

Kapitel 4

Eigenwerte

In Abschnitt 2.2.4 hatten wir für eine lineare Abbildung $F: V \rightarrow W$ ein Paar von Basen konstruiert, bezüglich derer F durch

$$\begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}$$

mit $r = \text{rang } F$ dargestellt wird. Die nötigen Transformationsmatrizen sind ganz einfach explizit zu berechnen (vgl. 2.7.6).

Zur Beschreibung eines Endomorphismus benutzt man eine einzige Basis, und ihre Anpassung an eine lineare Abbildung ist weit schwieriger als wenn man zwei Basen variieren kann. In die Matrizenrechnung übersetzt bedeutet diese Frage, zu einer quadratischen Matrix A eine möglichst einfache ähnliche Matrix

$$B = SAS^{-1}$$

zu finden (vgl. 2.6.7). Insbesondere wird sich zeigen, daß man hierzu stärkere Hilfsmittel aus der Algebra, nämlich Polynome höheren Grades benötigt, obwohl in der linearen Algebra zunächst nur der Grad 1 interessiert. Die benutzten Tatsachen über Polynome sind in Abschnitt 1.3 zusammengestellt.

4.1 Beispiele und Definitionen

4.1.1. Wir beginnen mit dem einfachsten Fall, nämlich einem K -Vektorraum V mit $\dim V = 1$. Ist $F \in \text{End}(V)$ und $0 \neq v \in V$, so gibt es ein eindeutig bestimmtes $\lambda \in K$ mit

$$F(v) = \lambda \cdot v,$$

und dieses λ hängt nicht von der Wahl von v ab. Denn ist $w = \mu v$ ein anderer Vektor von V , so ist

$$F(w) = F(\mu v) = \mu \lambda v = \lambda \mu v = \lambda w.$$

Die Abbildung F ist also durch die Zahl λ festgelegt.

Auch in höherdimensionalen Räumen darf man auf Vektoren hoffen, die bei Anwendung von F nur mit einem Faktor multipliziert werden, weil dann die Wirkung von F wenigstens in bestimmten Richtungen besonders einfach zu verstehen ist. Das führt zu der

Definition. Sei F ein Endomorphismus des K -Vektorraums V . Ein $\lambda \in K$ heißt *Eigenwert* von F , wenn es ein $v \in V$ mit $v \neq 0$ gibt, so daß gilt

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

Jedes vom Nullvektor verschiedene $v \in V$ mit $F(v) = \lambda v$ heißt *Eigenvektor* von F (zum Eigenwert λ).

Vorsicht! Man beachte, daß natürlich $0 \in K$ ein Eigenwert sein kann, der Nullvektor $0 \in V$ jedoch nach Definition nie Eigenvektor ist.

Das zentrale Problem dieses Kapitels ist die Frage nach der Existenz und der Vielfalt von Eigenvektoren. Dazu zunächst drei

Beispiele. a) Es sei $V = \mathbb{R}^2$ und F eine Drehung um den Winkel α , die beschrieben wird durch die Matrix

$$A = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}.$$

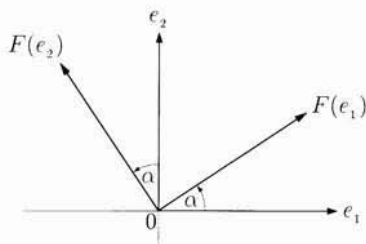


Bild 4.1

Anschaulich ist klar, daß es mit Ausnahme der Fälle $\alpha = 0$ und $\alpha = \pi$ keinen Eigenvektor geben kann.

b) Wir variieren das erste Beispiel, indem wir die Richtung von $F(e_2)$ umkehren. Dann ist die beschreibende Matrix

$$A' = \begin{pmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{pmatrix}.$$

Wenn man die Gerade mit dem Winkel $\frac{\alpha}{2}$ einzeichnet, erkennt man, daß die Abbildung F eine Spiegelung an dieser Geraden ist. Damit hat man zwei Eigenvektoren gefunden:

$$\begin{aligned} v_1 &= \left(\cos \frac{\alpha}{2}, \sin \frac{\alpha}{2} \right) && \text{zum Eigenwert } \lambda_1 = 1 \text{ und} \\ v_2 &= \left(\cos \frac{\alpha+\pi}{2}, \sin \frac{\alpha+\pi}{2} \right) && \text{zum Eigenwert } \lambda_2 = -1, \end{aligned}$$

und diese beiden bilden eine Basis $\mathcal{B} = (v_1, v_2)$ von $\mathbb{R}^2$ mit

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

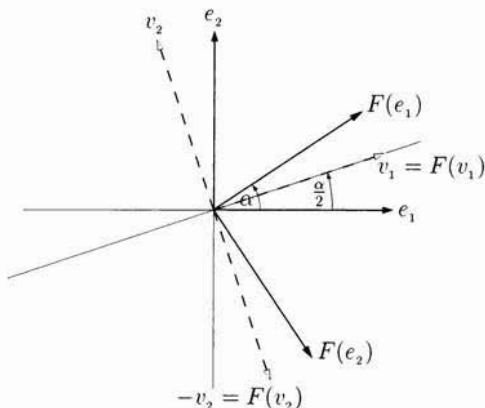


Bild 4.2

c) Sei $I \subset \mathbb{R}$ ein Intervall und $V = \mathcal{D}(I; \mathbb{R})$ der unendlichdimensionale $\mathbb{R}$ -Vektorraum der auf I beliebig oft differenzierbaren Funktionen. Ein Endomorphismus ist gegeben durch

$$F: V \rightarrow V, \quad \varphi \mapsto \varphi'.$$

Dieses F hat jedes beliebige $\lambda \in \mathbb{R}$ als Eigenwert, denn

$$\varphi(x) := ce^{\lambda x}$$

ist für jedes $c \in \mathbb{R}^*$ ein Eigenvektor zu λ . Das ist ein erster Hinweis auf die Bedeutung der Eigenwerttheorie für Differentialgleichungen (vgl. auch Aufgabe 3).

4.1.2. Endomorphismen wie im obigen Beispiel b) erhalten einen eigenen Namen:

Definition. Ein Endomorphismus heißt *diagonalisierbar*, wenn es eine Basis aus Eigenvektoren gibt.

Dieser Name erklärt sich aus der

Bemerkung. Ist $\dim V = n < \infty$, so ist $F \in \text{End}(V)$ genau dann diagonalisierbar, wenn es eine Basis $\mathcal{B} = (v_1, \dots, v_n)$ von V gibt, so daß $M_{\mathcal{B}}(F)$ eine Diagonalmatrix ist, d.h.

$$M_{\mathcal{B}}(F) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Das folgt unmittelbar aus der Definition der darstellenden Matrix. $\square$

Insbesondere nennt man eine Matrix $A \in M(n \times n; K)$ diagonalisierbar, wenn der durch A beschriebene Endomorphismus von K^n diagonalisierbar ist. Nach 2.6.7 ist das gleichwertig mit der Existenz einer Matrix $S \in GL(n; K)$ mit

$$SAS^{-1} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

d.h. A ist ähnlich zu einer Diagonalmatrix.

Es sei daran erinnert, daß die Spalten von S^{-1} eine Basis aus Eigenvektoren von A sind. Das sieht man zum Beispiel durch Multiplikation obiger Gleichung von links mit S^{-1} .

Vorsicht! Selbst wenn F diagonalisierbar ist, braucht nicht *jeder* Vektor ungleich Null aus V ein Eigenvektor zu sein. Man mache sich das an Beispielen klar (etwa Beispiel b) oben oder Bild 4.3)!

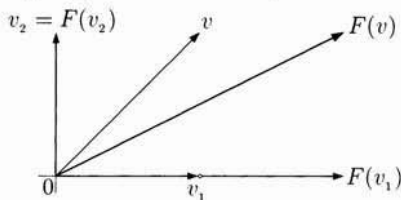


Bild 4.3

4.1.3. Bevor wir eine Methode angeben, Eigenwerte zu finden, beweisen wir den

Satz. Angenommen, $F \in \text{End}(V)$ hat paarweise verschiedene Eigenwerte $\lambda_1, \dots, \lambda_n$, wobei $n = \dim V$. Dann ist F diagonalisierbar.

Da es zu jedem Eigenwert λ_i mindestens einen Eigenvektor $v_i \neq 0$ gibt, folgt der Satz sofort aus dem

Lemma. Gegeben seien $F \in \text{End}(V)$ und Eigenvektoren $v_1, \dots, v_m$ zu paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_m$. Dann sind $v_1, \dots, v_m$ linear unabhängig. Insbesondere ist $m \leq \dim V$.

Beweis. Wir führen Induktion über m . Der Fall $m = 1$ ist wegen $v_1 \neq 0$ klar. Sei $m \geq 2$, und sei die Aussage für $m - 1$ bereits bewiesen. Wir betrachten die Bedingung

$$\alpha_1 v_1 + \dots + \alpha_m v_m = 0. \quad (*)$$

Darauf können wir einerseits F anwenden, andererseits mit λ_1 multiplizieren und die Ergebnisse voneinander subtrahieren. Das ergibt

$$\begin{aligned} \alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 + \dots + \alpha_m \lambda_m v_m &= 0, \\ \alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_1 v_2 + \dots + \alpha_m \lambda_1 v_m &= 0 \end{aligned}$$

und

$$\alpha_2 (\lambda_2 - \lambda_1) v_2 + \dots + \alpha_m (\lambda_m - \lambda_1) v_m = 0.$$

Nach Induktionsannahme sind $v_2, \dots, v_m$ linear unabhängig, also ist

$$\alpha_2 (\lambda_2 - \lambda_1) = \dots = \alpha_m (\lambda_m - \lambda_1) = 0,$$

und da die Eigenwerte verschieden sind, folgt

$$\alpha_2 = \dots = \alpha_m = 0.$$

Setzt man das in $(*)$ ein, so folgt $\alpha_1 v_1 = 0$ und wegen $v_1 \neq 0$ auch $\alpha_1 = 0$. $\square$

4.1.4. Wie wir gesehen haben, gibt es höchstens $n = \dim V$ Eigenwerte, dagegen im allgemeinen sehr viel mehr Eigenvektoren. Daher ist es nützlich, alle Eigenvektoren zu einem festen Eigenwert zusammenzufassen.

Definition. Ist F ein Endomorphismus von V und $\lambda \in K$, so nennen wir

$$\text{Eig}(F; \lambda) := \{v \in V : F(v) = \lambda v\}$$

den *Eigenraum* von F bezüglich λ .

Bemerkung.

- a) $\text{Eig}(F; \lambda) \subset V$ ist ein Untervektorraum.
- b) λ ist Eigenwert von $F \Leftrightarrow \text{Eig}(F; \lambda) \neq \{0\}$.
- c) $\text{Eig}(F; \lambda) \setminus \{0\}$ ist die Menge der zu λ gehörigen Eigenvektoren von F .

d) $\text{Eig}(F; \lambda) = \text{Ker}(F - \lambda \text{id}_V)$.

e) Sind $\lambda_1, \lambda_2 \in K$ verschieden, so ist $\text{Eig}(F; \lambda_1) \cap \text{Eig}(F; \lambda_2) = \{0\}$.

