

20. Die Jordansche Normalform

Im letzten Kapitel haben wir bereits ausführlich untersucht, wie man zu einer gegebenen quadratischen Matrix $A \in K^{n \times n}$ eine möglichst einfache ähnliche Matrix finden kann (bzw. wie man zu einem Endomorphismus $f: V \rightarrow V$ eines endlich-dimensionalen Vektorraums V eine Basis B von V finden kann, so dass die zugehörige Abbildungsmatrix A_f^B möglichst einfach wird). Wir haben dabei in Folgerung 19.40 gesehen, dass A genau unter den Bedingungen

- (a) χ_A zerfällt in Linearfaktoren, und
- (b) $\mu_g(A, \lambda) = \mu_a(A, \lambda)$ für alle Eigenwerte λ

diagonalisierbar, also zu einer Diagonalmatrix ähnlich ist, und dass eine solche Diagonalform in diesem Fall auch das bestmögliche Ergebnis ist (siehe Lemma 19.33). Wir wollen uns in diesem Kapitel nun anschauen, was wir erreichen können, wenn eine der beiden obigen Bedingungen nicht erfüllt ist (und A nach Folgerung 19.40 damit nicht diagonalisierbar ist).

Die Bedingung (a), dass das charakteristische Polynom in Linearfaktoren zerfällt, ist dabei eher harmlos. Ist dies nämlich z. B. bei einer reellen Matrix $A \in \mathbb{R}^{n \times n}$ nicht der Fall, so können wir A einfach als komplexe Matrix $A \in \mathbb{C}^{n \times n}$ auffassen. Nach dem Fundamentalsatz der Algebra (siehe Satz 6.11) zerfällt χ_A dann zumindest über \mathbb{C} in Linearfaktoren, und somit können wir wieder eine Diagonalform erreichen (wenn auch mit komplexen Eigenwerten und Eigenvektoren). So war z. B. die Matrix A aus Beispiel 19.37 (a) mit charakteristischem Polynom $\chi_A(t) = t^2 + 1$ zwar über \mathbb{R} nicht diagonalisierbar, über \mathbb{C} ist aber $\chi_A(t) = (t + i)(t - i)$ und A damit diagonalisierbar nach Bemerkung 19.41 (a).

Wie in der Vorlesung „Einführung in die Algebra“ des zweiten Studienjahres gezeigt wird, lässt sich dieser Trick nicht nur für reelle Matrizen, sondern über jedem Grundkörper K anwenden: Man kann stets zu einem größeren Körper übergehen, in dem das charakteristische Polynom dann in Linearfaktoren zerfällt. Das Problem eines nicht zerfallenden charakteristischen Polynoms ist also recht einfach zu beheben, und daher wollen wir hier auch nicht mehr weiter darauf eingehen. Wir werden in diesem Kapitel in der Regel also Matrizen betrachten, deren charakteristisches Polynom zwar in Linearfaktoren zerfällt, für die aber i. A. nicht die Bedingung (b) oben gilt, und wollen wiederum nach einer möglichst einfachen ähnlichen Matrix suchen. Nach Satz 19.35 bedeutet dies dann $\mu_g(A, \lambda) < \mu_a(A, \lambda)$ für einen Eigenwert λ von A .

20.A Haupträume

Es sei $A \in K^{n \times n}$ eine quadratische Matrix, deren charakteristisches Polynom in Linearfaktoren zerfällt. Nach Lemma 19.27 (a) ist A genau dann diagonalisierbar, wenn es eine Basis von K^n aus Eigenvektoren von A gibt – diese können wir dann als Spalten in eine Matrix $T \in \text{GL}(n, K)$ schreiben, und erhalten für $T^{-1}AT$ eine Diagonalform.

Ist nun $\mu_g(A, \lambda) < \mu_a(A, \lambda)$ für mindestens einen Eigenwert λ von A , so haben wir im Beweis von Folgerung 19.36 gezeigt, dass wir nicht genug linear unabhängige Eigenvektoren für eine solche Basis finden. Unsere Strategie wird es in diesem Fall daher sein, den Begriff der Eigenvektoren etwas zu verallgemeinern, so dass einerseits von diesen verallgemeinerten Eigenvektoren stets genügend viele existieren, um aus ihnen eine Basis von K^n zusammenstellen und in eine invertierbare Matrix T schreiben zu können, und andererseits die transformierte Matrix $T^{-1}AT$ auch im Fall einer Basis aus derartigen verallgemeinerten Eigenvektoren noch recht einfach ist. Die Idee hierfür ist, statt Eigenvektoren, also Vektoren $x \neq 0$ mit $(\lambda E - A)x = 0$, nun solche zu betrachten, für die zumindest $(\lambda E - A)^r x = 0$ für irgendein $r \in \mathbb{N}$ gilt.

Konstruktion 20.1 (Haupträume). Es seien $A \in K^{n \times n}$ und $\lambda \in K$.

(a) Für $r \in \mathbb{N}$ heißt

$$H_r(A, \lambda) := \text{Ker}((\lambda E - A)^r) \leq K^n$$

der **verallgemeinerte Eigenraum** der Stufe r von A zu λ .

Offensichtlich ist $H_0(A, \lambda) = \text{Ker} E = \{0\}$ und $H_1(A, \lambda) = \text{Eig}(A, \lambda)$.

(b) Die verallgemeinerten Eigenräume bilden eine aufsteigende Kette

$$\{0\} = H_0(A, \lambda) \leq H_1(A, \lambda) \leq H_2(A, \lambda) \leq \dots$$

von Unterräumen von K^n : Ist nämlich $x \in H_r(A, \lambda)$ für ein $r \in \mathbb{N}$, also $(\lambda E - A)^r x = 0$, so ist natürlich auch $(\lambda E - A)^{r+1} x = (\lambda E - A)(\lambda E - A)^r x = 0$ und damit $x \in H_{r+1}(A, \lambda)$. Die Dimensionen der verallgemeinerten Eigenräume sind also monoton wachsend, können aber natürlich nicht größer als die Dimension n des umgebenden Raumes werden. Es muss also ein $r \in \mathbb{N}$ geben mit

$$H_r(A, \lambda) = H_{r+1}(A, \lambda) = H_{r+2}(A, \lambda) = \dots =: H(A, \lambda).$$

Dieser Unterraum $H(A, \lambda)$, den wir uns also als „Grenzwert der verallgemeinerten Eigenräume $H_r(A, \lambda)$ für $r \rightarrow \infty$ “ vorstellen können, heißt der **Hauptraum** von A zu λ .

(c) Für $r \in \mathbb{N}_{>0}$ nennt man die Vektoren $x \in H_r(A, \lambda) \setminus H_{r-1}(A, \lambda)$, für die also $(\lambda E - A)^r x = 0$ und $(\lambda E - A)^{r-1} x \neq 0$ gilt, **Hauptvektoren** oder **verallgemeinerte Eigenvektoren** der Stufe r von A zum Eigenwert λ .

Der Hauptraum besteht also aus dem Nullvektor zusammen mit den Hauptvektoren einer beliebigen Stufe. Die gewöhnlichen Eigenvektoren aus Definition 19.11 (a) sind gerade die Hauptvektoren der Stufe 1, nämlich die Elemente von $H_1(A, \lambda) \setminus H_0(A, \lambda) = \text{Eig}(A, \lambda) \setminus \{0\}$.

Bemerkung 20.2 (Haupträume von Endomorphismen). Natürlich gibt es zu Konstruktion 20.1 auch wieder eine entsprechende Version für Endomorphismen $f: V \rightarrow V$ eines endlich-dimensionalen Vektorraums V : Analog zu Bemerkung 19.12 (b) sind die verallgemeinerten Eigenräume in diesem Fall $H_r(f, \lambda) = \text{Ker}((\lambda \text{id}_V - f)^r)$, wobei die Potenz hier für die r -fache Verkettung von $\lambda \text{id}_V - f$ mit sich selbst steht.

Bemerkung 20.3. Sind $A \in K^{n \times n}$, $\lambda \in K$ und $r \in \mathbb{N}$ mit $H_r(A, \lambda) = H_{r+1}(A, \lambda)$, so dass also in der aufsteigenden Kette der verallgemeinerten Eigenräume in Konstruktion 20.1 (b) an irgendeiner Stelle die Gleichheit gilt, so gilt auch an der folgenden Stelle $H_{r+1}(A, \lambda) = H_{r+2}(A, \lambda)$ die Gleichheit, denn für alle $x \in H_{r+2}(A, \lambda)$ folgt dann

$$\begin{aligned} (\lambda E - A)^{r+1} (\lambda E - A)x &= 0 \\ \Rightarrow (\lambda E - A)x &\in H_{r+1}(A, \lambda) = H_r(A, \lambda) \\ \Rightarrow (\lambda E - A)^r (\lambda E - A)x &= 0, \end{aligned}$$

und damit auch $x \in H_{r+1}(A, \lambda)$. Daraus erhalten wir sofort:

(a) Aus der Gleichheit $H_r(A, \lambda) = H_{r+1}(A, \lambda)$ an einer Stelle $r \in \mathbb{N}$ folgt induktiv bereits

$$H_r(A, \lambda) = H_{r+1}(A, \lambda) = H_{r+2}(A, \lambda) = \dots,$$

und damit $H(A, \lambda) = H_r(A, \lambda)$.

