indexschlacht aus und bringt vermutlich wenig Einsicht mit sich. Wir verzichten daher auf weitere Ausführungen dieser Art. \square

Bemerkung A.29 (Ausblick auf weitere Normalformen). [Satz A.24](#) trifft leider keine Aussage für Endomorphismen f deren charakteristisches Polynom χ_f nicht in Linearfaktoren zerfällt. Ja, tatsächlich ist eine Transformation auf Jordan-Normalform dann auch gar nicht möglich, da das charakteristische Polynom einer Jordan-Matrix (oder allgemeiner: einer beliebigen oberen Dreiecksmatrix) stets in Linearfaktoren zerfällt. Transformierbarkeit auf Jordan-Normalform zieht also das Zerfallen von χ_f unausweichlich nach sich.

Für Endomorphismen f mit nicht notwendiger Weise zerfallendem charakteristischem Polynom χ_f steht auch eine verallgemeinerung der Jordan-Normalform zur Verfügung, die beim Zerfallen von χ_f in Linearfaktoren mit der Jordan-Normalform

übereinstimmt. Eine solche Normalform ist als **Frobenius-Normalform** (oder **rationale Normalform**) bekannt. Grundlage für jene ist eine Verallgemeinerung der Hauptraumzerlegung, **Satz A.16**, die sogenannte **Primärzerlegung** (siehe [3, Korollar 8.3.10]). Für einen konzeptionelleren Zugang, der aber mehr algebraisches Vorwissen erfordert, siehe [6, § 8.3].

A.4. Hauptraumzerlegung: Der Beweis

In diesem Abschnitt sei $f: V \rightarrow V$ weiterhin ein Endomorphismus eines endlich-dimensionalen K -Vektorraums V . Wir schreiben $n = \dim V$. Wir haben mittlerweile gesehen, dass die Hauptraumzerlegung, **Satz A.16**, fundamental für den Existenzsatz über die Jordan-Normalform (**Satz A.24**) ist. Darum ist es höchste Zeit, dass wir uns an den Beweis von **Satz A.16** wagen. Unsere Darstellung orientiert sich stark an einer bekannten Arbeit von Sheldon Axler [1]. Wir benötigen zunächst einige Hilfsresultate. Für $n = 0$ ist **Satz A.16** die Hauptraumzerlegung trivialerweise richtig (und überhaupt haben wir auch nichts Spannendes über den Nullvektorraum zu erzählen). Demnach sei im Folgenden auch stets $1 \leq n < \infty$ vorausgesetzt.

Lemma A.30. *Es gibt ein $N \in \mathbb{N}$, sodass für alle λ die Gleichung*

$$\text{Hau}(f, \lambda) = \ker((\lambda \text{id}_V - f)^N)$$

besteht.

Beweis. Wenn λ kein Eigenwert von f ist, gilt die Gleichung schon mit $N = 1$. Nun durchlaufe λ die endlich vielen⁷ Eigenwerte von f . Für jedes solche λ können in der Kette

$$(A.26) \quad \ker((\lambda \text{id}_V - f)^1) \subseteq \ker((\lambda \text{id}_V - f)^2) \subseteq \ker((\lambda \text{id}_V - f)^3) \subseteq \dots$$

nur endlich viele Inklusionen strikt sein. Es gibt also einen Exponenten $N(\lambda) \in \mathbb{N}$ mit

$$\text{Hau}(f, \lambda) = \ker((\lambda \text{id}_V - f)^{N(\lambda)}).$$

(Dieses Argument hatten wir auch schon im Beweis von **Lemma A.19** angeführt.)

Wähle nun⁸ $N = \max\{N(\lambda) : \text{Eigenwerte } \lambda \text{ von } f\}$. □

Bemerkung. Axler [1, Lemma 3.1] zeigt (in elementarer Weise), dass man in **Lemma A.30** sogar $N = \dim V$ wählen kann. Wir könnten hier auch leicht dieses schärfere Resultat erhalten: Aus **Lemma A.21 (2)** folgt nämlich, dass in der Kette (A.26) ab dem ersten „ \subseteq “ in dem Gleichheit besteht, fortan auch weiterhin stets Gleichheit in der Kette besteht. Die dafür nötige Schlussweise kennen wir bereits aus Teil (2) der Beispiele auf **Seite 31**.