Beweis. a) bis d) ist ganz klar, e) folgt aus dem Lemma in 4.1.3. $\square$

Aufgaben zu 4.1

1. Zeigen Sie: Ein nilpotenter Endomorphismus hat null als einzigen Eigenwert.

2. Gegeben sei die lineare Abbildung $F: \mathcal{D}(I; \mathbb{R}) \rightarrow \mathcal{D}(I; \mathbb{R})$, $\varphi \mapsto \varphi'$, wobei $I \subset \mathbb{R}$ ein Intervall ist.

a) Bestimmen Sie die reellen Eigenwerte von F .

b) Bestimmen Sie eine Basis von $\text{Eig}(F, -1)$.

3. Sei $I \subset \mathbb{R}$ ein offenes Intervall. Durch eine Matrix $A \in M(n \times n; \mathbb{R})$ ist das homogene lineare Differentialgleichungssystem

$$y' = A \cdot y$$

bestimmt; nach Aufgabe 11 zu 3.2 hat der zugehörige Lösungsraum

$$\mathcal{L}_0 = \{\varphi \in \mathcal{D}(I; \mathbb{R}^n): \varphi' = A \cdot \varphi\} \subset \mathcal{D}(I; \mathbb{R}^n)$$

die Dimension n . Um Lösungen zu erhalten, kann man den Ansatz

$$\varphi(t) = e^{\lambda t} \cdot v$$

benutzen, wobei $\lambda \in \mathbb{R}$ und $v \in \mathbb{R}^n$. Zeigen Sie:

a) $\varphi(t) = e^{\lambda t} \cdot v$ ist eine Lösung $\neq 0$ von $y' = A \cdot y$ genau dann, wenn v Eigenvektor von A zum Eigenwert λ ist.

b) Lösungen $\varphi^{(1)}(t) = e^{\lambda_1 t} \cdot v_1, \dots, \varphi^{(k)}(t) = e^{\lambda_k t} \cdot v_k$ sind linear unabhängig genau dann, wenn $v_1, \dots, v_k$ linear unabhängig sind.

Insbesondere erhält man mit diesem Ansatz eine Basis des Lösungsraums, falls A diagonalisierbar ist.

4. Sei V ein K -Vektorraum und $F: V \rightarrow V$ linear. Zeigen Sie: Hat $F^2 + F$ den Eigenwert -1 , so hat F^3 den Eigenwert 1 .

5. Gegeben sei ein K -Vektorraum V und $F, G \in \text{End}(V)$. Beweisen Sie:

a) Ist $v \in V$ Eigenvektor von $F \circ G$ zum Eigenwert $\lambda \in K$, und ist $G(v) \neq 0$, so ist $G(v)$ Eigenvektor von $G \circ F$ zum Eigenwert λ .

b) Ist V endlichdimensional, so haben $F \circ G$ und $G \circ F$ dieselben Eigenwerte.

4.2 Das charakteristische Polynom

In diesem Abschnitt entwickeln wir eine systematische Methode zur Suche von Eigenwerten und zur Bestimmung der zugehörigen Eigenräume. Entscheidend dafür ist die Determinante eines Endomorphismus, von der wir in 3.4.1 gesehen hatten, daß sie wohldefiniert ist, wenn der zugrundeliegende Vektorraum *endlichdimensional* ist. Das muß in diesem ganzen Abschnitt stillschweigend vorausgesetzt werden.

4.2.1. Grundlegend ist die folgende einfache

Bemerkung. Für $F \in \text{End}(V)$ und $\lambda \in K$ sind folgende Bedingungen gleichwertig:

i) λ ist Eigenwert von F .

ii) $\det(F - \lambda \text{id}_V) = 0$.

Beweis. Für festes $\lambda \in K$ ist die Existenz eines $v \neq 0$ mit $F(v) = \lambda v$ gleichbedeutend mit

$$\begin{aligned} F(v) - \lambda v &= 0 \\ \Leftrightarrow (F - \lambda \text{id}_V)(v) &= 0 && \text{wegen der Linearität,} \\ \Leftrightarrow \text{Ker}(F - \lambda \text{id}_V) &\neq \{0\} && \text{nach der Definition des Kerns,} \\ \Leftrightarrow \text{Im}(F - \lambda \text{id}_V) &\neq V && \text{nach der Dimensionsformel aus 2.2.4,} \\ \Leftrightarrow \text{rang}(F - \lambda \text{id}_V) &< \dim V && \text{nach Definition des Ranges,} \\ \Leftrightarrow \det(F - \lambda \text{id}_V) &= 0 && \text{nach 3.4.1.} \quad \square \end{aligned}$$

4.2.2. Durch die obige Bemerkung ist die Suche nach Eigenwerten zurückgeführt auf die Suche nach Nullstellen der Abbildung

$$\tilde{P}_F: K \rightarrow K, \quad \lambda \mapsto \det(F - \lambda \text{id}_V).$$

Diese nennen wir die *charakteristische Funktion* von F . Wir zeigen, daß sie durch ein Polynom beschrieben wird.

Sei dazu $\mathcal{A}$ eine Basis von V und $A = M_{\mathcal{A}}(F)$. Ist t eine Unbestimmte, so definieren wir

$$P_A(t) := \det(A - t \cdot E_n) = \begin{vmatrix} a_{11} - t & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - t & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - t \end{vmatrix}.$$

Die Definition dieser Determinante ist etwas problematisch, weil in der Matrix die Unbestimmte t vorkommt. Man hat mehrere Möglichkeiten, damit umzugehen:

- 1) Man kann für t beliebige $\lambda \in K$ einsetzen und damit rechnen. Für unendlichen K ist das nach Korollar 2 aus 1.3.8 keine Einschränkung.
- 2) Man kann die Determinante formal nach der Formel von LEIBNIZ aus 3.2.5 ausrechnen. Eine Rechtfertigung dafür wurde in 3.2.8 gegeben.
- 3) Man kann die Einträge von $A - t \cdot E_n$ als Elemente des Körpers $K(t)$ der rationalen Funktionen ansehen (vgl. Aufgabe 10 in 1.3).

Das Ergebnis der Rechnung ist

$$P_A(t) = (a_{11} - t) \cdot \dots \cdot (a_{nn} - t) + Q,$$

wobei der erste Summand zur identischen Permutation gehört und Q die restliche Summe über $S_n \setminus \{\text{id}\}$ ist. Da in einem Summanden von Q als Faktoren höchstens $n - 2$ Diagonalkomponenten auftreten können, ist Q ein Polynom vom Grad höchstens $n - 2$. Nun ist

$$(a_{11} - t) \cdot \dots \cdot (a_{nn} - t) = (-1)^n t^n + (-1)^{n-1} (a_{11} + \dots + a_{nn}) t^{n-1} + Q_1,$$

wobei Q_1 ein Polynom vom Grad höchstens $n - 2$ ist. Also ist $P_A(t)$ ein Polynom vom Grad n mit Koeffizienten aus K , d.h. es gibt $\alpha_0, \dots, \alpha_n \in K$, so daß

$$P_A(t) = \alpha_n t^n + \alpha_{n-1} t^{n-1} + \dots + \alpha_1 t + \alpha_0.$$

Dabei ist

$$\begin{aligned} \alpha_n &= (-1)^n, \\ \alpha_{n-1} &= (-1)^{n-1} (a_{11} + \dots + a_{nn}) \quad \text{und} \\ \alpha_0 &= \det A. \end{aligned}$$

Man nennt $(a_{11} + \dots + a_{nn})$ auch die *Spur* von A . Die Koeffizienten $\alpha_1, \dots, \alpha_{n-2}$ sind nicht so leicht aufzuschreiben, sie haben auch keine speziellen Namen.

Diese Überlegung zeigt, daß $P_A(t)$ ein Element des Polynomrings $K[t]$ ist (siehe 1.3.6). Man nennt

$$P_A(t) = \det(A - t \cdot E_n) \in K[t]$$

das *charakteristische Polynom* der $(n \times n)$ -Matrix A .

Setzt man für die Unbekannte t ein $\lambda \in K$ ein, so erhält man eine Abbildung

$$\tilde{P}_A: K \rightarrow K, \quad \lambda \mapsto P_A(\lambda).$$

Nun zurück zu F . Für jedes $\lambda \in K$ ist

$$M_A(F - \lambda \cdot \text{id}_V) = A - \lambda \cdot E_n.$$

Also ist

$$\det(F - \lambda \cdot \text{id}_V) = \det(A - \lambda \cdot E_n) = P_A(\lambda),$$

d.h. die charakteristische Funktion von F ist beschrieben durch das charakteristische Polynom von A .

Ist nun $\mathcal{B}$ eine weitere Basis von V , so ist $B := M_{\mathcal{B}}(F)$ zu A ähnlich.

Lemma. Ähnliche Matrizen haben das gleiche charakteristische Polynom.

Beweis. Sei $B = SAS^{-1}$ mit $S \in \text{GL}(n; K)$. Eine formale Rechnung mit der Unbestimmten t (d.h. genauer eine Rechnung im Ring $M(n \times n; K[t])$, vgl. 3.2.8) ergibt

$$S \cdot t \cdot E_n \cdot S^{-1} = t \cdot E_n.$$

Also ist

$$B - t \cdot E_n = SAS^{-1} - S \cdot t \cdot E_n \cdot S^{-1} = S(A - t \cdot E_n)S^{-1}.$$

Anwendung der Determinante ergibt

$$\det(B - t \cdot E_n) = \det S \cdot \det(A - t \cdot E_n) \cdot (\det S)^{-1} = \det(A - t \cdot E_n)$$

und somit $P_B(t) = P_A(t)$. $\square$

4.2.3. Damit ist gezeigt, daß das charakteristische Polynom der darstellenden Matrix nicht von der Wahl der Basis abhängt. Also ist folgende Definition sinnvoll:

Sei F ein Endomorphismus und $\mathcal{A}$ eine Basis von V . Ist $A = M_{\mathcal{A}}(F)$, so nennen wir

$$P_F(t) := P_A(t)$$

das *charakteristische Polynom* von F .

Insgesamt haben wir folgendes bewiesen:

Satz. Sei V ein K -Vektorraum der Dimension $n < \infty$ und F ein Endomorphismus von V . Dann hat das charakteristische Polynom $P_F(t) \in K[t]$ die folgenden Eigenschaften:

1) $\deg P_F = n$.

2) P_F beschreibt die Abbildung $K \rightarrow K, \lambda \mapsto \det(F - \lambda \text{id}_V)$.

3) Die Nullstellen von P_F sind die Eigenwerte von F .

4) Ist A eine Matrix, die F darstellt, so ist $P_F(t) = \det(A - t \cdot E_n)$. $\square$

Damit ist das geometrische Problem, die Eigenwerte eines Endomorphismus zu finden, auf das algebraische Problem der Bestimmung der Nullstellen eines Polynoms zurückgeführt.