(b) Speziell für $r = 0$ bedeutet (a): Ist $\lambda \in K$ kein Eigenwert von A , gibt es also keine Eigenvektoren zu λ und ist damit $H_1(A, \lambda) = \text{Eig}(A, \lambda) = \{0\} = H_0(A, \lambda)$, so folgt auch $H(A, \lambda) = \{0\}$, d. h. dann gibt es auch keine verallgemeinerten Eigenvektoren zu λ . Anders formuliert verallgemeinern wir also nur den Begriff der **Eigenvektoren**, aber nicht den der **Eigenwerte**.

(c) Da alle verallgemeinerten Eigenräume von A in K^n liegen und somit höchstens Dimension n haben, kann in der aufsteigenden Kette dieser Räume höchstens n -mal die strikte Inklusion gelten. Zusammen mit (a) bedeutet dies, dass in jedem Fall $H(A, \lambda) = H_n(A, \lambda)$ gilt, also spätestens nach n Schritten in der Kette der Hauptraum erreicht sein muss.

Beispiel 20.4. Die reelle Matrix

$$A = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

hat nach Folgerung 18.18 das charakteristische Polynom $\chi_A(t) = t^2(t-2)$. Wir wollen zum Eigenwert $\lambda = 0$ den Hauptraum $H(A, 0)$ bestimmen. Dazu berechnen wir der Reihe nach die verallgemeinerten Eigenräume $H_r(A, 0)$:

$$\begin{aligned} H_1(A, 0) &= \text{Ker}(0E - A) = \text{Ker} \begin{pmatrix} 0 & -2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix} = \text{Lin}(e_1), \\ H_2(A, 0) &= \text{Ker}(0E - A)^2 = \text{Ker} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 4 \end{pmatrix} = \text{Lin}(e_1, e_2), \\ H_3(A, 0) &= \text{Ker}(0E - A)^3 = \text{Ker} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -8 \end{pmatrix} = \text{Lin}(e_1, e_2). \end{aligned}$$

Damit ist der gesuchte Hauptraum gleich $H(A, 0) = \text{Lin}(e_1, e_2)$: Dies folgt sowohl aus Bemerkung 20.3 (a) (da $H_2(A, 0)$ und $H_3(A, 0)$ übereinstimmen) als auch aus Bemerkung 20.3 (c) (da für 3×3 -Matrizen spätestens an der dritten Stelle der Hauptraum erreicht sein muss).

Wir sehen an diesem Beispiel also bereits, dass es mehr Hauptvektoren als Eigenvektoren geben kann, denn $\text{Eig}(A, 0) = H_1(A, 0)$ ist hier ja ein echter Unterraum von $H(A, 0) = H_2(A, 0)$. In unserem Beispiel führt dies dazu, dass die Standardbasis eine Basis von \mathbb{R}^3 aus Hauptvektoren von A ist (die ersten beiden Einheitsvektoren sind ja Hauptvektoren zum Eigenwert 0, der dritte ist offensichtlich ein Eigenvektor zum Eigenwert 2), wohingegen wir wegen $\mu_g(A, 0) = \mu_g(A, 2) = 1$ keine Basis von \mathbb{R}^3 aus Eigenvektoren von A finden können und A damit nicht diagonalisierbar ist.

Wir wollen nun zeigen, dass wir in der Tat zu jeder Matrix $A \in K^{n \times n}$ mit zerfallendem charakteristischem Polynom eine Basis aus Hauptvektoren finden können. Analog zu Abschnitt 19.C sind dies gerade die beiden Aussagen:

- (A) Wir finden in jedem Hauptraum genügend viele linear unabhängige Vektoren, d. h. es ist $\dim H(A, \lambda) = \mu_a(A, \lambda)$ für alle Eigenwerte λ .
- (B) Alle diese Vektoren zusammen genommen sind linear unabhängig, d. h. die Summe der Haupträume ist direkt.

Für den Beweis dieser Aussagen in den nächsten beiden Sätzen benötigen wir zuerst zwei kleine Bemerkungen.

Bemerkung 20.5. Es sei $A \in K^{n \times n}$.

- (a) (Kommutativität von Matrixprodukten aus A und E) Natürlich sind Matrixprodukte im Allgemeinen nicht kommutativ, d. h. in der Regel ist $AB \neq BA$ für eine weitere Matrix $B \in K^{n \times n}$. Jedoch gilt hier offensichtlich die Gleichheit, falls $B = A$ oder $B = E$ ist. Dementsprechend sind auch Produkte von Matrizen kommutativ, die aus A und E zusammengesetzt sind, wie z. B.

$$(\lambda E - A) \cdot A = A \cdot (\lambda E - A)$$

(in der Tat sind beide Seiten gleich $\lambda A - A^2$). Wir werden diese Kommutativität im Folgenden oft benutzen, ohne jedes Mal wieder darauf hinzuweisen.

- (b) (Invarianz der Haupträume) Die Abbildung $x \mapsto Ax$ lässt sich auf jeden verallgemeinerten Eigenraum $H_r(A, \lambda)$ (und damit auch auf $H(A, \lambda)$) zu einem Endomorphismus

$$H_r(A, \lambda) \rightarrow H_r(A, \lambda), x \mapsto Ax$$

einschränken: Ist nämlich $x \in H_r(A, \lambda)$, also $(\lambda E - A)^r x = 0$, so gilt nach (a) auch

$$(\lambda E - A)^r Ax = A(\lambda E - A)^r x = 0$$

und damit $Ax \in H_r(A, \lambda)$. Man sagt dafür auch, dass die verallgemeinerten Eigenräume und der Hauptraum A -invariant sind.

Satz 20.6 (Hauptraumdimension). *Es seien $A \in K^{n \times n}$ eine quadratische Matrix mit zerfallendem charakteristischen Polynom und $\lambda \in K$. Dann gilt:*

- (a) Für alle $\mu \in K$ mit $\mu \neq \lambda$ ist $H(A, \lambda) \cap H(A, \mu) = \{0\}$.
- (b) $\dim H(A, \lambda) = \mu_a(A, \lambda)$.

Beweis. Nach Bemerkung 20.5 (b) können wir die Abbildung $f: K^n \rightarrow K^n$, $x \mapsto Ax$ auf einen Endomorphismus des Hauptraums $H(A, \lambda)$ einschränken. Es gibt nach Aufgabe 17.24 (b) also lineare Abbildungen

$$g: H(A, \lambda) \rightarrow H(A, \lambda), x \mapsto Ax \quad \text{und} \quad h: K^n/H(A, \lambda) \rightarrow K^n/H(A, \lambda), \bar{x} \mapsto \overline{Ax}.$$

Für die Eigenwerte dieser beiden Abbildungen gilt:

- (1) Kein $\mu \neq \lambda$ ist ein Eigenwert von g : Ist $x \in H(A, \lambda)$ mit $Ax = \mu x$, so wissen wir einerseits $(\lambda E - A)^n x = 0$ nach Bemerkung 20.3 (c) und andererseits $(\lambda E - A)x = (\lambda - \mu)x$. Damit folgt

$$0 = (\lambda E - A)^n x = (\lambda - \mu)^n x,$$

was wegen $\mu \neq \lambda$ nur für $x = 0$ möglich ist. Also hat g keine Eigenvektoren zu μ , und damit nach Bemerkung 20.3 (b) auch keine verallgemeinerten Eigenvektoren zu μ . Mit anderen Worten ist $H(A, \lambda) \cap H(A, \mu) = H(g, \mu) = \{0\}$, was (a) zeigt.

- (2) Umgekehrt ist λ kein Eigenwert von h : Ist $\bar{x} \in K^n/H(A, \lambda)$ mit $\overline{Ax} = \lambda \bar{x}$, so folgt

$$\begin{aligned} & \overline{(\lambda E - A)x} = \bar{0} \\ \Rightarrow & (\lambda E - A)x \in H_n(A, \lambda) \quad (\text{Bemerkung 20.3 (c)}) \\ \Rightarrow & (\lambda E - A)^n (\lambda E - A)x = 0 \\ \Rightarrow & x \in H_{n+1}(A, \lambda) = H(A, \lambda), \end{aligned}$$

und damit $\bar{x} = \bar{0}$. Also besitzt h keine Eigenvektoren zu λ .