⁷Kombiniere [3, Korollar 6.3.10, § 6.3.5 und Korollar 6.2.12].

⁸Um den Fall $\dim V = 0$ auch einzuschließen sollte man hier $N = \max(\{1\} \cup \{N(\lambda) : \text{Eigenwerte } \lambda \text{ von } f\})$ schreiben. Laut unserer Grundannahme vom Beginn des Abschnittes ist aber ohnehin $\dim V \geq 1$.

Das nächste Ergebnis ist eine Verallgemeinerung der bekannten Aussage, dass Eigenvektoren zu paarweise verschiedenen Eigenwerten linear unabhängig sind, auf Haupträume. Man vergleiche mit [Aufgabe 2.2](#) (und deren Lösung) bzw. mit [[4](#), Lemma 6.4.8]. Die Beweismethode ist grundlegend dieselbe, präsentiert sich im hiesigen Fall für Haupträume aber in etwas komplizierterem Gewandt.

Lemma A.31. *Seien $\lambda_1, \dots, \lambda_m$ paarweise verschiedene Eigenwerte von f und $v_r \in \text{Hau}(f, \lambda_r)$ von Null verschieden für $r = 1, \dots, m$. Dann sind die Vektoren v_1, \dots, v_m linear unabhängig.*

Beweis. Wir schreiben $v = v_1$. Betrachte Skalare $\mu, \mu_2, \dots, \mu_m \in K$ mit

$$(A.27) \quad \mu v + \mu_2 v_2 + \dots + \mu_m v_m = 0.$$

Wir müssen zeigen, dass die Skalare μ_r alle gleich Null sind. Dazu genügt es $\mu = 0$ zu zeigen und die Aussage folgt dann per Induktion. Sei $i \in \mathbb{N}$ minimal mit $v \in \ker((\lambda \text{id}_V - f)^i)$. Wir betrachten nun den Endomorphismus

$$h = (\lambda \text{id}_V - f)^{i-1} \circ (\lambda_2 \text{id}_V - f)^N \circ \dots \circ (\lambda_m \text{id}_V - f)^N$$

mit N wie aus [Lemma A.30](#). Darin schreiben darin nun jeden bis auf den ersten „Faktor“ wie folgt um:

$$(\lambda_r \text{id}_V - f)^N = ((\lambda \text{id}_V - f) + (\lambda_r - \lambda) \text{id}_V)^N \quad (r = 2, 3, \dots, m).$$

Man zeigt nun leicht mittels Induktion durch „Entpacken“ der obigen Verkettungen, dass sich h in der Form

$$(A.28) \quad h = (\lambda \text{id}_V - f)^{i-1} \circ \left(\prod_{r=2}^m (\lambda_r - \lambda)^N \cdot \text{id}_V + \sum \dots \right),$$

schreiben lässt, wobei die Summe \sum nur Terme involviert, die mit einem Faktor $\lambda \text{id}_V - f$ beginnen. Wir wenden nun h auf [\(A.27\)](#) an. Die rechte Seite bleibt 0. Auf der linken Seite gilt einerseits $h(v_r) = 0$ für alle $r = 2, \dots, m$, denn $(\lambda_r \text{id}_V - f)^N(v) = 0$ dank [Lemma A.30](#). Darum wird die linke Seite von [\(A.27\)](#) nach Anwenden von h also zu $\mu h(v)$. Andererseits zeigt unsere zweite Darstellung [\(A.28\)](#) für h aber

$$h(v) = \prod_{r=2}^m (\lambda_r - \lambda)^N \cdot v.$$

Daraus ergibt sich $\mu h(v) = h(\mu v + \mu_2 v_2 + \dots + \mu_m v_m) = 0$ und also $\mu = 0$. □

Lemma A.32. *Sei f wie oben, λ sei ein Eigenwert von f und N sei wie in [Lemma A.30](#). Schreibe*

$$U = \ker((\lambda \text{id}_V - f)^N) \quad \text{und} \quad W = \text{im}((\lambda \text{id}_V - f)^N).$$

Dann gelten die folgenden Aussagen:

- (1) $V = U \oplus W$,
- (2) $U = \text{Hau}(f, \lambda)$,
- (3) U ist f -invariant,

- (4) W ist f -invariant,
 (5) $\dim W < n$,
 (6) für $h := f|_W: W \rightarrow W$ zerfällt χ_h in Linearfaktoren und die Eigenwerte μ von h sind die Eigenwerte $\neq \lambda$ von f .