4.2.4. Hat man einen Eigenwert gefunden, so ist die Bestimmung des Eigenraumes ganz einfach. Wir können uns dabei auf den Fall $V = K^n$ beschränken. Aus 4.1.4 folgt sofort die

Bemerkung. Ist ein Endomorphismus $A: K^n \rightarrow K^n$ durch die Matrix $A \in M(n \times n; K)$ gegeben, so ist der Eigenraum $\text{Eig}(A; \lambda)$ für jedes $\lambda \in K$ der Lösungsraum des homogenen linearen Gleichungssystems

$$(A - \lambda E_n)x = 0. \quad \square$$

Nach der vielen Theorie ist es höchste Zeit für einige

Beispiele. a) Ist

$$A = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$$

die Matrix einer *Drehung* im $\mathbb{R}^2$ (vgl. 2.1.1), so ist

$$P_A(t) = t^2 - 2t \cos \alpha + 1.$$

Dieses quadratische Polynom hat nach 1.3.9 genau dann eine reelle Wurzel, wenn

$$4 \cos^2 \alpha - 4 \geq 0, \quad \text{d.h.} \quad \cos^2 \alpha = 1$$

gilt. Das ist nur für $\alpha = 0$ und $\alpha = \pi$ der Fall. Diese beiden Drehungen sind trivialerweise diagonalisierbar, alle anderen Drehungen besitzen keine Eigenvektoren. Damit ist noch einmal durch die allgemeine Theorie bestätigt, was wir in 4.1.1 anschaulich gesehen hatten.

b) Für eine Spiegelung

$$A = \begin{pmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{pmatrix}$$

ist $P_A(t) = t^2 - 1 = (t+1)(t-1)$. Also ist A nach 4.1.3 diagonalisierbar, was wir in 4.1.1 schon direkt bewiesen hatten.

c) Für

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

ist $P_A(t) = t^2 - 3t + 2 = (t-1)(t-2)$. Setzen wir $A_1 = A - E_2$, $A_2 = A - 2E_2$, so haben wir die linearen Gleichungssysteme

$$A_1 \cdot x = 0 \quad \text{und} \quad A_2 \cdot x = 0 \quad \text{mit Lösungen} \quad \begin{pmatrix} 3 \\ 1 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 2 \\ 1 \end{pmatrix},$$

das sind also Eigenvektoren zu den Eigenwerten 1 und 2. Zur Kontrolle berechnet

man

$$S \cdot A \cdot S^{-1} = \begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

d) Ist

$$\begin{aligned} A &= \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix}, \quad \text{so ist} \quad P_A = \begin{vmatrix} -t & -1 & 1 \\ -3 & -2-t & 3 \\ -2 & -2 & 3-t \end{vmatrix} \\ &= -t \cdot \begin{vmatrix} -2-t & 3 \\ -2 & 3-t \end{vmatrix} + 3 \cdot \begin{vmatrix} -1 & 1 \\ -2 & 3-t \end{vmatrix} - 2 \cdot \begin{vmatrix} -1 & 1 \\ -2-t & 3 \end{vmatrix} \\ &= -t(t^2 - t) + 3(t - 1) - 2(t - 1) = -t^3 + t^2 + t - 1 \\ &= -(t - 1)^2(t + 1). \end{aligned}$$

Darauf kommen wir in 4.3.4 zurück.

Es darf nicht verschwiegen werden, daß die Beispiele c) und d) sehr künstlich sind, weil man die Nullstellen der charakteristischen Polynome mehr oder weniger schnell erkennen kann. Ist das nicht der Fall, so muß man die Methoden der Numerik verwenden, um die Nullstellen näherungsweise zu berechnen. Wie man ebenfalls in der Numerik lernt, gibt es Verfahren, Eigenwerte und Eigenvektoren in einem Aufwasch gemeinsam zu approximieren. Das hier angegebene Verfahren, zuerst die Eigenwerte und anschließend die Eigenvektoren zu suchen, hilft nur für die Theorie und in besonderen Glücksfällen, die bei den hier behandelten Beispielen inszeniert wurden. Dennoch haben derartige Beispiele einen Sinn: man kann erst einmal ohne Ablenkung durch größeren Rechenaufwand einen Lösungsweg deutlich erkennen.

Aufgaben zu 4.2

1. Berechnen Sie das charakteristische Polynom, die Eigenwerte und Eigenvektoren von

$$\begin{pmatrix} 2 & 2 & 3 \\ 1 & 2 & 1 \\ 2 & -2 & 1 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} -5 & 0 & 7 \\ 6 & 2 & -6 \\ -4 & 0 & 6 \end{pmatrix}.$$

2. Beweisen Sie: Ist $A \in M(2 \times 2; \mathbb{R})$ symmetrisch, so hat A reelle Eigenwerte.

3. Sei V ein endlichdimensionaler Vektorraum und $F \in \text{End}(V)$. Zeigen Sie, dass $P_F(0) \neq 0$ genau dann, wenn F ein Isomorphismus ist.

4. Zeigen Sie, dass die Matrix

$$A = \begin{pmatrix} 0 & \dots & 0 & -\alpha_0 \\ 1 & 0 & \dots & 0 & -\alpha_1 \\ & 1 & \ddots & 0 & \vdots \\ & & \ddots & 0 & \vdots \\ 0 & & & 1 & -\alpha_{n-1} \end{pmatrix}$$

das charakteristische Polynom $P_A(t) = (-1)^n(t^n + \alpha_{n-1}t^{n-1} + \dots + \alpha_1t + \alpha_0)$ besitzt.

5. Sei $A \in M(n \times n; K)$ und $\Phi: M(n \times n; K) \rightarrow M(n \times n; K)$ der Endomorphismus, der durch die Linksmultiplikation mit A gegeben ist, das heißt $\Phi(B) = AB$. Zeigen Sie, dass für die charakteristischen Polynome von A und Φ gilt: $P_\Phi = (P_A)^n$.

4.3 Diagonalisierung

4.3.1. Zunächst halten wir als Ergebnis der vorhergehenden Abschnitte folgendes fest:

Satz. Sei F ein Endomorphismus von V mit $n = \dim V$. Dann gilt:

- 1) Ist F diagonalisierbar, so ist $P_F = \pm(t - \lambda_1) \cdot \dots \cdot (t - \lambda_n)$, d.h. das charakteristische Polynom zerfällt in Linearfaktoren.
- 2) Ist $P_F = \pm(t - \lambda_1) \cdot \dots \cdot (t - \lambda_n)$ mit paarweise verschiedenen $\lambda_1, \dots, \lambda_n$, so ist F diagonalisierbar.

Beweis. 1) ist klar, weil man das charakteristische Polynom mit Hilfe einer Basis aus Eigenvektoren berechnen kann, und

$$P_F = \det \begin{pmatrix} \lambda_1 - t & & 0 \\ & \ddots & \\ 0 & & \lambda_n - t \end{pmatrix} = (\lambda_1 - t) \cdot \dots \cdot (\lambda_n - t).$$

2) folgt aus 4.1.3 und 4.2.3. □

4.3.2. Nach 4.3.1 bleibt also zu klären, wann F im Falle mehrfacher Nullstellen des charakteristischen Polynoms noch diagonalisierbar ist. Zu diesem Zweck fassen wir in P_F gleiche Linearfaktoren zusammen, d.h. wir schreiben

$$P_F = \pm(t - \lambda_1)^{r_1} \cdot \dots \cdot (t - \lambda_k)^{r_k},$$

wobei die $\lambda_1, \dots, \lambda_k$ paarweise verschieden sind, $1 \leq r_i \leq n$ für $i = 1, \dots, k$ und $r_1 + \dots + r_k = n$. Der Exponent r_i ist die Vielfachheit der Nullstelle λ_i von P_F , in der Notation von 1.3.8 ist $r_i = \mu(P_F; \lambda_i)$. Andererseits gehört zu λ_i der Eigenraum $\text{Eig}(F; \lambda_i)$.

Lemma. Ist λ Eigenwert von F , so gilt $1 \leq \dim \text{Eig}(F; \lambda) \leq \mu(P_F; \lambda)$.

Beweis. Sei $v_1, \dots, v_s$ eine Basis von $\text{Eig}(F; \lambda)$. Die Ungleichung $1 \leq s$ ist klar, da λ Eigenwert ist. Zum Beweis der zweiten Ungleichung ergänzen wir die Basis von $\text{Eig}(F; \lambda)$ zu einer Basis

$$\mathcal{A} = (v_1, \dots, v_s, v_{s+1}, \dots, v_n)$$

von V . Dann ist

$$A := M_{\mathcal{A}}(F) = \left(\begin{array}{ccc|ccc} \lambda & & 0 & & & \\ & \ddots & & & * & \\ 0 & & \lambda & & & \\ \hline & & & 0 & & A' \end{array} \right), \quad \left. \vphantom{\begin{pmatrix} \lambda & & 0 \\ & \ddots & \\ 0 & & \lambda \end{pmatrix}} \right\} s\text{-mal}$$

also $P_F = (t - \lambda)^s \cdot P_{A'}$ nach D8 und D9 in 3.1.3, und damit

$$\dim \operatorname{Eig}(F; \lambda) = s \leq \mu(P_F; \lambda).$$

□

4.3.3. Nun können wir zeigen, daß die obigen Ungleichungen ein leicht nachprüfbares Kriterium für die Diagonalisierbarkeit ergeben. Um den Beweis kürzer und die Aussage klarer zu machen, benutzen wir die in 1.6.4 bereitgestellten Tatsachen über direkte Summen.

Theorem. Sei V ein endlichdimensionaler K -Vektorraum und $F \in \operatorname{End}(V)$. Dann sind die folgenden Bedingungen äquivalent:

- i) F ist diagonalisierbar.
- ii) a) Das charakteristische Polynom P_F zerfällt in Linearfaktoren und
b) $\dim \operatorname{Eig}(F; \lambda) = \mu(P_F; \lambda)$ für alle Eigenwerte λ von F .
- iii) Sind $\lambda_1, \dots, \lambda_k$ die paarweise verschiedenen Eigenwerte von F , so ist

$$V = \operatorname{Eig}(F; \lambda_1) \oplus \dots \oplus \operatorname{Eig}(F; \lambda_k).$$

Beweis. i) $\Rightarrow$ ii). Ist F diagonalisierbar, so ordnen wir die Basis von V aus Eigenvektoren entsprechend den verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_k$, d.h. für $i = 1, \dots, k$ betrachten wir eine Basis

$$\left(v_1^{(i)}, \dots, v_{s_i}^{(i)} \right) \quad \text{von} \quad \operatorname{Eig}(F; \lambda_i).$$

Setzen wir $r_i := \mu(P_F, \lambda_i)$, so gilt

$$s_1 + \dots + s_k = n, \quad r_1 + \dots + r_k = n \quad \text{und} \quad s_i \leq r_i.$$

Das geht nur, wenn $s_i = r_i$ für alle i , das ist Bedingung ii) b).

ii) $\Rightarrow$ iii). Sei $W := \operatorname{Eig}(F; \lambda_1) + \dots + \operatorname{Eig}(F; \lambda_k)$. Nach Lemma 4.1.3 und der Definition der direkten Summe in 1.6.4 gilt

$$W = \operatorname{Eig}(F; \lambda_1) \oplus \dots \oplus \operatorname{Eig}(F; \lambda_k).$$

Aus ii) und Bedingung iii) in Satz 1.6.4 folgt $\dim W = r_1 + \dots + r_k = n$, also ist $W = V$.

iii) $\Rightarrow$ i). Für jedes $i \in \{1, \dots, k\}$ sei $\mathcal{B}_i = (v_1^{(i)}, \dots, v_{s_i}^{(i)})$ eine Basis von $\text{Eig}(F; \lambda_i)$. Nach 1.6.4 ist

$$\mathcal{B} := (v_1^{(1)}, \dots, v_{s_1}^{(1)}, \dots, v_1^{(k)}, \dots, v_{s_k}^{(k)})$$

eine Basis von V , und da sie nach Definition aus Eigenvektoren besteht, ist F diagonalisierbar. Insbesondere ist $s_i = r_i$ für alle i und

$$M_{\mathcal{B}}(F) = \left(\begin{array}{ccccccc} \lambda_1 & & & & & & \\ & \ddots & & & & & \\ & & \lambda_1 & & & & \\ & & & \ddots & & & \\ & & & & \ddots & & \\ & 0 & & & & \lambda_k & \\ & & & & & & \ddots \\ & & & & & & & \lambda_k \end{array} \right) \left. \begin{array}{l} \vdots \\ \vdots \\ \vdots \end{array} \right\} \begin{array}{l} r_1\text{-mal} \\ \vdots \\ r_k\text{-mal.} \end{array} \quad \square$$

4.3.4. Mit Hilfe der gerade abgeschlossenen theoretischen Überlegungen kann man ein praktisches Verfahren für die Diagonalisierung eines Endomorphismus F von einem endlichdimensionalen K -Vektorraum V angeben:

1. Schritt: Mit Hilfe einer Basis $\mathcal{A}$ von V und der Matrix $A = M_{\mathcal{A}}(F)$ berechne man das charakteristische Polynom. Wenn man eine Zerlegung in Linearfaktoren angeben kann, ist das erste und für die Rechnung entscheidende Hindernis überwunden.