Nach Aufgabe 19.25 ist nun $\chi_A = \chi_f = \chi_g \cdot \chi_h$, und dieses Polynom zerfällt nach Voraussetzung in Linearfaktoren. Aus χ_g können nach (1) aber ausschließlich Linearfaktoren $t - \lambda$ kommen, aus χ_h dagegen nach (2) kein Linearfaktor $t - \lambda$. Die algebraische Vielfachheit $\mu_a(A, \lambda)$, also die Anzahl der Linearfaktoren $t - \lambda$ in χ_A , ist damit genau gleich dem Grad von χ_g , also der Dimension $\dim H(A, \lambda)$ des Startraums von g . Damit ist auch (b) gezeigt. \square

Satz 20.7 (Hauptraumzerlegung). *Es sei $A \in K^{n \times n}$ eine quadratische Matrix, deren charakteristisches Polynom in Linearfaktoren zerfällt. Sind $\lambda_1, \dots, \lambda_k$ die verschiedenen Eigenwerte von A , so gilt*

$$H(A, \lambda_1) \oplus \dots \oplus H(A, \lambda_k) = K^n.$$

Beweis. Wir zeigen mit Induktion über $l = 1, \dots, k$, dass die Summe $H(A, \lambda_1) + \dots + H(A, \lambda_l)$ direkt ist. Für den Induktionsanfang $l = 1$ ist dabei nichts zu zeigen. Für den Induktionsschritt $l \rightarrow l + 1$ seien

$$x_1 + \dots + x_{l+1} = y_1 + \dots + y_{l+1} \tag{1}$$

zwei Darstellungen desselben Vektors in K^n mit $x_i, y_i \in H(A, \lambda_i)$ für alle $i = 1, \dots, l + 1$. Multiplizieren wir diese Gleichung von links mit $(\lambda_{l+1} E - A)^n$, so fallen x_{l+1} und y_{l+1} nach Bemerkung 20.3 (c) weg, und wir erhalten

$$(\lambda_{l+1} E - A)^n (x_1 - y_1) + \dots + (\lambda_{l+1} E - A)^n (x_l - y_l) = 0. \tag{2}$$

Nun ist aber $(\lambda_{l+1}E - A)^n(x_i - y_i) \in H_n(A, \lambda_i) = H(A, \lambda_i)$ für alle $i = 1, \dots, l$, denn nach Bemerkung 20.5 (a) ist

$$(\lambda_i E - A)^n (\lambda_{l+1} E - A)^n (x_i - y_i) = (\lambda_{l+1} E - A)^n \underbrace{(\lambda_i E - A)^n (x_i - y_i)}_{=0} = 0.$$

Da die Summe $H(A, \lambda_1) + \dots + H(A, \lambda_l)$ nach Induktionsvoraussetzung direkt ist, ergibt sich aus (2) also

$$(\lambda_{l+1} E - A)^n (x_i - y_i) = 0 \quad \text{für alle } i = 1, \dots, l.$$

Damit ist $x_i - y_i \in H(A, \lambda_i) \cap H(A, \lambda_{l+1})$, mit Satz 20.6 (a) also $x_i - y_i = 0$ und somit $x_i = y_i$ für $i = 1, \dots, l$. Einsetzen in (1) zeigt dann auch $x_{l+1} = y_{l+1}$, und damit ist die Summe der Haupträume direkt.

Da χ_A in Linearfaktoren zerfällt, ist nun aber auch

$$\dim(H(A, \lambda_1) \oplus \dots \oplus H(A, \lambda_k)) \stackrel{17.3}{=} \sum_{i=1}^k \dim H(A, \lambda_i) \stackrel{20.6(b)}{=} \sum_{i=1}^k \mu_a(A, \lambda_i) = n,$$

und damit $H(A, \lambda_1) \oplus \dots \oplus H(A, \lambda_k) = K^n$. \square

Bemerkung 20.8. Wollen wir den Hauptraum $H(A, \lambda)$ einer Matrix $A \in K^{n \times n}$ zu einem Eigenwert λ bestimmen, indem wir nacheinander die verallgemeinerten Eigenräume $H_r(A, \lambda)$ mit $r \in \mathbb{N}$ berechnen, so folgt aus Satz 20.6 (b) insbesondere, dass wir mit der Berechnung dieser Räume aufhören können, sobald einer von ihnen die Dimension $\mu_a(A, \lambda)$ hat – und nicht erst wie in Beispiel 20.4, wenn zwei von ihnen gleich sind oder $r = n$ gilt. In der Tat hätten wir in diesem Beispiel damit $H_3(A, 0)$ nicht mehr berechnen müssen, da dort bereits $\dim H_2(A, 0) = \mu_a(A, 0) = 2$ war und somit $H(A, 0) = H_2(A, 0)$ gelten musste.

46

20.B Jordandiagramme

Wir wollen nun endlich unser Problem lösen, zu einer quadratischen Matrix mit zerfallendem charakteristischem Polynom eine möglichst einfache ähnliche Matrix zu finden. Dazu haben wir im letzten Abschnitt gesehen, dass sich diese Frage auf die Untersuchung der einzelnen Haupträume reduzieren lässt. Wir fassen unser Ergebnis hier noch einmal in der jetzt benötigten Form zusammen.

Bemerkung 20.9 (Hauptraumzerlegung). Es sei $A \in K^{n \times n}$ eine Matrix, deren charakteristisches Polynom in Linearfaktoren zerfällt. Sind $\lambda_1, \dots, \lambda_k$ die verschiedenen Eigenwerte von A , so haben wir in Satz 20.7 gesehen, dass

$$H(A, \lambda_1) \oplus \dots \oplus H(A, \lambda_k) = K^n.$$

Wählen wir also Basen B_1, \dots, B_k dieser Haupträume, so erhalten wir nach Aufgabe 17.7 mit allen diesen Vektoren zusammen genommen eine Basis B von K^n . Damit können wir diese Vektoren in die Spalten einer invertierbaren Matrix T schreiben.

Wegen der Invarianz der Haupträume gemäß Bemerkung 20.5 (b) können wir nun die Abbildung $f: K^n \rightarrow K^n$, $x \mapsto Ax$ auf die Haupträume einschränken zu

$$f_i: H(A, \lambda_i) \rightarrow H(A, \lambda_i), \quad x \mapsto Ax$$

für $i = 1, \dots, k$. Die Matrix A_f^B (die nach Bemerkung 19.3 (b) gleich $T^{-1}AT$ ist) kann nach Satz 16.27 in den zu B_i gehörigen Spalten also auch nur in den entsprechenden Zeilen Einträge ungleich 0 haben, und ist dort gleich der Abbildungsmatrix $A_{f_i}^{B_i}$. Wir erhalten für diese Matrix damit

$$A_f^B = T^{-1}AT = \begin{pmatrix} \boxed{A_{f_1}^{B_1}} & & 0 \\ & \ddots & \\ 0 & & \boxed{A_{f_k}^{B_k}} \end{pmatrix}$$

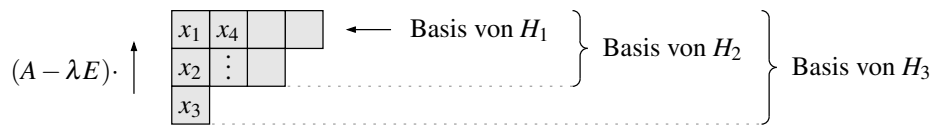
Man sagt, dass A_f^B eine *Blockdiagonalmatrix* mit den Blöcken $A_{f_1}^{B_1}, \dots, A_{f_k}^{B_k}$ ist.

Um unsere Suche nach einer möglichst einfachen zu A ähnlichen Matrix zu beenden, müssen wir jetzt also nur noch die Basen B_i der einzelnen Haupträume $H(A, \lambda_i)$ für $i = 1, \dots, k$ so wählen, dass die Abbildungsmatrizen $A_{f_i}^{B_i}$ der auf die Haupträume eingeschränkten Abbildung möglichst einfach werden. Hierfür ist das folgende Konzept der Jordandiagramme sehr nützlich, da es die etwas komplizierte Berechnung von B_i auf grafische Art gut veranschaulicht.

Definition 20.10 (Jordandiagramme). Es sei $A \in K^{n \times n}$ eine quadratische Matrix, deren charakteristisches Polynom in Linearfaktoren zerfällt. Wir wählen einen festen Eigenwert λ von A und setzen zur Abkürzung der Schreibweise $H_r := H_r(A, \lambda)$ für alle $r \in \mathbb{N}$.

Ein **Jordandiagramm** von A zum Eigenwert λ ist dann wie im Bild unten dargestellt ein in linksbündigen Zeilen angeordnetes Diagramm von endlich vielen Kästchen, mit von oben nach unten (nicht notwendig streng) monoton fallenden Zeilenlängen, und zusammen mit einem Vektor in K^n in jedem dieser Kästchen, so dass die folgenden beiden Bedingungen erfüllt sind:

- (a) (*Zeilenbedingung*) Für alle $r > 0$ bilden die Vektoren der ersten r Zeilen eine Basis von H_r . Insbesondere bilden also die Vektoren in allen Kästchen zusammen eine Basis des Hauptraums $H(A, \lambda)$.
- (b) (*Spaltenbedingung*) Steht ein Vektor y in einer Spalte des Diagramms unmittelbar über dem Vektor x , so ist $y = (A - \lambda E)x$ (im Diagramm unten ist also z. B. $x_2 = (A - \lambda E)x_3$).



Bemerkung 20.11 (Alternative Zeilenbedingung). Oft ist es nützlich, die Zeilenbedingung aus Definition 20.10 (a) äquivalent umzuformulieren zu:

(*Alternative Zeilenbedingung*) Für alle $r > 0$ liegen die Vektoren der Zeile r in H_r , und ihre Klassen bilden eine Basis von H_r/H_{r-1} .

Diese alternative Bedingung ist nämlich nach Bemerkung 17.21 äquivalent dazu, dass die Vektoren der Zeile r eine Basis von H_{r-1} zu einer Basis von H_r ergänzen – und dies ist ja gerade die ursprüngliche Zeilenbedingung.

Insbesondere heißt dies auch, dass alle Vektoren einer Zeile r in $H_r \setminus H_{r-1}$ liegen, also Hauptvektoren der Stufe r sind.

Wir werden in Satz 20.13 noch beweisen, dass Jordandiagramme immer existieren und auch einfach berechenbar sind. Zuerst wollen wir aber sehen, warum gerade eine Basiswahl wie in den Kästchen eines Jordandiagramms zu einer besonders einfachen Abbildungsmatrix führt.