Beweis. (1) ist beinahe schon aus dem Beweis der Dimensionsformel für lineare Abbildungen wohl-bekannt (siehe [3, Satz 4.2.11]). Wir schreiben $h = (\lambda \operatorname{id}_V - f)^N$ und zeigen $U \cap W = \{0\}$. Ist dies getan, so folgt bereits $V = U \oplus W$ aus Dimensionsgründen (dank der zitierten Dimensionsformel). Sei also $w \in U \cap W$. Dann ist $w = h(v)$ für ein $v \in V$ und wegen $w \in U$ folgt $0 = h(w) = h^2(v)$. Allerdings muss dann schon $w = h(v) = 0$ gelten, denn $\ker h = \ker(h^2)$ gemäß unserer Wahl von N .

(2) ist Aussage von **Lemma A.30**.

(3) folgt wegen (2) aus **Lemma A.18**.

(4) beweist man analog (siehe die Rechnung in **Lemma A.18**).

(5) folgt aus (1) und der Dimensionsformel, denn wegen $U \supseteq \operatorname{Eig}(f, \lambda)$ und $\operatorname{Eig}(f, \lambda) \neq \{0\}$ (nach Voraussetzung) ist $\dim U \geq \dim \operatorname{Eig}(f, \lambda) > 0$ und also $\dim W = \dim V - \dim U < n$.

(6) Da es sich bei h um eine Einschränkung von f handelt, ist jedes Eigenwert-Eigenvektor-Paar von h auch ein solches für f . Der Eigenwert λ von f kann nicht als Eigenwert von h vorkommen, da die zugehörigen Eigenvektoren alle in U liegen und sich W nach (1) nur trivial mit U schneidet.

Nun zur Aussage über χ_h . Wählt man Basen von U und W und fügt diese zu einer Basis von V zusammen, so hat die Darstellungsmatrix von f bezüglich dieser Basen dank (3) und (4) Blockdiagonalgestalt. Es folgt wie in der Lösung von **Aufgabe 10.4**, dass

$$(A.29) \quad \prod_{\lambda'} (X - \lambda')^{\mu(\chi_f, \lambda')} = \chi_f = \chi_h \cdot \chi_{f|_U}$$

gilt. Mit Argumenten wie aus dem Beweis von **Aufgabe 3.4** folgt nun, dass χ_h zerfällt: Jede Nullstelle von χ_f muss auch Nullstelle von χ_h oder $\chi_{f|_U}$ sein. In jedem Fall kann man von einem der beiden Polynome einen Linearfaktor abspalten und diesen dank **Aufgabe 3.4** (a) aus der gesamten Gleichung herauskürzen. Durch Wiederholung dieses Prozesses erhält man, dass χ_h zerfällt. \square

Bemerkung. In der Situation von (A.29) könnte man tatsächlich

$$\chi_{f|_U} = (X - \lambda)^{\mu(\chi_f, \lambda)} \quad \text{und} \quad \chi_h = \prod_{\lambda' \neq \lambda} (X - \lambda')^{\mu(\chi_f, \lambda')}$$

zeigen. Dieses Detail ist für das Gelingen des Beweises allerdings unerheblich.

Beweis von Satz A.16. Wir beweisen die Aussage per Induktion über $n = \dim V$. Der Induktionsanfang $n = 0$ ist klar. (Dann ist $f: V \rightarrow V$ die Nullabbildung $\{0\} \rightarrow \{0\}$ und es gilt $\operatorname{Hau}(f, 0) = \{0\} = V$.)