2. Schritt: Für jeden Eigenwert λ von F bestimme man durch Lösung eines linearen Gleichungssystems eine Basis von $\text{Eig}(F; \lambda)$. Damit kann man kontrollieren, ob

$$\dim \text{Eig}(F; \lambda) = \mu(P_F; \lambda)$$

gilt. Genau dann, wenn dies für alle λ der Fall ist, ist F diagonalisierbar. Entsprechend 4.3.3 bilden dann die Basen der Eigenräume zusammen eine Basis $\mathcal{B}$ von V .

Beispiel. Sei $F: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ gegeben durch

$$F(x, y, z) := (-y + z, -3x - 2y + 3z, -2x - 2y + 3z).$$

Bezeichnet $\mathcal{K}$ die kanonische Basis von $\mathbb{R}^3$, so ist

$$A := M_{\mathcal{K}}(F) = \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix}$$

und $P_F = -t^3 + t^2 + t - 1 = -(t-1)^2(t+1)$. Also sind $\lambda_1 = 1$ und $\lambda_2 = -1$ die beiden einzigen Eigenwerte von F . $\text{Eig}(F; 1)$ ist der Lösungsraum von

$$\begin{pmatrix} -1 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3-1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

umgeformt $-x_1 - x_2 + x_3 = 0$. Es ist also $\mu(P_F; 1) = 2 = \dim \text{Eig}(F; 1)$, und $((1, 0, 1), (0, 1, 1))$ ist eine Basis von $\text{Eig}(F; 1)$.

Ebenso ist $\text{Eig}(F; -1)$ der Lösungsraum von

$$\begin{pmatrix} 1 & -1 & 1 \\ -3 & -2+1 & 3 \\ -2 & -2 & 3+1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

umgeformt

$$x_1 - x_2 + x_3 = 0, \quad -4x_2 + 6x_3 = 0.$$

Es ist also $\mu(P_F; -1) = 1 = \dim \text{Eig}(F; -1)$, und $(1, 3, 2)$ ist eine Basis von $\text{Eig}(F; -1)$. Somit ist

$$\mathcal{B} = ((1, 0, 1), (0, 1, 1), (1, 3, 2))$$

eine Basis von $\mathbb{R}^3$, bestehend aus Eigenvektoren von F . Für die Transformationsmatrix $S = T_{\mathcal{B}}^{\mathcal{K}}$ gilt

$$S^{-1} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix}, \quad \text{also} \quad S = \frac{1}{2} \begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ 1 & 1 & -1 \end{pmatrix}.$$

Für

$$D := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

muß somit $D = SAS^{-1}$ gelten, was man zur Kontrolle der Rechnung benutzen kann.

4.3.5. Ist eine Masse an einer Feder aufgehängt und zur Zeit $t = 0$ in senkrechter Richtung in die Position $y(0) = \alpha$ mit Geschwindigkeit $\dot{y}(0) = \beta$ ausgelenkt, so ist die weitere Bewegung bestimmt durch die Differentialgleichung der gedämpften Schwingung

$$\ddot{y} + 2\mu\dot{y} + \omega^2 y = 0, \quad y(0) = \alpha, \quad \dot{y}(0) = \beta.$$

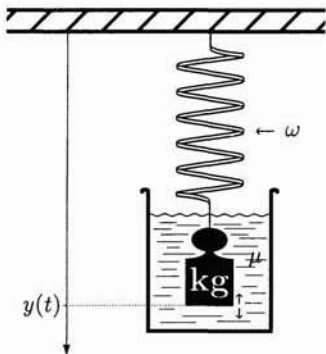


Bild 4.4

Dabei sind $\omega, \mu \in \mathbb{R}_+$ Konstanten, ω ist durch die Feder und μ durch die Reibung bestimmt. Wie üblich macht man daraus mit $y_0 = y$ und $y_1 = \dot{y}$ das lineare System erster Ordnung

$$\begin{aligned} \dot{y}_0 &= y_1, & y_0(0) &= \alpha, \\ \dot{y}_1 &= -\omega^2 y_0 - 2\mu y_1, & y_1(0) &= \beta. \end{aligned} \quad (*)$$

Das führt auf die Matrix mit von t unabhängigen Einträgen

$$A = \begin{pmatrix} 0 & 1 \\ -\omega^2 & -2\mu \end{pmatrix}.$$

Einer Diagonalisierung von A entspricht eine Entkoppelung des Systems (*), und wenn A zu einer oberen Dreiecksmatrix gemacht werden kann, ist das System von unten nach oben explizit lösbar (vgl. [Fo2], §14). Daher betrachten wir das charakteristische Polynom

$$P_A(\lambda) = \lambda^2 + 2\mu\lambda + \omega^2, \quad \text{also} \quad \lambda = -\mu \pm \sqrt{\mu^2 - \omega^2}.$$

Entscheidend für die Art der Bewegung ist die Diskriminante $\mu^2 - \omega^2$. Es sind drei Fälle möglich.

$$1) \quad 0 \leq \mu < \omega, \text{ d.h. } \mu^2 - \omega^2 < 0, \quad (\text{schwache Dämpfung})$$

2) $\mu = \omega$, d.h. $\mu^2 - \omega^2 = 0$, (aperiodischer Grenzfall)

3) $\mu > \omega$, d.h. $\mu^2 - \omega^2 > 0$ (starke Dämpfung)

Im Fall 3) gibt es zwei verschiedene negative Eigenwerte

$$\lambda_1 = -\mu + \sqrt{\mu^2 - \omega^2}, \quad \lambda_2 = -\mu - \sqrt{\mu^2 - \omega^2},$$

also ist A diagonalisierbar.

Im Fall 2) ist $\lambda = -\mu = -\omega$ ein 2-facher Eigenwert, die Matrix

$$A - \lambda E_2 = \begin{pmatrix} \omega & 1 \\ -\omega^2 & -\omega \end{pmatrix}$$

hat den Rang 1, also ist $\dim \text{Eig}(A; -\omega) = 1$, und A ist nicht diagonalisierbar. In 4.4 werden wir sehen, daß A trigonalisierbar ist.

Im Fall 1) gibt es keine reellen Eigenwerte, dafür aber zwei verschiedene komplexe. Also ist A komplex diagonalisierbar, man kann also zunächst komplexe und daraus reelle Lösungen berechnen. Die explizite Durchführung der Lösung von (*) überlassen wir dem Leser (Aufgabe 4 und Aufgabe 4 zu 4.4).

4.3.6. Gelegentlich tritt das Problem auf, daß man zwei oder mehrere Endomorphismen mit einer gemeinsamen Basis diagonalisieren will (man nennt das *simultane Diagonalisierung*). Daß dies nicht immer gehen kann, sieht man am schnellsten mit Matrizen: Gibt es zu $A, B \in M(n \times n; K)$ ein $S \in GL(n; K)$ mit

$$SAS^{-1} = D \quad \text{und} \quad SBS^{-1} = \tilde{D},$$

wobei D und $\tilde{D}$ Diagonalmatrizen bezeichnen, so folgt aus $D \cdot \tilde{D} = \tilde{D} \cdot D$, daß

$$B \cdot A = S^{-1} \tilde{D} S \cdot S^{-1} D S = S^{-1} D S \cdot S^{-1} \tilde{D} S = A \cdot B.$$

Erfreulicherweise ist die Vertauschbarkeit auch hinreichend. Es gilt also:

Satz. Zwei diagonalisierbare Endomorphismen $F, G \in \text{End}_K(V)$ sind genau dann simultan diagonalisierbar, wenn $F \circ G = G \circ F$.

Beweis. Nach Voraussetzung hat man die Zerlegungen in Eigenräume

$$\begin{aligned} V &= \text{Eig}(F; \lambda_1) \oplus \dots \oplus \text{Eig}(F; \lambda_k) \\ &= \text{Eig}(G; \mu_1) \oplus \dots \oplus \text{Eig}(G; \mu_l), \end{aligned}$$

wobei λ_i bzw. μ_j die Eigenwerte von F bzw. G bezeichnen. Wir halten ein λ fest und betrachten

$$W := \text{Eig}(F; \lambda).$$

Ist $w \in W$, so ist $F(G(w)) = G(F(w)) = G(\lambda w) = \lambda G(w)$, also ist W auch G -invariant (vgl. 4.4.1). Wir definieren

$$W_j := W \cap \text{Eig}(G; \mu_j) \quad \text{für } j = 1, \dots, l,$$

und es genügt zu zeigen, daß $W = W_1 \oplus \dots \oplus W_l$, da dies dann für alle Eigenwerte von F gilt. Wegen Lemma 4.1.3 ist nur zu zeigen, daß

$$W = W_1 + \dots + W_l.$$

Zu $w \in W$ gibt es $w_j \in \text{Eig}(G; \mu_j)$, so daß $w = w_1 + \dots + w_l$. Daraus folgt

$$\begin{aligned} F(w) &= F(w_1) + \dots + F(w_l) \\ &= \lambda w = \lambda w_1 + \dots + \lambda w_l, \end{aligned}$$

daher ist wegen der Eindeutigkeit in der direkten Summe

$$F(w_j) = \lambda w_j, \quad \text{also } w_j \in W \text{ und somit } w_j \in W_j. \quad \square$$

Aufgaben zu 4.3

1. Beweisen Sie Teil 2) von Satz 4.3.1 mit Hilfe von Theorem 4.3.3.

2. Sind die folgenden Matrizen diagonalisierbar?

$$\begin{pmatrix} 1 & 2 & 0 & 4 \\ 0 & 2 & 3 & 1 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}, \quad \begin{pmatrix} -5 & 0 & 7 \\ 6 & 2 & -6 \\ -4 & 0 & 6 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 & 2 \\ -2 & -2 & -6 \\ 1 & 2 & 5 \end{pmatrix}.$$

3. Für welche $a, b \in \mathbb{R}$ ist die Matrix

$$\begin{pmatrix} -3 & 0 & 0 \\ 2a & b & a \\ 10 & 0 & 2 \end{pmatrix}$$

diagonalisierbar?