Konstruktion 20.12 (Abbildungsmatrizen aus Jordandiagrammen). Wir betrachten zunächst eine Spalte der Länge m in einem Jordandiagramm zu einer Matrix $A \in K^{n \times n}$ zum Eigenwert λ ; die Vektoren von oben nach unten gelesen seien x_1, \dots, x_m (im Bild rechts ist eine solche Situation für den Fall $m = 4$ dargestellt).

$$\begin{matrix} (A - \lambda E) \cdot \left(\begin{matrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{matrix} \right) \\ (A - \lambda E) \cdot \left(\begin{matrix} x_2 \\ x_3 \\ x_4 \end{matrix} \right) \\ (A - \lambda E) \cdot \left(\begin{matrix} x_3 \\ x_4 \end{matrix} \right) \end{matrix}$$

Wegen $x_1 \in H_1$ ist dann $Ax_1 = \lambda x_1$, während für alle $i = 2, \dots, m$ nach der Spaltenbedingung (b) aus Definition 20.10

$$(A - \lambda E)x_i = x_{i-1}, \quad \text{also} \quad Ax_i = \lambda x_i + x_{i-1}$$

gilt. Der Unterraum $U = \text{Lin}(x_1, \dots, x_m)$ ist damit im Sinne von Bemerkung 20.5 (b) A -invariant, d. h. die Abbildung $K^n \rightarrow K^n$, $x \mapsto Ax$ lässt sich zu einer Abbildung $U \rightarrow U$, $x \mapsto Ax$ einschränken, und die Abbildungsmatrix dieser eingeschränkten Abbildung ist bezüglich der Basis (x_1, \dots, x_m) von

U nach Satz 16.27 gleich

$$J_m(\lambda) := \begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix} \in K^{m \times m}.$$

Wir nennen diese $m \times m$ -Matrix $J_m(\lambda)$, bei der alle Einträge auf der Diagonale gleich λ , die unmittelbar darüber gleich 1, und alle anderen gleich 0 sind, einen **Jordanblock** (oder manchmal auch ein **Jordankästchen**) der Größe m zum Eigenwert λ .

Wir gehen nun zum gesamten Jordandiagramm zu λ über, nehmen also alle Spalten zusammen. Wählen wir als Basis von $H(A, \lambda)$ die Vektoren im Diagramm, spaltenweise von oben nach unten gelesen, so erhalten wir für die Abbildungsmatrix zu $x \mapsto Ax$ auf $H(A, \lambda)$ analog zu Bemerkung 20.9 eine Blockdiagonalform mit Jordanblöcken wie oben beschrieben, also

$$\begin{pmatrix} \boxed{J_{m_1}(\lambda)} & & 0 \\ & \ddots & \\ 0 & & \boxed{J_{m_l}(\lambda)} \end{pmatrix},$$

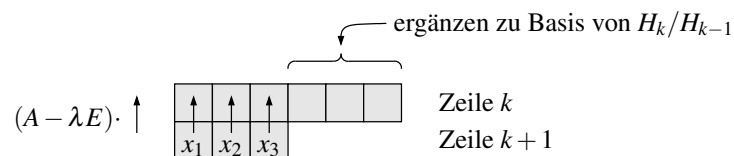
wobei jeder Jordanblock einer Spalte im Jordandiagramm entspricht und die Größen m_1, \dots, m_l der Jordanblöcke genau die Längen dieser Spalten (in der gewählten Reihenfolge) sind. Dies ist die einfache Form der Abbildungsmatrix, nach der wir gesucht haben.

Wir haben nun also gesehen, dass man mit Hilfe von Jordandiagrammen Haupttraumbasen bestimmen kann, die letztlich zu sehr einfachen Abbildungsmatrizen führen. Im nächsten Satz wollen wir daher zeigen, dass solche Jordandiagramme auch wirklich existieren. Der Beweis des Satzes gibt gleichzeitig auch ein einfaches konstruktives Verfahren zur Berechnung eines solchen Diagramms.

Satz 20.13 (Existenz und Berechnung von Jordandiagrammen). *Es seien $A \in K^{n \times n}$ eine quadratische Matrix, deren charakteristisches Polynom in Linearfaktoren zerfällt, und λ ein Eigenwert von A . Dann gibt es ein Jordandiagramm von A zum Eigenwert λ .*

Beweis. Wie oben setzen wir wieder $H_k := H_k(A, \lambda)$ für alle $k \in \mathbb{N}$. Es sei $r \in \mathbb{N}$ die kleinste Zahl mit $H_r(A, \lambda) = H(A, \lambda)$, also die benötigte Anzahl Zeilen im Jordandiagramm. Wir konstruieren das gesuchte Diagramm nun nach folgendem Verfahren zeilenweise von unten nach oben:

- In die letzte Zeile r schreiben wir beliebige Vektoren in H_r , deren Klassen wie in Bemerkung 20.11 eine Basis von H_r/H_{r-1} bilden – z. B. indem wir wie in Bemerkung 17.21 eine Basis von H_{r-1} zu einer Basis von H_r ergänzen und die hinzugenommenen Vektoren in die letzte Zeile des Diagramms schreiben. (Die dafür bestimmte Basis von H_{r-1} wird dabei im Folgenden nicht mehr benötigt.)
- Um für $k = r - 1, \dots, 1$ die Zeile k aus Zeile $k + 1$ zu konstruieren, schreiben wir wie im Bild unten dargestellt zunächst unmittelbar über die Vektoren x_1, \dots, x_m der Zeile $k + 1$ die Vektoren $(A - \lambda E)x_1, \dots, (A - \lambda E)x_m$ in Zeile k , und ergänzen diese Vektoren dann so, dass die Klassen der Vektoren in Zeile k insgesamt eine Basis von H_k/H_{k-1} bilden.



Nach Konstruktion sind im ausgefüllten Diagramm dann überall die alternative Zeilenbedingung aus Bemerkung 20.11 und die Spaltenbedingung aus Definition 20.10 (b) erfüllt. Außerdem müssen die Zeilenlängen dann natürlich von oben nach unten monoton fallend sein. Um sicherzustellen, dass Schritt (b) immer funktioniert (wenn das Diagramm unterhalb dieser Zeile schon korrekt ausgefüllt ist), müssen wir aber noch zwei Dinge überprüfen:

- Die Vektoren $(A - \lambda E)x_i$ liegen wirklich in H_k für alle $i = 1, \dots, m$: Da x_i in Zeile $k + 1$ steht, ist $x_i \in H_{k+1}$, also $(\lambda E - A)^{k+1}x_i = 0$ bzw. $(A - \lambda E)^{k+1}x_i = 0$. Dies bedeutet aber auch $(A - \lambda E)^k(A - \lambda E)x_i = 0$, und damit $(A - \lambda E)x_i \in H_k$.
- Die Klassen der Vektoren $(A - \lambda E)x_1, \dots, (A - \lambda E)x_m$ sind linear unabhängig in H_k/H_{k-1} , so dass wir sie zu einer Basis dieses Raumes ergänzen können: Für $\mu_1, \dots, \mu_m \in K$ gilt

$$\begin{aligned} & \mu_1 \overline{(A - \lambda E)x_1} + \dots + \mu_m \overline{(A - \lambda E)x_m} = \bar{0} \in H_k/H_{k-1} \\ \Rightarrow & \mu_1 (A - \lambda E)x_1 + \dots + \mu_m (A - \lambda E)x_m \in H_{k-1} && \text{(Bemerkung 17.16)} \\ \Rightarrow & (A - \lambda E)^{k-1}(A - \lambda E)(\mu_1 x_1 + \dots + \mu_m x_m) = 0 && \text{(Definition von } H_{k-1}) \\ \Rightarrow & \mu_1 x_1 + \dots + \mu_m x_m \in H_k && \text{(Definition von } H_k) \\ \Rightarrow & \mu_1 \bar{x}_1 + \dots + \mu_m \bar{x}_m = \bar{0} \in H_{k+1}/H_k && \text{(Bemerkung 17.16)} \\ \Rightarrow & \mu_1 = \dots = \mu_m = 0, \end{aligned}$$

da die Vektoren $\bar{x}_1, \dots, \bar{x}_m$ eine Basis von H_{k+1}/H_k bilden und damit in diesem Quotientenraum linear unabhängig sind. □

Ein konkretes Beispiel für dieses Verfahren werden wir in Beispiel 20.17 noch angeben. Zunächst aber wollen wir unsere Ergebnisse zusammenfassen und damit das Hauptergebnis dieses Kapitels zeigen.

Folgerung 20.14 (Jordansche Normalform). *Es sei $A \in K^{n \times n}$ eine quadratische Matrix, deren charakteristisches Polynom in Linearfaktoren zerfällt. Dann ist A ähnlich zu einer Blockdiagonalmatrix der Form*

$$J = \begin{pmatrix} \boxed{J_{m_1}(\lambda_1)} & & 0 \\ & \ddots & \\ 0 & & \boxed{J_{m_k}(\lambda_k)} \end{pmatrix},$$

wobei die $J_{m_1}(\lambda_1), \dots, J_{m_k}(\lambda_k)$ Jordanblöcke wie in Konstruktion 20.12 für gewisse $m_1, \dots, m_k \in \mathbb{N}_{>0}$ und nicht notwendig verschiedene Eigenwerte $\lambda_1, \dots, \lambda_k$ von A sind, d. h. für $i = 1, \dots, k$ ist

$$J_{m_i}(\lambda_i) = \begin{pmatrix} \lambda_i & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{pmatrix} \in K^{m_i \times m_i}.$$

Man nennt eine solche Matrix J die **Jordanform** oder **Jordansche Normalform** von A (wir werden in Aufgabe 20.21 noch sehen, dass sie bis auf die Reihenfolge der Blöcke eindeutig bestimmt ist).