Lösung besitzt. Dies ist äquivalent dazu, dass die Matrix

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \in K^{n \times n}$$

invertierbar ist. In dieser Situation gilt

$$x \text{ löst (A.30)} \iff Ax = b \iff x = A^{-1}b.$$

Aus diesem Grund sind wir daran interessiert A^{-1} auszurechnen. Hierfür kann der bekannte **Gauß-Jordan-Algorithmus** verwendet werden, der ausgehend von der Blockmatrix $(A \mid E_n)$ durch elementare Zeilenumformungen zu einer Blockmatrix $(E_n \mid B)$ mit einem $B \in K^{n \times n}$ übergeht. Dann gilt $B = A^{-1}$ (siehe [3, Algorithmus 4.4.10]).

A.5.1. Der 2×2 -Fall. Wir sind daher an einer Formel für A^{-1} interessiert. Aus [3, § 5.0.1] kennen wir bereits die Formel

$$(A.31) \quad A^{-1} = \frac{1}{\det A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \quad \text{für } A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Diese Formel bestätigt man leicht, indem man

$$(A.32) \quad \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \det A & 0 \\ 0 & \det A \end{pmatrix}$$

verifiziert. Tatsächlich ist die Formel (A.32) der Formel (A.31) überlegen, indes (A.32) auch für nicht invertierbare Matrizen ihre Gültigkeit behält.

Wir zeigen, dass man mittels des Gauß-Jordan-Algorithmus auf diese Formel kommen kann. Wir beginnen mit

$$(A \mid E_2) = \left(\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array} \right)$$

Wir gehen nun davon aus, dass $a \neq 0$ gilt. Das ist natürlich eine *echte Annahme* und ist nicht für jede Matrix $A \in K^{2 \times 2}$ erfüllt. — Auch nicht für jede invertierbare Matrix A . Tatsächlich soll uns das aber nicht stören: Sobald wir die gesuchte Formel für A^{-1} gefunden haben, lässt sich diese ja auch von Hand verifizieren und es ist egal, ob während des Prozesses der Formel-Findung einige potentiell ungerechtfertigte Annahmen getroffen haben. Wir multiplizieren nun die zweite Spalte mit a :

$$(A \mid E_2) \rightsquigarrow \left(\begin{array}{cc|cc} a & b & 1 & 0 \\ ac & ad & 0 & a \end{array} \right).$$

Nun ziehen wir das c -fache der ersten Zeile von der zweiten Zeile ab und erhalten somit

$$(A \mid E_2) \rightsquigarrow \left(\begin{array}{cc|cc} a & b & 1 & 0 \\ 0 & ad - cb & -c & a \end{array} \right) = \left(\begin{array}{cc|cc} a & b & 1 & 0 \\ 0 & \det A & -c & a \end{array} \right).$$

Da wir mit einer invertierbaren Matrix A arbeiten gilt $\det A \neq 0$ und wir dürfen das $b/\det A$ -fache der zweiten Zeile von der ersten abziehen. Daraus ergibt sich

$$(A | E_2) \rightsquigarrow \left(\begin{array}{cc|cc} a & 0 & 1 + \frac{bc}{\det A} & \frac{-ab}{\det A} \\ 0 & \det A & -c & a \end{array} \right).$$

Durch Division der Zeilen durch a bzw. $(\det A)/a$ entsteht nun schließlich

$$\begin{aligned} (A | E_2) &\rightsquigarrow \left(\begin{array}{cc|cc} 1 & 0 & \frac{1}{a} + \frac{bc}{\det A} \frac{1}{a} & \frac{-ab}{\det A} \frac{1}{a} \\ 0 & 1 & \frac{-c}{\det A} & \frac{a}{\det A} \end{array} \right) = \left(\begin{array}{cc|cc} 1 & 0 & \frac{\det A + bc}{a \det A} & \frac{-b}{\det A} \\ 0 & 1 & \frac{-c}{\det A} & \frac{a}{\det A} \end{array} \right) \\ &= \left(\begin{array}{cc|cc} 1 & 0 & \frac{d}{\det A} & \frac{-b}{\det A} \\ 0 & 1 & \frac{-c}{\det A} & \frac{a}{\det A} \end{array} \right) = \left(E_2 \mid \frac{1}{\det A} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \right). \end{aligned}$$

Die letzte Zeile ist von der Form $(E_2 | B)$, wobei die Matrix B durch genau der in (A.31) als A^{-1} angegebenen Matrix entspricht.