4. Wir betrachten das Differentialgleichungssystem mit Anfangswertbedingung

$$\dot{y} = A \cdot y, \quad y_0(0) = \alpha, \quad y_1(0) = \beta \quad (*)$$

für die gedämpfte Schwingung (siehe 4.3.5), wobei

$$A = \begin{pmatrix} 0 & 1 \\ -\omega^2 & -2\mu \end{pmatrix}.$$

- Im Fall $\mu > \omega$ ist A (reell) diagonalisierbar. Bestimmen Sie eine Basis des $\mathbb{R}^2$ aus Eigenvektoren von A und geben Sie eine Basis des Lösungsraums von $\dot{y} = A \cdot y$ an (vgl. Aufgabe 3 zu 4.1). Wie sieht die Lösung von $(*)$ aus?
- Im Fall $\mu < \omega$ ist $A \in M(2 \times 2; \mathbb{C})$ komplex diagonalisierbar. Bestimmen Sie die Eigenwerte von A und geben Sie eine Basis des $\mathbb{C}^2$ aus Eigenvektoren von A an. Ist

$\lambda \in \mathbb{C}$ Eigenwert von A zum Eigenvektor $v \in \mathbb{C}^2$, so ist $re^{i\lambda t} \cdot v$, im $e^{i\lambda t} \cdot v$ eine Basis des Lösungsraums von $\dot{y} = A \cdot y$ ([Fo2], §13). Bestimmen Sie auch in diesem Fall die Lösung von (*).

5. Diagonalisieren Sie die Matrizen

$$A = \begin{pmatrix} -5 & 1 & 6 & 6 \\ -12 & 2 & 12 & 12 \\ 1 & 1 & 0 & -2 \\ -4 & 0 & 4 & 6 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 0 & -1 & -4 \\ -3 & 1 & 3 & 0 \\ 2 & 0 & -1 & -2 \\ 1 & 0 & -1 & -3 \end{pmatrix}$$

aus $M(4 \times 4; \mathbb{R})$ simultan, d. h. bestimmen Sie eine Matrix $S \in GL(4; \mathbb{R})$, so dass SAS^{-1} und SBS^{-1} Diagonalmatrizen sind.

6. Seien $A, B \in M(n \times n; K)$ mit $AB = BA$ und alle Eigenwerte von A und B seien einfach. Dann gilt: A und B haben die gleichen Eigenvektoren.

7. Zeigen Sie, dass es für $\lambda \in K$ und natürliche Zahlen μ, n mit $1 \leq \mu \leq n$ stets eine Matrix $A \in M(n \times n; K)$ gibt mit $\mu(P_A; \lambda) = \mu$ und $\dim \text{Eig}(A; \lambda) = 1$.

8. Es sei K ein Körper mit $\text{char } K \neq 2$. Zeigen Sie, dass die Lösungen der Gleichung $A^2 = E_2$ in $M(2 \times 2; K)$ genau von der folgenden Gestalt sind:

$$A = E_2, \quad A = -E_2 \quad \text{oder} \quad A = SDS^{-1} \quad \text{mit} \quad D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{und} \quad S \in GL(2; K).$$

9. Sei F ein diagonalisierbarer Endomorphismus eines endlichdimensionalen $\mathbb{R}$ -Vektorraums, für den gilt: Sind v und w Eigenvektoren von F , so ist $v + w$ ein Eigenvektor von F oder $v + w = 0$. Zeigen Sie, dass es ein $\lambda \in \mathbb{R}$ gibt mit $F = \lambda \cdot \text{id}$.

10. Seien $A, B \in M(3 \times 3; \mathbb{R})$ zwei Matrizen mit den charakteristischen Polynomen $P_A(t) = -t^3 + 2t^2 - t$ und $P_B(t) = -t^3 + 7t^2 - 9t + 3$. Zeigen Sie, dass der Kern von AB die Dimension 1 hat.

4.4 Trigonalisierung*

Wie wir in 4.3.3 gesehen hatten, gibt es zwei Bedingungen für die Diagonalisierbarkeit eines Endomorphismus:

- Das charakteristische Polynom muß in Linearfaktoren zerfallen, und
- die Dimensionen der Eigenräume müssen gleich den Vielfachheiten der Eigenwerte sein.

In diesem Abschnitt zeigen wir, daß ein Endomorphismus, der nur Bedingung a) erfüllt, wenigstens durch eine obere Dreiecksmatrix beschrieben werden kann.

4.4.1. Um die geometrischen Eigenschaften komplizierter Endomorphismen besser verstehen zu können, sind ein paar neue Begriffe nützlich.

Definition. Sei $F: V \rightarrow V$ ein Endomorphismus und $W \subset V$ ein Untervektorraum. W heißt *F-invariant*, wenn $F(W) \subset W$.

Es gibt stets die trivialen, aber nutzlosen Beispiele $W = \{0\}$ und $W = V$. Je mehr weitere invariante Unterräume existieren, desto übersichtlicher ist ein Endomorphismus. Ist etwa F diagonalisierbar und $(v_1, \dots, v_n)$ eine Basis aus Eigenvektoren, so ist

$$V = W_1 \oplus \dots \oplus W_n \quad \text{mit} \quad W_i = K \cdot v_i$$

eine Zerlegung in eindimensionale invariante Unterräume. Im Falle mehrfacher Eigenwerte (4.3.3) hat man eine weitere Zerlegung

$$V = \text{Eig}(F; \lambda_1) \oplus \dots \oplus \text{Eig}(F; \lambda_k)$$

in invariante Unterräume der Dimensionen $r_1, \dots, r_k$.

Die Beziehung zwischen invarianten Unterräumen und charakteristischem Polynom ist in einer Richtung ganz klar.

Bemerkung. Ist $W \subset V$ ein F -invarianter Unterraum, so ist $P_{F|W}$ ein Teiler von P_F .

Beweis. Wir ergänzen eine Basis $\mathcal{B}'$ von W zu einer Basis $\mathcal{B}$ von V . Dann ist

$$M_{\mathcal{B}}(F) = \left(\begin{array}{c|c} M_{\mathcal{B}'}(F|W) & * \\ \hline 0 & A \end{array} \right),$$

also $P_F = P_{F|W} \cdot P_A$ nach D9 in 3.1.3. □

Beispiel. In $\mathbb{Q}[t]$ hat das Polynom $t^n - 2$ für $n \geq 2$ keinen Teiler kleineren Grades (Aufgabe 1). Also hat der durch

$$A = \begin{pmatrix} 0 & \cdots & 0 & 2 \\ 1 & & 0 & 0 \\ & \ddots & & \vdots \\ 0 & & 1 & 0 \end{pmatrix}$$

beschriebene Endomorphismus des $\mathbb{Q}^n$ nur die trivialen invarianten Unterräume.

4.4.2. Sei $A \in M(n \times n; K)$ eine obere Dreiecksmatrix, d.h. $a_{ij} = 0$ für $i > j$. Unter dem Endomorphismus A des K^n sind die Untervektorräume

$$W_r := \text{span}(e_1, \dots, e_r)$$

für $1 \leq r \leq n$ invariant. Das motiviert die

Definition. Unter einer Fahne (V_r) in einem n -dimensionalen Vektorraum V versteht man eine Kette

$$\{0\} = V_0 \subset V_1 \subset \dots \subset V_n = V$$

von Untervektorräumen mit $\dim V_r = r$. Ist $F \in \text{End}(V)$, so heißt die Fahne F -invariant, wenn

$$F(V_r) \subset V_r \quad \text{für alle } r \in \{0, \dots, n\}.$$

Man kann sich V_0 als Befestigungspunkt, V_1 als Stange und V_2 als Tuch der Fahne vorstellen.

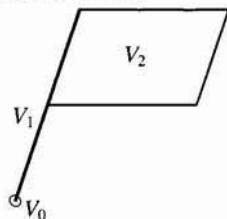


Bild 4.5

Mit Hilfe jeder Basis eines Vektorraumes kann man viele Fahnen konstruieren, aber aus der Existenz einer F -invarianten Fahne folgt insbesondere wegen $F(V_1) \subset V_1$, daß es einen Eigenvektor geben muß.

Bemerkung. Für $F \in \text{End}(V)$ sind folgende Bedingungen gleichwertig:

- i) Es gibt eine F -invariante Fahne in V .
- ii) Es gibt eine Basis $\mathcal{B}$ von V , so daß $M_{\mathcal{B}}(F)$ obere Dreiecksmatrix ist.

Ist das der Fall, so heißt F trigonalisierbar.

Beweis. Die Beziehung zwischen Fahne (V_r) und Basis $\mathcal{B} = (v_1, \dots, v_n)$ ist geregelt durch

$$V_r = \text{span}(v_1, \dots, v_r) \quad \text{für alle } r. \quad \square$$

Übersetzt in den Matrizenkalkül ergibt das die

Definition. $A \in M(n \times n; K)$ heißt trigonalisierbar, wenn es ein $S \in \text{GL}(n; K)$ gibt, so daß SAS^{-1} obere Dreiecksmatrix ist.

4.4.3. Ergebnis dieses Abschnittes ist der

Trigonalisierungssatz. Für einen Endomorphismus F eines n -dimensionalen K -Vektorraumes sind folgende Bedingungen äquivalent:

- i) F ist trigonalisierbar.
- ii) Das charakteristische Polynom P_F zerfällt in Linearfaktoren, d.h.

$$P_F = \pm(t - \lambda_1) \cdot \dots \cdot (t - \lambda_n) \quad \text{mit } \lambda_1, \dots, \lambda_n \in K.$$

Mit Hilfe des Fundamentalsatzes der Algebra (1.3.9) folgt das

Korollar. Jeder Endomorphismus eines endlich-dimensionalen komplexen Vektorraumes ist trigonalisierbar. $\square$

Eine weitere wichtige Anwendung betrifft die Lösung linearer Systeme von Differentialgleichungen ([Fo2], §14).