Ist $T = (x_1 | \dots | x_n) \in GL(n, K)$ eine Matrix mit $J = T^{-1}AT$, so dass J also die Abbildungsmatrix von $K^n \rightarrow K^n$, $x \mapsto Ax$ bezüglich der Basis $B = (x_1, \dots, x_n)$ ist, so heißt B eine **Jordanbasis** von A .

Beweis. Auch der Beweis dieser Folgerung ist konstruktiv: Man bestimme zunächst mit Hilfe des charakteristischen Polynoms χ_A die Eigenwerte von A , und dann zu jedem dieser Eigenwerte ein Jordandiagramm mit dem Verfahren aus Satz 20.13. Aus jedem Diagramm erhalten wir eine Basis des entsprechenden Hauptraums, und so nach der Hauptraumzerlegung aus Satz 20.7 insgesamt eine Basis von K^n . Dies ist dann eine Jordanbasis: Lesen wir die Basisvektoren in den Diagrammen spaltenweise von oben nach unten und schreiben sie in die Matrix T , so ist $T^{-1}AT$ nach Bemerkung 20.9 eine Blockdiagonalmatrix mit einem Block für jeden Eigenwert λ , wobei nach Konstruktion 20.12 jeder dieser Blöcke selbst wieder eine Blockdiagonalfom aus Jordanblöcken zu λ ist. □

Bemerkung 20.15 (Bestimmung der Jordanform ohne Jordanbasis). Möchte man zu einer Matrix A (mit in Linearfaktoren zerfallendem charakteristischem Polynom) nur die Jordanform, aber keine Jordanbasis bestimmen, so ist das Verfahren hierfür sehr viel einfacher. Weil wir genau einen Jordanblock der Größe m zum Eigenwert λ für jede Spalte der Länge m im Jordandiagramm zu λ bekommen, benötigen wir zur Bestimmung der Jordanform nämlich nur die äußere Form der Jordandiagramme, aber nicht die in ihnen stehenden Vektoren. Da die äußere Form eines solchen

Diagramms zum Eigenwert λ nach Bemerkung 20.11 aber wiederum dadurch eindeutig bestimmt ist, dass Zeile r für alle r die Länge $\dim H_r(A, \lambda) - \dim H_{r-1}(A, \lambda)$ hat, genügt also die Information der Dimensionen aller Räume $H_r(A, \lambda)$, um die Jordanform zu bestimmen.







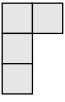

In vielen Fällen reicht sogar noch viel weniger Information zur Bestimmung der Jordanform aus. So haben z. B. die algebraische und geometrische Vielfachheit eines Eigenwerts eine direkte Interpretation in den Jordandiagrammen und damit auch in der Jordanform:

- (a) Die algebraische Vielfachheit $\mu_a(A, \lambda)$ ist nach Satz 20.6 (b) die Dimension von $H(A, \lambda)$, also gleich der Anzahl der Kästchen im Jordandiagramm zu λ . Natürlich ist dies dann gleichzeitig auch die Summe der Längen aller Spalten in diesem Diagramm, und damit gleich der Summe der Größen aller Jordanblöcke in der Jordanform von A .
- (b) Die geometrische Vielfachheit $\mu_g(A, \lambda) = \dim \text{Eig}(A, \lambda) = \dim H_1(A, \lambda)$ ist nach der Zeilenbedingung aus Definition 20.10 (a) genau die Länge der ersten Zeile des Jordandiagramms zu λ . Da die Zeilenlängen in einem Jordandiagramm von oben nach unten monoton fallend sind, ist dies dasselbe wie die Anzahl der Spalten im Jordandiagramm, und damit wie die Anzahl der Jordanblöcke zum Eigenwert λ in der Jordanform von A .

Wir halten also fest:

$$\begin{aligned} \mu_a(A, \lambda) &= \text{Anzahl der Kästchen im Jordandiagramm zum Eigenwert } \lambda \\ &= \text{Summe der Größen der Jordanblöcke zum Eigenwert } \lambda \text{ in der Jordanform} \\ \mu_g(A, \lambda) &= \text{Anzahl der Spalten im Jordandiagramm zum Eigenwert } \lambda \\ &= \text{Anzahl der Jordanblöcke zum Eigenwert } \lambda \text{ in der Jordanform} \end{aligned}$$

Im Fall $\mu_a(A, \lambda) \leq 3$ reichen diese beiden Zahlen in der Tat bereits aus, um das Jordandiagramm und damit auch die von diesem Eigenwert kommenden Jordanblöcke in der Jordanform eindeutig zu bestimmen: Im Bild unten links sind alle möglichen Jordandiagramme mit höchstens 3 Kästchen (und die sich daraus ergebenden Jordanblöcke) angegeben, und diese unterscheiden sich alle in der Anzahl ihrer Kästchen oder Spalten, also durch die algebraische oder geometrische Vielfachheit des Eigenwerts. Erst im Fall $\mu_a(A, \lambda) = 4$ gibt es zum ersten Mal zwei Jordandiagramme, die sich durch diese beiden Zahlen nicht unterscheiden lassen, nämlich die beiden Diagramme unten rechts mit 4 Kästchen und 2 Spalten – und dementsprechend auch zwei verschiedene mögliche Jordanformen. Möchte man zwischen diesen beiden Fällen unterscheiden, muss man auch noch $\dim H_2(A, \lambda)$ berechnen: Diese Zahl ist nach Definition 20.10 ja gleich der Anzahl der Kästchen in den ersten beiden Zeilen des Jordandiagramms, in den beiden Fällen im Bild unten rechts also 4 bzw. 3.

$\mu_a(A, \lambda):$	1	2	2	3	3	3		4	4
$\mu_g(A, \lambda):$	1	1	2	1	2	3		2	2
Jordan- diagramm									
Jordan- blöcke	$\begin{bmatrix} \lambda \end{bmatrix}$	$\begin{bmatrix} \lambda & 1 \\ & \lambda \end{bmatrix}$	$\begin{bmatrix} \lambda & & \\ & \lambda & \end{bmatrix}$	$\begin{bmatrix} \lambda & 1 & \\ & \lambda & 1 \\ & & \lambda \end{bmatrix}$	$\begin{bmatrix} \lambda & 1 & \\ & \lambda & \\ & & \lambda \end{bmatrix}$	$\begin{bmatrix} \lambda & & \\ & \lambda & \\ & & \lambda \end{bmatrix}$		$\begin{bmatrix} \lambda & 1 & \\ & \lambda & 1 \\ & & \lambda \end{bmatrix}$	$\begin{bmatrix} \lambda & 1 & \\ & \lambda & \\ & & \lambda & 1 \\ & & & \lambda \end{bmatrix}$

Bemerkung 20.16.

- (a) Natürlich gibt es einen zu Folgerung 20.14 analogen Satz auch wieder für Endomorphismen $f: V \rightarrow V$ eines endlich erzeugten Vektorraums V : Zerfällt χ_f in Linearfaktoren, so gibt es eine Jordanbasis B von V für f – also eine Basis, so dass die zugehörige Abbildungsmatrix A_f^B in Jordanform ist.

- (b) Beachte, dass der Fall von diagonalisierbaren Matrizen, also Folgerung 19.40, im Satz über die Jordansche Normalform enthalten ist: Ist A eine quadratische Matrix, deren charakteristisches Polynom in Linearfaktoren zerfällt, so ergibt sich aus Folgerung 20.14 und Bemerkung 20.15

A diagonalisierbar \Leftrightarrow alle Jordanblöcke haben die Größe 1
 \Leftrightarrow alle Spalten der Jordandiagramme haben die Länge 1
 \Leftrightarrow alle Jordandiagramme haben genauso viele Spalten wie Kästchen
 $\Leftrightarrow \mu_g(A, \lambda) = \mu_a(A, \lambda)$ für alle Eigenwerte λ ,

und damit Folgerung 19.40.

47

Beispiel 20.17. Es sei

$$A = \begin{pmatrix} 0 & 1 & 1 & 0 \\ -4 & 4 & 2 & 0 \\ 0 & 0 & 2 & 0 \\ 1 & 0 & 1 & 3 \end{pmatrix} \in \mathbb{R}^{4 \times 4}.$$

Das charakteristische Polynom von A berechnet sich wohl am einfachsten durch Laplace-Entwicklung gemäß Satz 18.15, zunächst nach der letzten Spalte und dann nach der letzten Zeile:

$$\begin{aligned} \chi_A(t) &= \det \begin{pmatrix} t & -1 & -1 & 0 \\ 4 & t-4 & -2 & 0 \\ 0 & 0 & t-2 & 0 \\ -1 & 0 & -1 & t-3 \end{pmatrix} = (t-3) \cdot \det \begin{pmatrix} t & -1 & -1 \\ 4 & t-4 & -2 \\ 0 & 0 & t-2 \end{pmatrix} \\ &= (t-3)(t-2) \cdot \det \begin{pmatrix} t & -1 \\ 4 & t-4 \end{pmatrix} = (t-3)(t-2)(t^2 - 4t + 4) = (t-3)(t-2)^3. \end{aligned}$$

Da χ_A in Linearfaktoren zerfällt, besitzt A nach Folgerung 20.14 also eine Jordanform.