A.5.2. Der $n \times n$ -Fall: Vorbereitung. Man könnte nun in derselben Weise eine Formel für die Inverse einer invertierbaren 3×3 -Matrix, einer 4×4 -Matrix, etc. angeben. Es braucht nicht viel Fantasie, um sich vorzustellen, dass die dabei entstehenden Rechnungen schnell unerfreulich werden. Es ist daher an der Zeit systematischer vorzugehen. Wir beachten zunächst, dass die Inverse A^{-1} von $A \in K^{n \times n}$ genau die Matrix $B \in K^{n \times n}$ ist, für welche $BA = E_n$ gilt.¹⁰ Für $B = (b_{ij})_{i,j=1}^n$ und $A = (a_{ij})_{i,j=1}^n$ schreibt sich die Gleichung $E_n = BA$ als

$$\begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} = \begin{pmatrix} - & b_{1\bullet} & - \\ \vdots & \vdots & \vdots \\ - & b_{n\bullet} & - \end{pmatrix} \begin{pmatrix} | & \dots & | \\ a_{\bullet 1} & \dots & a_{\bullet n} \\ | & \dots & | \end{pmatrix} = \begin{pmatrix} b_{1\bullet} a_{\bullet 1} & \dots & b_{1\bullet} a_{\bullet n} \\ \vdots & \ddots & \vdots \\ b_{n\bullet} a_{\bullet 1} & \dots & b_{n\bullet} a_{\bullet n} \end{pmatrix},$$

mit

$$b_{i\bullet} a_{\bullet j} = \sum_{\nu=1}^n b_{i\nu} a_{\nu j} \in K^{1 \times 1} = K.$$

Wir wollen also Zeilen $b_{i\bullet} \in K^{1 \times n}$ ($i = 1, \dots, n$) mit

$$(A.33) \quad b_{i\bullet} a_{\bullet j} = \begin{cases} 1 & \text{falls } i = j, \\ 0 & \text{falls } i \neq j, \end{cases} \quad \text{für } j = 1, \dots, n$$

finden.

¹⁰Aus $BA = E_n$ folgt nämlich aus Dimensionsgründen schon die Invertierbarkeit von A und somit die Existenz von A^{-1} und durch Rechtsmultiplikation mit A^{-1} erhält man sodann $B = A^{-1}$. Weiß man umgekehrt schon, dass A invertierbar ist, so liefert Rechtsmultiplikation mit A^{-1} auch schon $B = A^{-1}$.

A.5.3. Der 3×3 -Fall. Wir denken für den Moment an den Fall $n = 3$ und $K = \mathbb{R}$ und fassen die Zeile $b_{i\bullet} \in K^{1 \times n}$ als den Vektor $b_{i\bullet}^\sharp := b_{i\bullet}^T \in K^{n \times 1} = K^n$ auf.¹¹ Dann ist

$$b_{i\bullet} a_{\bullet j} = (b_{i\bullet}^\sharp)^T a_{\bullet j} = \langle b_{i\bullet}^\sharp, a_{\bullet j} \rangle,$$

mit dem Standardskalarprodukt $\langle _, _ \rangle$ auf dem \mathbb{R}^3 . Die Gleichung (A.33) sagt uns also, dass wir zu den drei Vektoren $a_{\bullet 1}, a_{\bullet 2}, a_{\bullet 3}$ drei weitere Vektoren $b_{1\bullet}^\sharp, b_{2\bullet}^\sharp, b_{3\bullet}^\sharp$ mit

$$(A.34) \quad \begin{cases} \langle b_{1\bullet}^\sharp, a_{\bullet 1} \rangle = 1, & b_{1\bullet}^\sharp \perp a_{\bullet 2}, & b_{1\bullet}^\sharp \perp a_{\bullet 3}, \\ b_{2\bullet}^\sharp \perp a_{\bullet 1}, & \langle b_{2\bullet}^\sharp, a_{\bullet 2} \rangle = 1, & b_{2\bullet}^\sharp \perp a_{\bullet 3}, \\ b_{3\bullet}^\sharp \perp a_{\bullet 1}, & b_{3\bullet}^\sharp \perp a_{\bullet 2}, & \langle b_{3\bullet}^\sharp, a_{\bullet 3} \rangle = 1, \end{cases}$$

suchen.