Beweis des Satzes. i) $\Rightarrow$ ii) ist klar, denn ist $A = (a_{ij}) = M_{\mathcal{B}}(F)$ eine obere Dreiecksmatrix, so folgt aus D8 in 3.1.3, daß

$$P_F = (a_{11} - t) \cdot \dots \cdot (a_{nn} - t).$$

ii) $\Rightarrow$ i) wird durch Induktion über $n = \dim V$ bewiesen. Für $n = 0, 1$ ist nichts zu zeigen. Ist $n \geq 2$, so wähle man einen Eigenvektor v_1 zu λ_1 und ergänze ihn zu einer Basis $\mathcal{B} = (v_1, w_2, \dots, w_n)$ von V . Dann ist

$$V = V_1 \oplus W \quad \text{mit } V_1 := \text{span}(v_1) \text{ und } W := \text{span}(w_2, \dots, w_n),$$

und

$$M_B(F) = \begin{pmatrix} \lambda_1 & a_{12} & \cdots & a_{1n} \\ 0 & \boxed{} & & \\ \vdots & & \ddots & \\ 0 & & & \end{pmatrix}.$$

V_1 ist F -invariant, W im allgemeinen nicht; das Hindernis dagegen sind die Einträge $a_{12}, \dots, a_{1n}$. Der Ausweg ist folgender: Durch

$$H(w_j) = a_{1j}v_1 \quad \text{und} \quad G(w_j) = a_{2j}w_2 + \dots + a_{nj}w_n$$

sind lineare Abbildungen $H: W \rightarrow V_1$ und $G: W \rightarrow W$ erklärt mit

$$F(w) = H(w) + G(w) \quad \text{für alle } w \in W.$$

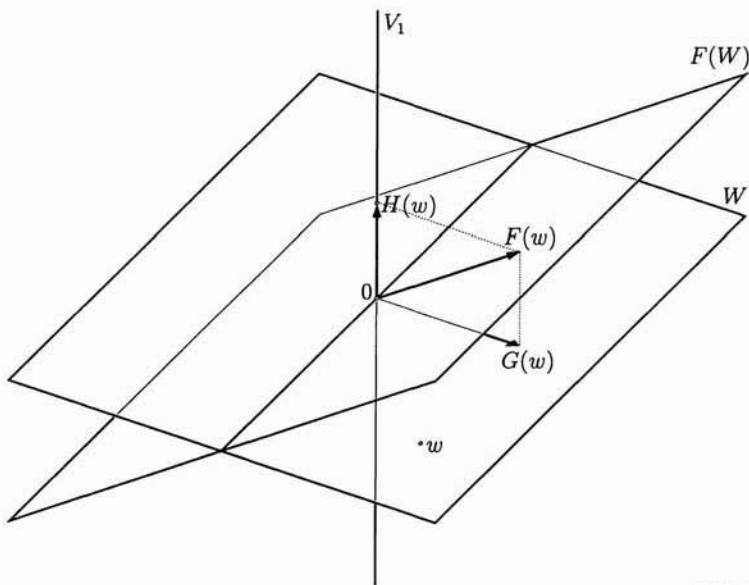


Bild 4.6

Für die charakteristischen Polynome gilt

$$P_F = (\lambda_1 - t) \cdot P_G, \quad \text{also} \quad P_G = (\lambda_2 - t) \cdot \dots \cdot (\lambda_n - t).$$

Die Induktionsvoraussetzung angewandt auf W ergibt eine G -invariante Fahne

$$\{0\} = W_0 \subset \dots \subset W_{n-1} = W,$$

und wir behaupten, daß durch $V_r := V_1 + W_{r-1}$ eine F -invariante Fahne gegeben ist. Das folgt aus

$$F(\mu v_1 + w) = \lambda_1 \mu v_1 + H(w) + G(w)$$

wegen $H(w) \in V_1$ und $G(w) \in W_{r-1}$ für $w \in W_{r-1}$. $\square$

4.4.4. Obwohl obiger Satz sehr nützlich ist, muß man doch bemerken, daß die Beweismethode nicht genügend sorgfältig auf die Geometrie der Abbildung achtet.

Beispiel. Ist $F \in \text{End}(\mathbb{R}^2)$ gegeben durch die Diagonalmatrix

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

so kann man mit $v_1 = e_1$ und $\lambda_1 = 1$ beginnen. Die Ergänzung durch w_2 ist beliebig: wählt man etwa $w_2 = (-1, 1)$, so wird

$$A = M_B(F) = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix},$$

und damit ist die schöne Diagonalmatrix versaut. Auch vom algebraischen Standpunkt hat A eine unangenehme Eigenschaft. Es ist

$$A = D + N \quad \text{mit} \quad N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Dabei ist N nilpotent (vgl. 4.5.7) mit $N^2 = 0$, aber $D \cdot N \neq N \cdot D$.

In 4.6 wird gezeigt, wie man den nilpotenten Anteil N auf eine der Abbildung angepaßte Normalform bringen kann.

4.4.5. Wir geben nun ein *Rechenverfahren zur Trigonalisierung eines Endomorphismus* an. Es genügt dazu, eine Matrix $A \in M(n \times n; K)$ zu betrachten, für die

$$P_A = \pm(t - \lambda_1) \cdot \dots \cdot (t - \lambda_n) \quad \text{mit} \quad \lambda_1, \dots, \lambda_n \in K.$$

Gesucht ist eine Matrix $S \in \text{GL}(n; K)$, so daß

$$D := S \cdot A \cdot S^{-1}$$

eine obere Dreiecksmatrix ist. Der Beweis des Trigonalisierungssatzes 4.4.3 ergibt das folgende Iterationsverfahren.

1. Schritt: Wir betrachten $W_1 = K^n$ mit der Basis $\mathcal{B}_1 = \mathcal{K}$ und dem Endomorphismus $A_1 = A$. Zu λ_1 berechnet man einen Eigenvektor $v_1 \in K^n$. Nach dem Austauschlemma 1.5.4 bestimmt man ein $j_1 \in \{1, \dots, n\}$, so daß

$$\mathcal{B}_2 := (v_1, e_1, \dots, \widehat{e}_{j_1}, \dots, e_n)$$

wieder eine Basis von K^n ist. Das Zeichen $\widehat{}$ bedeutet dabei, daß e_{j_1} ausgelassen wird. Wir betrachten die Transformationsmatrix

$$S_1^{-1} := T_{\mathcal{B}_1}^{\mathcal{B}_2}$$

mit der Basis $\mathcal{B}_2$ als Spalten. Dann ist

$$A_2 := S_1 \cdot A \cdot S_1^{-1} = \begin{pmatrix} \lambda_1 & * & \dots & * \\ 0 & \boxed{A'_2} \\ \vdots & & & \\ 0 & & & \end{pmatrix}.$$

2. Schritt: Wir betrachten W_2 mit der Basis $\mathcal{B}'_2 = (e_1, \dots, \widehat{e}_{j_1}, \dots, e_n)$ und dem Endomorphismus A'_2 . Es ist

$$P_{A'_2} = \pm(t - \lambda_2) \cdot \dots \cdot (t - \lambda_n).$$

Zu λ_2 berechnet man einen Eigenvektor $v_2 \in W_2$, und man wählt ein $j_2 \neq j_1$, so daß

$$\mathcal{B}'_3 = (v_2, e_1, \dots, \widehat{e}_{j_1}, \dots, \widehat{e}_{j_2}, \dots, e_n)$$

eine Basis von W_2 , also

$$\mathcal{B}_3 = (v_1, v_2, e_1, \dots, \widehat{e}_{j_1}, \dots, \widehat{e}_{j_2}, \dots, e_n)$$

eine Basis von K^n ist. Mit der Transformationsmatrix $S_2^{-1} = T_{\mathcal{B}_1}^{\mathcal{B}_3}$ erhält man

$$A_3 = S_2 \cdot A \cdot S_2^{-1} = \begin{pmatrix} \lambda_1 & * & \dots & \dots & * \\ 0 & \lambda_2 & * & \dots & * \\ \vdots & 0 & \boxed{A'_3} \\ \vdots & \vdots & & & \\ 0 & 0 & & & \end{pmatrix}.$$

Bei der Berechnung von S_2 kann man benutzen, daß

$$T_{\mathcal{B}_1}^{\mathcal{B}_3} = T_{\mathcal{B}_1}^{\mathcal{B}_2} \cdot T_{\mathcal{B}_2}^{\mathcal{B}_3} \quad \text{und} \quad T_{\mathcal{B}_2}^{\mathcal{B}_3} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \boxed{T_{\mathcal{B}'_2}^{\mathcal{B}'_3}} \\ \vdots & & & \\ 0 & & & \end{pmatrix}.$$

Spätestens im $(n-1)$ -ten Schritt erhält man eine obere Dreiecksmatrix A_n , denn A'_n ist eine (1×1) -Matrix. Also ist

$$D := A_n = S_{n-1} \cdot A \cdot S_{n-1}^{-1}$$

eine obere Dreiecksmatrix.

Vorsicht! Die im ersten Schritt berechnete erste Zeile von A_2 kann sich in jedem der folgenden Schritte wieder ändern, analog für die anderen Zeilen. Das ist nicht angenehm.

Beispiel. Sei $K = \mathbb{R}$, $n = 3$ und

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

Dann ist $P_A = -(t-2)^3$, also $\lambda_1 = \lambda_2 = \lambda_3 = 2$. Wegen $\text{rang}(A - 2E_3) = 2$ ist

$$\dim \text{Eig}(A; 2) = 1 < 3 = \mu(P_A; 2),$$

also ist A nicht diagonalisierbar. Die Iteration läuft wie folgt:

1. Schritt: $v_1 = (1, -1, 1)$, $j_1 = 1$, $B_2 = (v_1, e_2, e_3)$,

$$S_1^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \quad S_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix},$$

$$A_2 = S_1 \cdot A \cdot S_1^{-1} = \begin{pmatrix} 2 & 4 & 3 \\ 0 & \boxed{4} & 2 \\ 0 & -2 & 0 \end{pmatrix}.$$

2. Schritt: $W_2 = \text{span}(e_2, e_3)$, $v_2 = e_2 - e_3$, $j_2 = 2$,

$$S_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 2 & 1 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & \boxed{2} \end{pmatrix} = D.$$

Also ist $S = S_2$.

Aufgaben zu 4.4

1. Zeigen Sie, dass das Polynom $t^n - 2 \in \mathbb{Q}[t]$ für $n \geq 2$ keinen Teiler $P \in \mathbb{Q}[t]$ mit $1 \leq \deg P \leq n-1$ besitzt.

2. Trigonalisieren Sie mit dem Verfahren aus 4.4.5 die Matrizen

$$\begin{pmatrix} 3 & 0 & -2 \\ -2 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} -1 & -3 & -4 \\ -1 & 0 & 3 \\ 1 & -2 & -5 \end{pmatrix}.$$

3. Zeigen Sie mit Induktion nach $n = \dim V$: Ist V ein endlichdimensionaler K -Vektorraum und $F: V \rightarrow V$ ein nilpotenter Endomorphismus, so existiert eine Basis $\mathcal{B}$ von V mit

$$M_{\mathcal{B}}(F) = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix},$$

und es gilt $P_F(t) = \pm t^n$.

4. (Fortsetzung von Aufgabe 4 in 4.3.) Zeigen Sie, dass die Matrix

$$A = \begin{pmatrix} 0 & 1 \\ -\omega^2 & -2\mu \end{pmatrix}$$

im Fall $\mu = \omega$ trigonalisierbar ist, und bestimmen Sie eine Matrix $S \in GL(2; \mathbb{R})$, so dass $B = SAS^{-1}$ obere Dreiecksmatrix ist. Das System $\dot{y} = A \cdot y$ geht somit durch die Substitution $z = Sy$ über in $\dot{z} = B \cdot z$, und es reicht, das (einfachere) System $\dot{z} = B \cdot z$ zu lösen. Bestimmen Sie auf diese Weise eine Basis des Lösungsraums von $\dot{y} = A \cdot y$ und lösen (*) in 4.3.5 auch im aperiodischen Grenzfall.

4.5 Potenzen eines Endomorphismus*

Zur genauen Beschreibung eines Endomorphismus – insbesondere bei der Suche nach invarianten Unterräumen – muß man systematisch das Ergebnis der wiederholten Anwendung der Abbildung betrachten. Das legt die Weiterentwicklung der algebraischen Hilfsmittel nahe. Das wichtigste Ergebnis ist der Satz von CAYLEY-HAMILTON.