Wollen wir diese berechnen, müssen wir die Jordandiagramme zu den beiden Eigenwerten $\lambda_1 = 2$ und $\lambda_2 = 3$ bestimmen. Für $\lambda_2 = 3$ ist dies einfach: Aus $\mu_a(A, 3) = 1$ folgt natürlich bereits, dass das zugehörige Jordandiagramm aus nur einem Kästchen besteht und die Jordanform von A daher genau einen Jordanblock der Größe 1 zum Eigenwert 3 hat. Für den Eigenwert $\lambda_1 = 2$ hingegen ist $\mu_a(A, 2) = 3$, und daher gibt es für das Jordandiagramm noch die drei im Bild von Bemerkung 20.15 aufgelisteten Möglichkeiten, die drei Kästchen im Jordandiagramm anzuordnen. Um zu entscheiden, welcher dieser Fälle hier vorliegt, berechnen wir die geometrische Vielfachheit $\mu_g(A, 2)$: Mit dem Gauß-Algorithmus ergibt sich

$$H_1(A, 2) = \text{Ker}(2E - A) = \text{Ker} \begin{pmatrix} 2 & -1 & -1 & 0 \\ 4 & -2 & -2 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & -1 & -1 \end{pmatrix} = \text{Lin} \left(\underbrace{\begin{pmatrix} 1 \\ 2 \\ 0 \\ -1 \end{pmatrix}}_{=: x_1}, \underbrace{\begin{pmatrix} 1 \\ 3 \\ -1 \\ 0 \end{pmatrix}}_{=: x_2} \right), \quad (*)$$

und damit $\mu_g(A, 2) = 2$. Das Jordandiagramm zum Eigenwert 2 hat gemäß Bemerkung 20.15 also 3 Kästchen und 2 Spalten und muss damit wie im Bild am Ende dieses Beispiels in der ersten Spalte zwei und in der zweiten Spalte ein Kästchen haben. Zum Eigenwert 3 gibt es also zwei Jordanblöcke, deren Größen gerade diese Spaltenlängen 2 bzw. 1 sind. Die Jordanform von A ist damit

$$J = \left(\begin{array}{cc|cc} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ \hline 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{array} \right).$$

Wollen wir auch noch eine Jordanbasis bestimmen, müssen wir die beiden Jordandiagramme zu den Eigenwerten 2 und 3 noch mit Vektoren füllen. Für den Eigenwert $\lambda_2 = 3$ ist dies wieder einfach: Da der vierte Einheitsvektor offensichtlich ein Eigenvektor zu diesem Eigenwert ist, können wir ihn wie im Bild unten in das eine Kästchen des Jordandiagramms schreiben. Um auch das Jordandiagramm

für $\lambda_1 = 2$ zu füllen, verwenden wir das Verfahren aus Satz 20.13 und beginnen also in der zweiten Zeile des Diagramms, wofür wir eine Basis für den eindimensionalen Raum $H_2(A, 2)/H_1(A, 2)$ brauchen. Wir müssen also erst einmal den verallgemeinerten Eigenraum $H_2(A, 2)$ bestimmen. Dies können wir durch explizite Berechnung von $H_2(A, 2) = \text{Ker}(2E - A)^2$ mit dem Gauß-Algorithmus machen und erhalten

$$H_2(A, 2) = \text{Ker}(2E - A)^2 = \text{Ker} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 1 & 2 & 1 \end{pmatrix} = \text{Lin} \left(\underbrace{\begin{pmatrix} 1 \\ 2 \\ 0 \\ -1 \end{pmatrix}}_{=: y_1}, \underbrace{\begin{pmatrix} 1 \\ 3 \\ -1 \\ 0 \end{pmatrix}}_{=: y_2}, \underbrace{\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}}_{=: y_3} \right)$$

Offensichtlich ist y_3 ein Vektor, der eine Basis von $H_1(A, 2)$ zu einer von $H_2(A, 2)$ ergänzt. Wir schreiben ihn also in das untere Kästchen des Jordandiagramms. Direkt darüber gehört nun nach dem Verfahren aus Satz 20.13 der Vektor

$$(A - 2E) \cdot y_3 = \begin{pmatrix} -2 & 1 & 1 & 0 \\ -4 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ -2 \\ 0 \\ 1 \end{pmatrix}.$$

Beachte, dass dieser Vektor nach (*) tatsächlich in $H_1(A, 2)$ liegt (dies ist eine gute Kontrolle der Rechnung – wir hätten irgendwo einen Rechenfehler gemacht, wenn dies nicht so wäre). Für das rechte Kästchen des Jordandiagramms müssen wir ihn noch zu einer Basis von $H_1(A, 2)$ ergänzen, nach (*) z. B. mit x_2 .

Insgesamt haben wir also die folgenden Jordandiagramme erhalten:

$$\begin{array}{|c|c|} \hline \begin{pmatrix} -1 \\ -2 \\ 0 \\ 1 \end{pmatrix} & \begin{pmatrix} 1 \\ 3 \\ -1 \\ 0 \end{pmatrix} \\ \hline \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} & \\ \hline \end{array} \quad \lambda_1 = 2 \qquad \begin{array}{|c|} \hline \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \\ \hline \end{array} \quad \lambda_2 = 3$$

Für eine Jordanbasis müssen wir die Vektoren in diesen Kästchen nun nur noch spaltenweise von oben nach unten lesen bzw. sie in dieser Reihenfolge als Spalten in die Transformationsmatrix T schreiben. Dabei müssen wir die Spalten der Diagramme in der gleichen Reihenfolge durchgehen, in der wir oben die Jordanblöcke in J angeordnet haben:

$$T := \begin{pmatrix} -1 & 1 & 1 & 0 \\ -2 & 1 & 3 & 0 \\ 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}.$$

Damit gilt dann $T^{-1}AT = J$ nach Folgerung 20.14 (was wir durch direkte Berechnung von T^{-1} und $T^{-1}AT$ natürlich auch explizit überprüfen könnten).

Aufgabe 20.18. Berechne die Jordanschen Normalformen der reellen Matrizen

$$A = \begin{pmatrix} -1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 2 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & -1 & -1 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 \\ 0 & 1 & 3 & 0 & 0 \\ 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 1 & 4 \end{pmatrix}.$$

Für die Matrix A bestimme man dabei zusätzlich eine Jordanbasis; für die Matrix B hingegen versuche man, mit möglichst wenig Rechenaufwand lediglich die Jordanform zu ermitteln.

Aufgabe 20.19. Für gegebenes $n \in \mathbb{N}$ sei $V = \text{Pol}_n(\mathbb{R}, \mathbb{R})$ der Vektorraum aller reellen Polynome vom Grad höchstens n . Berechne die Jordanform der linearen Abbildung $f: V \rightarrow V$ mit $f(\varphi)(x) = \varphi(x+1)$.

Aufgabe 20.20 (Anwendung der Jordanform auf Systeme von Differentialgleichungen). In dieser Aufgabe wollen wir ein vermutlich recht unerwartetes Beispiel für die Anwendung der Jordanform aus dem Bereich der Analysis geben. Das Ziel soll es sein, reelle differenzierbare Funktionen $f_1, f_2, f_3: \mathbb{R} \rightarrow \mathbb{R}$ zu bestimmen, so dass $f_1(0) = f_2(0) = f_3(0) = 1$ und

$$\begin{aligned}f_1' &= f_2 + 2f_3 \\f_2' &= f_1 + f_2 + 3f_3 \\f_3' &= -f_1 - f_3\end{aligned}$$

(an jeder Stelle $x \in \mathbb{R}$) gilt, wobei f_i' wie üblich die Ableitung von f_i bezeichnet.

Derartige Systeme von sogenannten *Differentialgleichungen*, die also in jedem Punkt die Änderung von Funktionen durch die Funktionswerte selbst ausdrücken, kommen in der Praxis überall vor. Ausgehend von einem Anfangswert (hier bei $x = 0$) kann man mit ihnen die gesamten Funktionen f_i rekonstruieren.

Zur Lösung schreibe man die gegebenen Gleichungen in Matrixform $f' = A \cdot f$ mit $A \in \mathbb{R}^{3 \times 3}$ und bringe A in Jordanform, d. h. bestimme eine Matrix $T \in \text{GL}(3, \mathbb{R})$, so dass $T^{-1}AT = J$ eine Jordanmatrix ist. Wenn ihr die Gleichungen dann umschreibt in Gleichungen für $g = T^{-1}f$ (also $f = Tg$), sollte sich dieses neue Differentialgleichungssystem für g (mit den passenden Werten bei $x = 0$) leicht lösen lassen.

Aufgabe 20.21 (Eindeutigkeit der Jordanform). Wir wollen nun die bereits in Folgerung 20.14 behauptete Eindeutigkeit der Jordanform (bis auf die Reihenfolge der Jordanblöcke) beweisen. Man zeige dazu:

- Sind A und B zwei ähnliche Matrizen, so gilt $\dim H_r(A, \lambda) = \dim H_r(B, \lambda)$ für alle $r \in \mathbb{N}$ und $\lambda \in K$.
- Ist A eine Matrix in Jordanscher Normalform, λ ein Eigenwert von A und $k \in \mathbb{N}_{>0}$, so ist die Anzahl der Jordanblöcke der Größe r zum Eigenwert λ in A genau

$$2 \dim H_r(A, \lambda) - \dim H_{r-1}(A, \lambda) - \dim H_{r+1}(A, \lambda).$$

- Zwei Matrizen in Jordanscher Normalform sind genau dann ähnlich zueinander, wenn sie aus den gleichen Jordanblöcken, nur evtl. in anderer Reihenfolge bestehen.