Wie so etwas geht wissen wir schon aus Aufgabe 6.4. Wir definieren wie dort für Vektoren $v, w \in K^3$

$$(A.35) \quad v \times w := \text{„det} \begin{pmatrix} e_1 & | & | \\ e_2 & v & w \\ e_3 & | & | \end{pmatrix} \text{“} := e_1 \det \begin{pmatrix} v_2 & w_2 \\ v_3 & w_3 \end{pmatrix} - e_2 \det \begin{pmatrix} v_1 & w_1 \\ v_3 & w_3 \end{pmatrix} \\ + e_3 \det \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix}.$$

Diese Formel ergibt auch für von \mathbb{R} verschiedene Körper K Sinn. Wie in Aufgabe 6.4 bestätigt man (auch für beliebige Körper)¹²

$$(A.36) \quad \langle v \times w, z \rangle = \langle z, v \times w \rangle = z^T(v \times w) = \det \begin{pmatrix} | & | & | \\ z & v & w \\ | & | & | \end{pmatrix}$$

für alle $z \in K^3$. Insbesondere gilt

$$(A.37) \quad \langle z, v \times w \rangle = 0 \quad \text{für} \quad z \in \{v, w\}.$$

Wir wählen nun

$$b_{1\bullet}^\sharp := \frac{1}{\det A} (a_{\bullet 2} \times a_{\bullet 3}), \quad b_{2\bullet}^\sharp := -\frac{1}{\det A} (a_{\bullet 1} \times a_{\bullet 3}), \quad b_{3\bullet}^\sharp := \frac{1}{\det A} (a_{\bullet 1} \times a_{\bullet 2}).$$

Unter Verwendung von (A.36) für die Diagonaleinträge und (A.37) für die übrigen Einträge erhalten wir (A.34). Wir haben also die folgende Formel für die Inverse A^{-1}

¹¹Das Zeichen „ \sharp “ kennt man aus der Musik und dort wird es im Englischen „sharp“ ausgesprochen. Diese Notation wählen wir hier, da wir die Zeile $b_{i\bullet} \in K^{1 \times n}$ zu dem Vektor $b_{i\bullet}^\sharp \in K^{n \times 1} = K^n$ „anspitzen“.

¹²Für beliebige Körper K ist $\langle _, _ \rangle: V \times V \rightarrow K, (v, w) \mapsto v^T w$, zwar im Allgemeinen kein Skalarprodukt mehr, sondern bloß noch eine nicht ausgeartete Bilinearform, aber das muss uns nicht stören.

zu A :

$$(A.38) \quad A^{-1} = \begin{pmatrix} | & | & | \\ a_{\bullet 1} & a_{\bullet 2} & a_{\bullet 3} \\ | & | & | \end{pmatrix}^{-1} = \frac{1}{\det A} \begin{pmatrix} - & +(a_{\bullet 2} \times a_{\bullet 3})^T & - \\ - & -(a_{\bullet 1} \times a_{\bullet 3})^T & - \\ - & +(a_{\bullet 1} \times a_{\bullet 2})^T & - \end{pmatrix}.$$

A.5.4. Der $n \times n$ -Fall. Wir verallgemeinern nun (A.35) auf Vektoren in K^n , wobei K wieder ein völlig beliebiger Körper sein darf. Die rechte Seite von (A.35) stammt von einer formal angewendeten Laplace-Entwicklung nach der ersten Spalte (siehe [3, Satz 5.4.2]). Diese Entwicklung benutzt Streichungsmatrizen. Für Vektoren v_1, \dots, v_{n-1} schreiben wir $V_{i\bullet}$ für die $(n-1) \times (n-1)$ -Matrix, welche man erhält, wenn man in der $n \times (n-1)$ -Matrix

$$\begin{pmatrix} | & \dots & | \\ v_1 & \dots & v_{n-1} \\ | & \dots & | \end{pmatrix}$$

die i -te Spalte streicht. Wir definieren dann

$$(A.39) \quad v_1 \times \dots \times v_{n-1} := \text{„det} \begin{pmatrix} e_1 & | & \dots & | \\ \vdots & v_1 & \dots & v_{n-1} \\ e_n & | & \dots & | \end{pmatrix} \text{“} := \sum_{i=1}^n (-1)^i e_i \det(V_{i\bullet}).$$