4.5.1. Die Eigenwertbedingung $F(v) = \lambda v$ kann man auch komplizierter ausdrücken: in dem linearen Polynom

$$L(t) = t - \lambda$$

setzt man F statt t und id_V statt 1 , das ergibt

$$L(F) := F - \lambda \text{id}_V,$$

und es ist $\text{Eig}(F; \lambda) = \text{Ker } L(F)$. Was hat man damit gewonnen? Ist

$$P(t) = \alpha_r t^r + \dots + \alpha_1 t + \alpha_0 \in K[t]$$

ein beliebiges Polynom, so kann man analog

$$P(F) := \alpha_r F^r + \dots + \alpha_1 F + \alpha_0 \text{id}_V \in \text{End}(V)$$

setzen. Ist $0 \neq v \in \text{Ker } P(F)$, so gilt

$$\alpha_r F^r(v) + \dots + \alpha_1 F(v) + \alpha_0 v = 0,$$

d.h. die Vektoren $v, F(v), \dots, F^r(v)$ sind linear abhängig, wenn $P \neq 0$. Ist $\alpha_r \neq 0$, so folgt, daß

$$W = \text{span}(v, F(v), \dots, F^{r-1}(v))$$

ein F -invarianter Unterraum von V mit $1 \leq \dim W \leq r$ ist. Ist $P(F) = 0$, so kann man diese Konstruktion mit jedem beliebigen v starten.

Beispiel. Sei F diagonalisierbar mit den paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_k$ und

$$Q(t) = (t - \lambda_1) \cdot \dots \cdot (t - \lambda_k).$$

Jedes $v \in V$ hat nach 4.3.3 eine Darstellung $v = v_1 + \dots + v_k$ mit $v_i \in \text{Eig}(F; \lambda_i)$. Benutzt man, daß für $\lambda \neq \mu$ und $0 \neq v \in \text{Eig}(F; \lambda)$

$$0 \neq (F - \mu \text{id}_V)(v) = (\lambda - \mu)v \in \text{Eig}(F; \lambda),$$

so folgt sofort $Q(F) = 0$ und $Q'(F) \neq 0$ für jeden echten Teiler Q' von Q . Weiter folgt $P_F(F) = 0$, wenn

$$P_F = (t - \lambda_1)^{r_1} \cdot \dots \cdot (t - \lambda_k)^{r_k}$$

das charakteristische Polynom bezeichnet.

4.5.2. Das Einsetzen eines Endomorphismus in Polynome ist beschrieben durch die Abbildung

$$\Phi_F: K[t] \rightarrow \text{End}(V), \quad P(t) \mapsto P(F).$$

Dies ist ein Homomorphismus von Ringen und auch von K -Vektorräumen. Das Bild

$$K[F] = \{P(F): P(t) \in K[t]\} \subset \text{End } V$$

ist ein kommutativer Unterring des (nicht kommutativen) Ringes $\text{End } V$, und der Kern

$$\mathcal{I}_F := \{P(t) \in K[t]: P(F) = 0\} \subset K[t]$$

heißt *Ideal von F* . Wegen $\dim K[t] = \infty$ und $\dim \text{End}(V) < \infty$ gibt es in $\mathcal{I}_F$ ein normiertes Polynom vom Grad $\leq n^2$ (vgl. Aufgabe 3). Ist F diagonalisierbar, so hatten wir in 4.5.1 sogar ein Polynom $Q \in \mathcal{I}_F$ vom Grad $k \leq n$ angegeben.

4.5.3. Wir zeigen nun, daß sich die Schranke der Grade in $\mathcal{I}_F$ für alle Endomorphismen F auf $n = \dim V$ verbessern läßt. Es gilt sogar noch mehr:

Satz von CAYLEY-HAMILTON. *Sei V ein endlichdimensionaler K -Vektorraum, $F \in \text{End}(V)$ und $P_F \in K[t]$ sein charakteristisches Polynom. Dann ist*

$$P_F(F) = 0 \in \text{End}(V).$$

Insbesondere gilt für jede Matrix $A \in M(n \times n; K)$

$$P_A(A) = 0 \in M(n \times n; K).$$

Vorsicht! Es ist entscheidend, daß die Null ein Endomorphismus oder eine Matrix, nicht die Null in K ist. Daher ist der naheliegende „Beweis“

$$P_A(A) = \det(A - A \cdot E_n) = \det(A - A) = \det(0) = 0$$

nicht richtig. Am Ende steht die $0 \in K$, der Fehler steht schon in der ersten Gleichung, weil ganz naiv bei einem Eintrag der Determinante $t = A$ gesetzt wurde, dann aber mit Einträgen aus K weitergerechnet wird.

Mit einem Trick von SERGE LANG [L] kann man diesen falschen Beweis zu einem richtigen aufmöbeln. Das Problem dabei ist folgendes: die Einträge einer Determinante müssen aus einem kommutativen Ring stammen (3.2.8), die Ringe quadratischer Matrizen sind i.a. nicht kommutativ. Der Ausweg ist, daß man sich auf den kommutativen Unterring (4.5.2)

$$K[A] \subset M(n \times n; K)$$

beschränken kann. Nach diesen Vorbemerkungen zum

Beweis. Man kann den Endomorphismus F von V mit Hilfe einer Basis durch eine Matrix $A \in M(n \times n; K)$ beschreiben. Also genügt es, den Satz dafür zu beweisen.

Für $n = 1$ ist die Aussage trivial, für $n = 2$ kann man sie noch durch direkte Rechnung mit Matrizen beweisen. Im allgemeinen Fall benutzen wir, wie oben erläutert, Matrizen mit Einträgen aus den kommutativen Ringen $K[t]$ und $K[A]$. Zunächst sei

$$B(t) := {}^t(A - tE_n) \in M(n \times n; K[t]).$$

Die Einträge von $B(t)$ außerhalb der Diagonalen sind Elemente von K , und in der Diagonalen stehen lineare Polynome, es ist

$$\det B(t) = P_A(t) \in K[t]. \quad (*)$$

Nun wird die Unbestimmte t durch die Matrix A und jeder Eintrag a_{ij} durch $a_{ij}E_n$ ersetzt. Das ergibt

$$B(A) = \begin{pmatrix} a_{11}E_n - A & a_{21}E_n & \cdots & a_{n1}E_n \\ \vdots & \vdots & & \vdots \\ a_{1n}E_n & a_{2n}E_n & \cdots & a_{nn}E_n - A \end{pmatrix} \in M(n \times n; K[A]).$$

Diese Matrix kann mit einer Spalte multipliziert werden, deren Einträge Spaltenvektoren des K^n sind. Insbesondere ist

$$B(A) \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} a_{11}e_1 - Ae_1 + a_{21}e_2 + \dots + a_{n1}e_n \\ \vdots \\ a_{1n}e_1 + a_{2n}e_2 + \dots + a_{nn}e_n - Ae_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}. \quad (**)$$

Nun sei $B^z(t) \in M(n \times n; K[t])$ die in 3.3.1 erklärte, zu $B(t)$ komplementäre Matrix. Ihre Einträge sind entsprechend der Definition Polynome vom Grad $\leq n - 1$, und es gilt wegen $(*)$

$$B^z(t) \cdot B(t) = P_A(t) \cdot E_n.$$

Setzt man auf beiden Seiten A für t ein, so folgt mit Hilfe von $(**)$

$$\begin{pmatrix} P_A(A) & & 0 \\ & \ddots & \\ 0 & & P_A(A) \end{pmatrix} \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} P_A(A)e_1 \\ \vdots \\ P_A(A)e_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Da $P_A(A)$ auf einer Basis die Werte $0 \in K^n$ hat, ist $P_A(A) = 0$ als Matrix. $\square$

Dieser Beweis ist extrem kurz und elegant, aber für einen Geometer mag er wie algebraischer Hokuspokus aussehen. Daher noch ein anderer

Beweis, bei dem man vielleicht besser ein Gefühl dafür bekommen kann, warum die Kombination $P_F(F)$ von Potenzen von F alles in V zu Null macht.

Zunächst betrachten wir den Spezialfall $K = \mathbb{C}$. Nach 4.4.3 ist F trigonalisierbar, d.h. es gibt eine F -invariante Fahne

$$\{0\} = V_0 \subset V_1 \subset \dots \subset V_n = V.$$

Ist $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V mit $V_i = \text{span}(v_1, \dots, v_i)$ für $i = 1, \dots, n$, so ist

$$M_{\mathcal{B}}(F) = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

wobei $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ die Eigenwerte von F sind. Da

$$P_F = (\lambda_1 - t) \cdot \dots \cdot (\lambda_n - t),$$

kann man $P_F(F)$ schrittweise aufbauen: Ist

$$G_i := (\lambda_1 \text{id}_V - F) \circ (\lambda_2 \text{id}_V - F) \circ \dots \circ (\lambda_i \text{id}_V - F) \in \text{End}(V)$$

für $i = 1, \dots, n$, so ist $G_n = P_F(F)$. Es genügt also durch Induktion zu zeigen, daß

$$G_i(V_i) = \{0\} \quad \text{für } i = 1, \dots, n.$$

Der Fall $i = 1$ ist klar, da $F(v_1) = \lambda_1 v_1$. Sei nun $i \geq 2$ und $v \in V_i$. Dann gibt es ein $w \in V_{i-1}$ und ein $\mu \in \mathbb{C}$, so daß $v = w + \mu v_i$. Da

$$\lambda_i w - F(w) \in V_{i-1} \quad \text{und} \quad \lambda_i v_i - F(v_i) \in V_{i-1},$$

ist nach Induktionsvoraussetzung

$$G_i(w) = (G_{i-1} \circ (\lambda_i \text{id}_V - F))(w) = G_{i-1}(\lambda_i w - F(w)) = 0$$

und

$$G_i(v_i) = (G_{i-1} \circ (\lambda_i \text{id}_V - F))(v_i) = G_{i-1}(\lambda_i v_i - F(v_i)) = 0.$$

Also ist $G_i(v) = G_i(w) + \mu G_i(v_i) = 0$.

Der Leser möchte sich überlegen, wie viel einfacher der Beweis wird, wenn F diagonalisierbar ist, d.h. wenn $F(v_i) = \lambda_i v_i$ (vgl. Aufgabe 5).

Den Fall $K = \mathbb{R}$ kann man durch *Komplexifizierung* auf den obigen Fall zurückführen, obwohl F nicht reell trigonalisierbar sein muß. Am einfachsten geht das mit Hilfe einer Matrix

$$A = M_{\mathcal{B}}(F),$$

wobei $\mathcal{B}$ irgend eine Basis von V ist (die „professionelle“ Methode der Komplexifizierung ist in 6.3.4 b) ausgeführt). A beschreibt einen Endomorphismus

$$A: \mathbb{C}^n \rightarrow \mathbb{C}^n,$$

also ist $P_A(A) = 0 \in \text{End } \mathbb{C}^n$. Nach 2.6.4 ist

$$M_B(P_F(F)) = P_F(M_B(F)) = P_A(A),$$

also auch $P_F(F) = 0 \in \text{End } V$.

Diese Methode ist nicht auf $\mathbb{R}$ und $\mathbb{C}$ beschränkt, denn zu jedem Körper K und jedem Polynom $f \in K[t]$ gibt es einen Oberkörper $K' \supset K$, so daß $f \in K'[t]$ in Linearfaktoren zerfällt: etwa den *Zerfällungskörper* von f oder den *algebraischen Abschluß* von K (siehe z.B. [F-S]).