Zum Abschluss dieses Abschnitts wollen wir schließlich noch ein paar Aufgaben betrachten, in denen man die Nützlichkeit der einfachen Matrixdarstellung der Jordanform in theoretischen Problemen sieht. Ihnen allen ist gemeinsam, dass man die zu zeigende Aussage für Jordanblöcke bzw. Matrizen in Jordanform recht einfach sehen und sie dann mit Hilfe einer Ähnlichkeitstransformation problemlos auf beliebige Matrizen (mit zerfallendem charakteristischem Polynom) übertragen kann – während ein direkter Beweis für beliebige Matrizen deutlich schwieriger wäre.

Aufgabe 20.22. Beweise, dass jede quadratische komplexe Matrix $A \in \mathbb{C}^{n \times n}$ zu ihrer transponierten Matrix A^T ähnlich ist.

Aufgabe 20.23. Es sei $A \in \mathbb{C}^{n \times n}$. Zeige, dass $\dim\{B \in \mathbb{C}^{n \times n} : AB = BA\} \geq n$.

Aufgabe 20.24. Eine quadratische Matrix $N \in K^{n \times n}$ heißt *nilpotent*, wenn es ein $k \in \mathbb{N}$ gibt mit $N^k = 0$.

- Zeige, dass sich jede komplexe Matrix $A \in \mathbb{C}^{n \times n}$ als Summe $A = D + N$ schreiben lässt, wobei D diagonalisierbar und N nilpotent ist, sowie $DN = ND$ gilt.

- (b) Berechne mit Hilfe von (a) für alle
- $n \in \mathbb{N}$
- die Potenzen

$$\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}^n \quad \text{und} \quad \begin{pmatrix} 4 & -4 \\ 1 & 0 \end{pmatrix}^n$$

in $\mathbb{R}^{2 \times 2}$ durch eine direkte Rechnung (also ohne erst durch Ausprobieren eine Formel zu raten und sie danach durch Induktion zu beweisen).

- (c) Wir definieren eine reelle Folge
- $(a_n)_{n \in \mathbb{N}}$
- rekursiv durch

$$a_0 = a_1 = 1 \quad \text{und} \quad a_{n+2} = 4a_{n+1} - 4a_n \quad \text{für alle } n \in \mathbb{N}.$$

Bestimme eine explizite Formel für alle a_n .

(Hinweis: Die Rekursionsgleichung ist offensichtlich äquivalent zu

$$\begin{pmatrix} a_{n+2} \\ a_{n+1} \end{pmatrix} = \begin{pmatrix} 4 & -4 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix}$$

für alle $n \in \mathbb{N}$.)

20.C Minimalpolynome

Als eine interessante Anwendung der Jordanschen Normalform wollen wir in diesem Abschnitt Polynomausdrücke in Matrizen betrachten. Hier sind zunächst einmal zwei Beispiele dafür.

Beispiel 20.25.

- (a) Wenn wir im Polynom $p(t) = (t - \lambda)^r$ „für t eine quadratische Matrix A einsetzen“, erhalten wir $(A - \lambda E)^r$. Diese Matrix haben wir in diesem Kapitel bereits oft betrachtet, denn ihr Kern ist ja gerade der verallgemeinerte Eigenraum $H_r(A, \lambda)$.
- (b) Es sei A eine quadratische Matrix mit $A^2 = A$, also eine „Nullstelle des Polynoms $t^2 - t$ “. Wir können uns fragen, was wir dann über A aussagen können, also ob wir die quadratische Gleichung $t^2 - t = 0$ „im Matrizenraum lösen“ können. In der Tat werden wir die allgemeine Matrixlösung dieser Gleichung in Beispiel ?? angeben können. Wir können hier aber schon einmal am Beispiel der reellen Matrix

$$A = \begin{pmatrix} 2 & 1 \\ -2 & -1 \end{pmatrix} \quad \text{mit} \quad A^2 - A = \begin{pmatrix} 2 & 1 \\ -2 & -1 \end{pmatrix} - \begin{pmatrix} 2 & 1 \\ -2 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

sehen, dass wir (im Gegensatz zur Lösung der Gleichung $t^2 - t = t(t - 1) = 0$ in K) hier nicht nur die „offensichtlichen“ Lösungen $A = 0$ und $A = E$ erhalten werden.

Als Erstes sollten wir aber das Einsetzen einer Matrix in ein Polynom exakt definieren.

Definition 20.26 (Polynomausdrücke in Matrizen). Es seien $p(t) = c_k t^k + \dots + c_1 t + c_0$ ein Polynom mit Koeffizienten in K und $A \in K^{n \times n}$. Dann setzen wir

$$p(A) := c_k A^k + \dots + c_1 A + c_0 E \in K^{n \times n}.$$

Im Rest dieses Abschnitts wollen wir der Einfachheit halber annehmen, dass unser Grundkörper K gleich \mathbb{C} ist, so dass jedes Polynom $p \neq 0$ nach dem Fundamentalsatz 6.11 der Algebra in Linearfaktoren zerfällt und somit als $p(t) = c(t - \lambda_1) \dots (t - \lambda_k)$ für $c, \lambda_1, \dots, \lambda_k \in \mathbb{C}$ mit $c \neq 0$ geschrieben werden kann. Einsetzen einer Matrix ergibt dann $p(A) = c(A - \lambda_1 E) \dots (A - \lambda_k E)$. Da wir jede reelle Matrix auch als komplexe auffassen können, ist diese Beschränkung auf \mathbb{C} für uns kein wesentliches Problem. Wie schon in der Einleitung zu diesem Kapitel erwähnt, kann man in der Tat sogar zu jedem Körper einen größeren finden, in dem Polynome immer in Linearfaktoren zerfallen. Mit einer solchen Aussage würden sich die Ergebnisse, die wir jetzt zeigen werden, dann auch auf beliebige Körper übertragen lassen.

Statt des in Beispiel 20.25 (b) erwähnten Problems, zu einem gegebenen komplexen Polynom p alle quadratischen Matrizen $A \in \mathbb{C}^{n \times n}$ mit $p(A) = 0$ zu finden, wollen wir diese Frage nun zunächst umdrehen und zu einer gegebenen Matrix A alle Polynome p mit $p(A) = 0$ bestimmen. Es wird

sich dabei herausstellen, dass es unter diesen Polynomen ein *eindeutiges* normiertes Polynom mit minimalem Grad gibt, das wir dann das Minimalpolynom von A nennen werden.

Für die konkrete Berechnung, ob $p(A) = 0$ gilt, ist es natürlich nützlich, wenn die Matrix A eine möglichst einfache Form hat. Daher werden wir diese Frage zunächst für Jordanblöcke untersuchen (wo sie einfach zu beantworten ist), dann für Matrizen in Jordanform, und schließlich mit Hilfe von Ähnlichkeitstransformationen für beliebige Matrizen.

Lemma 20.27. *Es seien $p \neq 0$ ein komplexes Polynom und $A = J_m(\lambda) \in \mathbb{C}^{m \times m}$ ein Jordanblock wie in Konstruktion 20.12.*

Dann gilt $p(A) = 0$ genau dann, wenn λ in p eine Nullstelle der Vielfachheit mindestens m ist.

Beweis. Wir können p als $p(t) = c(t - \lambda_1)^{a_1} \cdots (t - \lambda_k)^{a_k}$ mit $c \in \mathbb{C} \setminus \{0\}$, $a_1, \dots, a_k \in \mathbb{N}$ und verschiedenen $\lambda_1, \dots, \lambda_k$ faktorisieren, so dass also

$$p(A) = c(A - \lambda_1 E)^{a_1} \cdots (A - \lambda_k E)^{a_k}. \tag{1}$$

Ein einzelner Faktor in $p(A)$ hat dabei die Form

$$A - \lambda_i E = \begin{pmatrix} \lambda - \lambda_i & 1 & & 0 \\ & & \ddots & \\ & & & \ddots & \\ 0 & & & & 1 \\ & & & & & \lambda - \lambda_i \end{pmatrix}. \tag{2}$$

Für alle i mit $\lambda_i \neq \lambda$ hat diese Matrix Determinante $(\lambda - \lambda_i)^m \neq 0$ und ist damit invertierbar. In (1) kann also höchstens dann $p(A) = 0$ gelten, wenn $\lambda_i = \lambda$ für ein i gilt; und in diesem Fall ist $p(A) = 0$ äquivalent zu $(A - \lambda_i E)^{a_i} = 0$. Die Matrix $A - \lambda_i E$ hat dann die Form (2) mit Nullen auf der Diagonale, und bildet damit die Einheitsvektoren e_1, \dots, e_m auf

$$(A - \lambda_i E) e_j = \begin{cases} e_{j-1} & \text{für } j > 1, \\ 0 & \text{für } j = 1 \end{cases}$$

ab, verschiebt also jeden Einheitsvektor auf den vorherigen und bildet e_1 auf 0 ab. Die a_1 -fache Anwendung $(A - \lambda_i E)^{a_1}$ dieser Operation bildet damit e_j auf e_{j-a_1} ab für $j > a_1$, und auf 0 für $j \leq a_1$. Sie ist also genau dann gleich 0, bildet also alle Einheitsvektoren e_1, \dots, e_m auf 0 ab, wenn $m \leq a_1$ gilt, also wie behauptet wenn die Vielfachheit a_1 der Nullstelle λ in p mindestens m ist. \square

Bemerkung 20.28. Es seien p ein komplexes Polynom und $A \in \mathbb{C}^{n \times n}$.