Sei $z = (z_1, \dots, z_n) \in K^n$. Aus dem Laplaceschen Entwicklungssatz [3, Satz 5.4.2] folgt sofort die natürliche Verallgemeinerung von (A.36):

$$(A.40) \quad z^T(v_1 \times \dots \times v_{n-1}) = \sum_{i=1}^n (-1)^i z_i \det(V_{i\bullet}) = \det \begin{pmatrix} | & | & \dots & | \\ z & v_1 & \dots & v_{n-1} \\ | & | & \dots & | \end{pmatrix}.$$

Aus jener erhalten wir auch sofort die folgende Verallgemeinerung von (A.37):

$$(A.41) \quad z^T(v_1 \times \dots \times v_{n-1}) = 0 \quad \text{für } z \in \{v_1, \dots, v_{n-1}\}.$$

Damit erhalten wir die folgende Verallgemeinerung von (A.38):

Satz A.33 (Cramersche Regel). Sei $A = (a_{ij})_{i,j=1}^n$ eine $n \times n$ -Matrix. Setze

$$(A.42) \quad A^{\text{adj}} = \begin{pmatrix} - & +(a_{\bullet 2} \times a_{\bullet 3} \times a_{\bullet 4} \times \dots \times a_{\bullet n-1} \times a_{\bullet n})^T & - \\ - & -(a_{\bullet 1} \times a_{\bullet 3} \times a_{\bullet 4} \times \dots \times a_{\bullet n-1} \times a_{\bullet n})^T & - \\ - & +(a_{\bullet 1} \times a_{\bullet 2} \times a_{\bullet 4} \times \dots \times a_{\bullet n-1} \times a_{\bullet n})^T & - \\ & \vdots & \vdots & \vdots \\ - & (-1)^{n-2} (a_{\bullet 1} \times a_{\bullet 2} \times a_{\bullet 3} \times \dots \times a_{\bullet n-2} \times a_{\bullet n})^T & - \\ - & (-1)^{n-1} (a_{\bullet 1} \times a_{\bullet 2} \times a_{\bullet 3} \times \dots \times a_{\bullet n-2} \times a_{\bullet n-1})^T & - \end{pmatrix}.$$

(In der i -ten Zeile steht $(-1)^{i-1}$ mal das Kreuzprodukt der Vektoren $a_{\bullet 1}, \dots, a_{\bullet n}$, wobei der Vektor $a_{\bullet i}$ auszulassen ist.) Dann gilt die folgende Verallgemeinerung von (A.32):

$$(A.43) \quad A^{\text{adj}} A = \det(A) E_n = A A^{\text{adj}}.$$

Beweis. Die Zeilen von A^{adj} sind genau so konstruiert, dass beim Berechnen von $A^{\text{adj}}A$ für die Einträge abseits der Diagonale stets die Formel (A.41) zum Tragen kommt. In den Diagonaleinträgen findet man dank (A.40) stets die Determinante von A wieder; Dafür ist auch die in der Definition von A^{adj} gewählte alternierende Vorzeichenwahl essentiell. Das zeigt $A^{\text{adj}}A = \det(A)E_n$.

Für die zweite Gleichung $\det(A)E_n = AA^{\text{adj}}$ überlege man sich durch „scharfes Hinsehen“ auf (A.39) und die Definition von A^{adj} , dass $(A^{\text{adj}})^T = (A^T)^{\text{adj}}$ gilt. Aus dem bereits Bewiesenen folgt nun $(A^T)^{\text{adj}}A^T = \det(A^T)E_n = \det(A)E_n$ und also $(A^{\text{adj}})^T A^T = \det(A)E_n$. Durch Transponieren dieser Gleichung erhält man nun $\det(A)E_n = AA^{\text{adj}}$. \square

Die Matrix A^{adj} wird **Adjunkte** von A , oder auch **klassische Adjungierte** von A oder **komplementäre Matrix** zu A genannt. Trotz des Namens „Adjungierte“ sollte man A^{adj} nicht mit der adjungierten Matrix A^* im Falle des euklidischen Raumes \mathbb{R}^n oder des unitären Raumes \mathbb{C}^n verwechseln. Tatsächlich sind diese Begriffe dennoch eng miteinander verwandt, aber um diesen Sachverhalt darlegen zu können müssten wir zunächst einen tiefen Streifzug durch Dualitätstheorie und äußere Potenzen von Vektorräumen unternehmen. Dazu fehlt uns hier die Zeit.