- (a) ($p(A)$ für Blockdiagonalmatrizen) Ist A in Blockdiagonalform mit quadratischen Blöcken A_1, \dots, A_k , so gilt aufgrund der Blockmultiplikation aus Bemerkung 15.8

$$A = \begin{pmatrix} \boxed{A_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{A_k} \end{pmatrix} \Rightarrow A^i = \begin{pmatrix} \boxed{A_1^i} & & 0 \\ & \ddots & \\ 0 & & \boxed{A_k^i} \end{pmatrix} \text{ für alle } i \in \mathbb{N},$$

und damit auch für den Ausdruck $p(A)$, der ja eine Linearkombination solcher Potenzen ist,

$$p(A) = \begin{pmatrix} \boxed{p(A_1)} & & 0 \\ & \ddots & \\ 0 & & \boxed{p(A_k)} \end{pmatrix}.$$

Also ist $p(A) = 0$ genau dann, wenn $p(A_j) = 0$ für alle $j = 1, \dots, k$.

- (b) ($p(A)$ für ähnliche Matrizen) Ist $T \in GL(n, \mathbb{C})$ und damit $A' := T^{-1}AT$ ähnlich zu A , so gilt $p(A') = 0$ genau dann, wenn $p(A) = 0$: Es ist nämlich

$$(A')^i = \underbrace{(T^{-1}AT)(T^{-1}AT) \cdots (T^{-1}AT)}_{i\text{-mal}} = T^{-1}A^i T$$

für alle $i \in \mathbb{N}$, da sich die Matrixprodukte TT^{-1} hier in der Mitte herauskürzen, und damit für $p(t) = c_n t^n + \dots + c_1 t + c_0$

$$p(A') = T^{-1}(c_n A^n + \dots + c_1 A + c_0 E)T = T^{-1}p(A)T,$$

was wegen der Invertierbarkeit von T genau dann gleich 0 ist, wenn $p(A) = 0$ gilt.

Folgerung und Definition 20.29 (Minimalpolynom). *Es seien $\lambda_1, \dots, \lambda_k$ die verschiedenen Eigenwerte einer Matrix $A \in \mathbb{C}^{n \times n}$. Ferner sei a_i für $i = 1, \dots, k$ die maximale Größe eines Jordanblocks zum Eigenwert λ_i in der Jordanform von A .*

(a) *Es gibt ein eindeutiges normiertes Polynom p_A minimalen Grades mit $p_A(A) = 0$, nämlich*

$$p_A(t) = (t - \lambda_1)^{a_1} \dots (t - \lambda_k)^{a_k}.$$

*Man nennt p_A das **Minimalpolynom** von A .*

(b) *Für ein beliebiges Polynom p gilt*

$$p(A) = 0 \Leftrightarrow \text{es gibt ein Polynom } q \text{ mit } p = q \cdot p_A$$

(man sagt dafür auch, dass p ein Vielfaches von p_A bzw. p_A ein Teiler von p ist).

Beweis. Nach Folgerung 20.14 und Bemerkung 20.28 (b) können wir annehmen, dass A eine Matrix in Jordanform ist. Ist $p \neq 0$ dann ein beliebiges komplexes Polynom, so ist nach Bemerkung 20.28 (a) genau dann $p(A) = 0$, wenn p ausgewertet an jedem Jordanblock von A gleich 0 ist, was wiederum nach Lemma 20.27 genau dann der Fall ist, wenn die Nullstellenordnung von allen λ_i in p mindestens so groß ist wie jeder Jordanblock zu λ_i , also mindestens so groß wie a_i . Die Polynome p mit $p(A) = 0$ sind also genau diejenigen der Form

$$p(t) = q \cdot (t - \lambda_1)^{a_1} \dots (t - \lambda_k)^{a_k}$$

für ein Polynom q . Dies zeigt bereits beide Teile der Folgerung. \square

Bemerkung 20.30.

- (a) Nach Folgerung 20.29 (a) sind die Nullstellen des Minimalpolynoms p_A genau alle Eigenwerte von A .
- (b) Da jeder Jordanblock der Größe m einer Spalte der Länge m im Jordandiagramm entspricht, ist die maximale Größe eines solchen Blocks genau die Anzahl der Zeilen im Jordandiagramm. Zusammen mit Bemerkung 20.15 können wir also für jeden Eigenwert einer komplexen Matrix als Merkregel für die Jordandiagramme festhalten:

Anzahl der Kästchen im Jordandiagramm = algebraische Vielfachheit
 Anzahl der Spalten im Jordandiagramm = geometrische Vielfachheit
 Anzahl der Zeilen im Jordandiagramm = Vielfachheit im Minimalpolynom

- (c) Für alle, die schon die Vorlesung „Algebraische Strukturen“ gehört haben, kann man Folgerung 20.29 auch so formulieren: Man sieht leicht, dass die Menge aller Polynome p mit $p(A) = 0$ ein Ideal im Polynomring $\mathbb{C}[t]$ ist. Da $\mathbb{C}[t]$ ein Hauptidealring ist [G, Beispiel 10.23], kann dieses Ideal von einem Element erzeugt werden – und dieser Erzeuger ist bis auf Multiplikation mit Konstanten in $\mathbb{C} \setminus \{0\}$ eindeutig. Das damit eindeutig bestimmte normierte Polynom, das dieses Ideal erzeugt, ist genau das Minimalpolynom von A .

Beispiel 20.31. Beachte, dass Folgerung 20.29 das Minimalpolynom einer Matrix A auf zwei ganz unterschiedliche Arten beschreibt: als normiertes Polynom kleinsten Grades, das beim Einsetzen von A Null ergibt, und als Polynom, dessen Nullstellenordnung bei jedem λ die maximale Größe eines Jordanblocks zum Eigenwert λ in der Jordanform von A ist. Um die Nützlichkeit beider

Beschreibungen zu sehen, betrachten wir noch einmal die reellen Matrizen

$$A = \begin{pmatrix} 2 & 1 \\ -2 & -1 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 0 & 1 & 1 & 0 \\ -4 & 4 & 2 & 0 \\ 0 & 0 & 2 & 0 \\ 1 & 0 & 1 & 3 \end{pmatrix}.$$

- (a) Für die Matrix A haben wir in Beispiel 20.25 (a) bereits gesehen, dass $p(A) = 0$ für $p(t) = t^2 - t = t(t - 1)$. Nach Folgerung 20.29 (b) muss dieses Polynom ein Vielfaches des Minimalpolynoms p_A sein. Für p_A kommen also nur die Polynome $t(t - 1)$, t und $t - 1$ in Frage. Da die letzten beiden Polynome beim Einsetzen von A aber offensichtlich nicht 0 ergeben (es ist $A \neq 0$ und $A - E \neq 0$), ist $p_A(t) = t^2 - t$.

In diesem Fall konnten wir das Minimalpolynom also einfach berechnen, ohne irgendetwas über die Eigenwerte oder die Jordanform von A zu wissen.

- (b) Von der Matrix B haben wir in Beispiel 20.17 schon die Jordanform berechnet; sie bestand aus zwei Jordanblöcken zum Eigenwert 2 mit den Größen 1 und 2 sowie einem Jordanblock zum Eigenwert 3 der Größe 1. Mit Folgerung 20.29 (a) können wir daraus also sofort ablesen, dass $p_B(t) = (t - 2)^2(t - 3)$.

Hier konnten wir durch die Kenntnis der Jordanform also das Minimalpolynom bestimmen, ohne irgendwelche Polynomausdrücke in Matrizen berechnen zu müssen.

Aus unserer Charakterisierung des Minimalpolynoms ergeben sich zwei unmittelbare interessante Folgerungen.

Folgerung 20.32 (Satz von Cayley-Hamilton). Für jede quadratische Matrix $A \in \mathbb{C}^{n \times n}$ gilt $\chi_A(A) = 0$.

Beweis. Das charakteristische Polynom von A hat die Form

$$\chi_A(t) = (t - \lambda_1)^{\mu_a(A, \lambda_1)} \dots (t - \lambda_k)^{\mu_a(A, \lambda_k)},$$

wobei $\lambda_1, \dots, \lambda_k$ die verschiedenen Eigenwerte von A sind. Für alle i ist aber nun $\mu_a(A, \lambda_i)$ nach Bemerkung 20.30 (b) die Anzahl der Kästchen im Jordandiagramm zu λ_i , und damit sicher mindestens so groß wie die Anzahl der Zeilen darin, also wie die Vielfachheit von λ_i im Minimalpolynom p_A . Wie in Folgerung 20.29 ist χ_A also ein Vielfaches des Minimalpolynoms p_A , und damit ist $\chi_A(A) = 0$. \square

Folgerung 20.33 (Diagonalisierbarkeit und Minimalpolynom). Eine Matrix $A \in \mathbb{C}^{n \times n}$ ist genau dann diagonalisierbar, wenn ihr Minimalpolynom p_A nur einfache Nullstellen hat.

Beweis. Nach Bemerkung 20.16 (b) ist A genau dann diagonalisierbar, wenn alle Jordandiagramme von A nur Spalten der Länge 1 haben, also genau eine Zeile besitzen. Nach Bemerkung 20.30 (b) ist dies äquivalent dazu, dass alle Eigenwerte einfache Nullstellen von p_A sind. \square