Bemerkung. Im Skript [4] von Hans Franzen findet man **Satz A.33** als Satz 8.1.8. Dort werden die Einträge von A^{adj} direkt angegeben, während wir A^{adj} in (A.42) zeilenweise angeben. Beide Formeln liefern dasselbe Ergebnis und es kann als gute Übung dienen, sich von diesem Sachverhalt zu überzeugen.

A.5.5. Noch einmal der 2×2 -Fall.

Beispiel A.34 (Cramersche Regel im 2×2 -Fall). Wir überlegen uns noch, dass die Formel (A.42) aus **Satz A.33** sich für $n = 2$ wirklich auf die Formel (A.32) reduziert. Wir betrachten die Matrix

$$A = \begin{pmatrix} | & | \\ a_1 & a_2 \\ | & | \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in K^{2 \times 2}, \quad \text{also} \quad a_1 = \begin{pmatrix} a \\ c \end{pmatrix}, \quad a_2 = \begin{pmatrix} b \\ d \end{pmatrix}.$$

Die Formel (A.39) definiert ein Kreuzprodukt von $n - 1$ Vektoren im K^n . Im aktuellen Fall gilt $n = 2$ und wir sprechen von einem *Kreuzprodukt von einem Vektor* $v = (x, y) \in K^2$. Um dieses Kreuzprodukt von v selbst unterscheiden zu können schreiben wir

$$\times(v) := \text{das Kreuzprodukt von } v \in K^2 \text{ gemäß (A.39)}.$$

Dann gilt

$$\times(v) = \text{„det} \begin{pmatrix} e_1 & | \\ e_2 & v \\ & | \end{pmatrix} \text{“} = e_1 \cdot (\text{zweiter Eintrag von } v) - e_2 \cdot (\text{erster Eintrag von } v).$$

(Man beachte, dass $\times(v)$ im Allgemeinen *nicht* gleich der Vektor v ist. Es ist daher wirklich wichtig, dass wir nicht der Versuchung erliegen, „ v “ für $\times(v)$ zu schreiben.) Damit erhalten wir

$$\times(a_1) = \times\begin{pmatrix} a \\ c \end{pmatrix} = \begin{pmatrix} c \\ -a \end{pmatrix} \quad \text{und} \quad \times(a_2) = \times\begin{pmatrix} b \\ d \end{pmatrix} = \begin{pmatrix} d \\ -b \end{pmatrix}.$$

Die Adjunkte A^{adj} von A aus (A.42) nimmt die Gestalt

$$A^{\text{adj}} = \begin{pmatrix} - & \times(a_2)^{\text{T}} & - \\ - & -\times(a_1)^{\text{T}} & - \end{pmatrix} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

an.¹³ Die Gleichung $A^{\text{adj}}A = (\det A)E_2$ aus (A.43) in Satz A.33 ist also tatsächlich genau (A.32).

¹³Man achte hier penibel auf die Reihenfolge der a s und die Vorzeichen!

Literaturverzeichnis

- [1] S. Axler. Down with determinants! *Am. Math. Mon.*, 102(2):139–154, 1995.
- [2] G. Fischer and B. Springborn. *Lineare Algebra. Eine Einführung für Studienanfänger*. Berlin: Springer, 19. edition, 2020.
- [3] H. Franzen. Skript zur Vorlesung Lineare Algebra I. Online verfügbar: <https://panda.uni-paderborn.de/course/view.php?id=55381>, 2024.
- [4] H. Franzen. Skript zur Vorlesung Lineare Algebra II. Online verfügbar: <https://panda.uni-paderborn.de/course/view.php?id=55381>, 2024.
- [5] S. Lang. *Algebra*. New York, NY: Springer, 3rd edition, 2002.
- [6] M. Technau. Algebra. Vorlesungsskriptum. Online verfügbar: <https://www.math.tugraz.at/~mtechnau/downloads/algebra.pdf>, 2021.