4.5.4. Als Beispiel für eine Anwendung des Satzes von CAYLEY-HAMILTON wollen wir in Analogie zu Korollar 4.4.3 zeigen, daß sich jeder Endomorphismus eines reellen Vektorraumes beinahe trigonalisieren läßt.

Theorem. Ist F Endomorphismus eines reellen Vektorraumes V , so gibt es eine Basis B derart, daß

$$M_B(F) = \begin{pmatrix} \lambda_1 & & & & \\ & \ddots & & & \\ & & \lambda_r & & \\ & & & \boxed{B_1} & \\ & 0 & & & \ddots \\ & & & & & \boxed{B_m} \end{pmatrix}$$

mit $\lambda_1, \dots, \lambda_r \in \mathbb{R}$ und $B_i = \begin{pmatrix} 0 & -c_i \\ 1 & -b_i \end{pmatrix}$ mit $b_i, c_i \in \mathbb{R}$ und $b_i^2 - 4c_i < 0$ für $i = 1, \dots, m$.

Lemma. Jeder Endomorphismus F eines reellen Vektorraumes V mit $\dim V \geq 1$ hat einen invarianten Unterraum W mit $1 \leq \dim W \leq 2$.

Beweis des Lemmas. Nach 1.3.10 gibt es eine Zerlegung

$$P_F = \pm(t - \lambda_1) \cdot \dots \cdot (t - \lambda_r) \cdot Q_1(t) \cdot \dots \cdot Q_m(t)$$

mit $Q_i(t) = t^2 + b_i t + c_i$ und $b_i^2 - 4c_i < 0$. Ist $r \geq 1$, so gibt es einen Eigenvektor v_1 zu λ_1 , und $W = \mathbb{R} \cdot v_1$ ist invariant.

Im Fall $r = 0$ starten wir mit einem beliebigen $0 \neq v \in V$. Nach CAYLEY-HAMILTON ist

$$P_F(F)(v) = (Q_1(F) \circ \dots \circ Q_m(F))(v) = 0.$$

Also gibt es ein eindeutig bestimmtes $i \in \{1, \dots, m\}$ mit

$$w := (Q_{i+1}(F) \circ \dots \circ Q_m(F))(v) \neq 0 \quad \text{und} \quad Q_i(F)(w) = 0,$$

wobei $w := v$ gesetzt wird, falls $i = m$, d.h. $Q_m(F)(v) = 0$. Wir setzen

$$W := \text{span}(w, F(w)).$$

Aus $Q_i(F)(w) = 0$ folgt

$$F(F(w)) + b_i F(w) + c_i w = 0,$$

also ist $F(W) \subset W$. Da F keinen Eigenwert hat, ist $\mathcal{B} = (w, F(w))$ eine Basis von W , und es gilt

$$M_{\mathcal{B}}(F|_W) = \begin{pmatrix} 0 & -c_i \\ 1 & -b_i \end{pmatrix},$$

also $P_{F|_W} = Q_i(t)$. □

Den *Beweis des Theorems* führt man schrittweise nach dem Muster des Induktionsbeweises von Satz 4.4.3. In der ersten Etappe baut man die Eigenräume ab, in der zweiten die zweidimensionalen invarianten Unterräume. Das ergibt eine Folge von Zerlegungen

$$\begin{aligned} V &= V_0 \oplus W_1 = V_1 \oplus W_2 = V_2 \oplus W_3 = \dots = V_r \oplus W_{r+1} \\ &= V_{r+2} \oplus W_{r+3} = V_{r+4} \oplus W_{r+5} = \dots = V_{r+2m} \oplus W_{r+2m+1}, \end{aligned}$$

wobei $W_1 = V$, $W_{r+2m+1} = \{0\}$ und

$$\{0\} = V_0 \subset V_1 \subset \dots \subset V_r \subset V_{r+2} \subset \dots \subset V_{r+2m} = V$$

mit $\dim V_j = j$ und $F(V_j) \subset V_j$. Dazu verwendet man die eindeutige Zerlegung von

$$F|_{W_i}: W_i \rightarrow V = V_{i-1} \oplus W_i \quad \text{in} \quad F|_{W_i}: H_i + F_i \quad \text{mit}$$

$$H_i: W_i \rightarrow V_{i-1} \quad \text{und} \quad F_i: W_i \rightarrow W_i.$$

Die Einzelheiten seien getrost dem Leser überlassen. □

4.5.5. Nun wollen wir das in 4.5.2 begonnene und durch den Satz von CAYLEY-HAMILTON unterbrochene Spiel mit den Polynomen noch etwas weiter treiben.

Definition. Eine Teilmenge $\mathcal{I}$ eines kommutativen Ringes R heißt *Ideal*, wenn gilt

$$\text{I 1 } P, Q \in \mathcal{I} \Rightarrow P - Q \in \mathcal{I},$$

$$\text{I 2 } P \in \mathcal{I}, Q \in R \Rightarrow Q \cdot P \in \mathcal{I}.$$

Ein Beispiel dafür ist unser Ideal $\mathcal{I}_F \subset K[t]$.

Satz. Zu jedem Ideal $\mathcal{I} \subset K[t]$ mit $\mathcal{I} \neq \{0\}$ gibt es ein eindeutiges Polynom M mit folgenden Eigenschaften:

- 1) M ist normiert, d.h. $M = t^d + \dots$, wobei $d = \deg M$.
- 2) Für jedes $P \in \mathcal{I}$ gibt es ein $Q \in K[t]$ mit $P = Q \cdot M$.

M heißt Minimalpolynom von $\mathcal{I}$, im Fall $\mathcal{I} = \mathcal{I}_F$ Minimalpolynom von F .

Beweis. Sei $d := \min \{r : \text{es gibt ein } P \in \mathcal{I} \text{ mit } P \neq 0 \text{ und } r = \deg P\} \in \mathbb{N}$. Wir nehmen ein normiertes $M \in \mathcal{I}$ mit $\deg M = d$. Ein beliebiges $P \in \mathcal{I}$ dividieren wir entsprechend 1.3.7 mit Rest durch M :

$$P = Q \cdot M + R,$$

wobei $R = 0$ oder $\deg R < d$. Es genügt, den Fall $R \neq 0$ auszuschließen. Dann wäre

$$R = P - Q \cdot M \in \mathcal{I}$$

im Widerspruch zur Minimalität von d . Die Eindeutigkeit ist klar. $\square$

4.5.6. Es ist in der Geometrie eines Endomorphismus begründet, wie weit sich das Minimalpolynom vom charakteristischen Polynom unterscheidet.

Beispiel. Sei

$$A = \begin{pmatrix} 0 & 1 & & & & \\ & \ddots & \ddots & & & \\ & & \ddots & 1 & & \\ & & & \ddots & \ddots & \\ & 0 & & & \ddots & 0 \\ & & & & & \ddots \\ & & & & & & 0 \end{pmatrix} \in M(n \times n; K),$$

wobei $d - 1$ mit $1 \leq d \leq n$ die Anzahl der Einsen in der Nebendiagonalen bezeichnet. Das bedeutet

$$A(e_1) = 0, \quad A(e_2) = e_1, \dots, \quad A(e_d) = e_{d-1} \quad \text{und} \quad A(e_i) = 0 \quad \text{sonst.}$$

Für die Potenzen A^l mit $1 \leq l \leq d$ gilt $A^l(e_d) = e_{d-l}$, also ist $A^d = 0$, aber $A^{d-1} \neq 0$. Daraus folgt

$$P_A = \pm t^n, \quad M_A = t^d.$$

Im Extremfall $d = 1$ ist $A = 0$ und $P_A = \pm M_A^n$.

Allgemein gilt folgende Beziehung zwischen charakteristischem Polynom P_F und Minimalpolynom M_F :

Satz. Sei V ein n -dimensionaler K -Vektorraum, $F \in \text{End}(V)$. Dann gilt:

1) M_F teilt P_F .

2) P_F teilt M_F^n .

Beweis. 1) folgt sofort aus dem Satz von CAYLEY-HAMILTON. Zum Beweis von 2) wollen wir voraussetzen, daß K Unterkörper von $\mathbb{C}$ ist. Im allgemeinen Fall kann man einen *algebraischen Abschluß* $\bar{K}$ verwenden, aber so weit wollen wir nicht in die Algebra ausschweifen (vgl. [F-S]).

Wir können F ersetzen durch $A \in M(n \times n; K)$. Die Polynome M_A und P_A zerfallen über $\mathbb{C}$ in Linearfaktoren (vgl. 1.3.9). Zunächst betrachten wir

$$P_A = \pm(t - \lambda_1)^{r_1} \cdot \dots \cdot (t - \lambda_k)^{r_k},$$

wobei $\lambda_1, \dots, \lambda_k \in \mathbb{C}$ paarweise verschieden sind. Aus Teil 1) folgt

$$M_A = (t - \lambda_1)^{s_1} \cdot \dots \cdot (t - \lambda_k)^{s_k} \quad \text{mit} \quad 0 \leq s_i \leq r_i \quad \text{für} \quad i = 1, \dots, k,$$

d.h. M_A hat keine weiteren Linearfaktoren. Wir zeigen, daß keiner fehlt, d.h. daß $s_i \geq 1$ ist für alle i . Angenommen $s_i = 0$ für ein i . Dann nehmen wir einen Eigenvektor $v_i \in V$ zu λ_i und wenden $M_A(A)$ darauf an. Mit der gleichen Rechnung wie im Beispiel 4.5.1 folgt

$$M_A(A)(v_i) \neq 0$$

im Widerspruch zu $M_A(A) = 0$. Nun definieren wir

$$Q := (t - \lambda_1)^{ns_1 - r_1} \cdot \dots \cdot (t - \lambda_k)^{ns_k - r_k} \in \mathbb{C}[t].$$

Nach Definition von Q gilt $M_A^n = \pm Q \cdot P_A$, und da $M_A^n, P_A \in K[t]$ folgt $Q \in K[t]$ (Aufgabe 5 zu 1.3). $\square$

4.5.7. Manchmal kann das Minimalpolynom helfen, Beweise etwas abstrakter und kürzer zu machen. Dafür geben wir ein naheliegendes Beispiel.

Definition. Man nennt $F \in \text{End}_K(V)$ *nilpotent*, wenn $F^k = 0$ für ein $k \geq 1$.

Satz. Ist $F \in \text{End}_K(V)$ und $n = \dim V$, so sind folgende Aussagen äquivalent:

i) F ist nilpotent.

ii) $F^d = 0$ für ein d mit $1 \leq d \leq n$.

iii) $P_F = \pm t^n$.

iv) Es gibt eine Basis B von V , so daß

$$M_B(F) = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}.$$

Beweis. i) bedeutet $t^k \in \mathcal{I}_F$. Also ist $M_F = t^d$ mit $1 \leq d \leq n$ und $P_F = \pm t^n$. Das zeigt i) $\Rightarrow$ ii) und iii). Der Trigonalisierungssatz 4.4.3 enthält iii) $\Rightarrow$ iv). Schließlich folgt iv) $\Rightarrow$ i) nach Aufgabe 2 zu 2.5. $\square$

Aufgabe 3 zu 4.4 zeigt, daß es auch ohne Minimalpolynom geht.

Aufgaben zu 4.5

1. Sei $F: V \rightarrow V$ linear und $P \in K[t]$. Zeigen Sie: Ist $\lambda \in K$ ein Eigenwert von F , so ist $P(\lambda)$ ein Eigenwert von $P(F)$.

2. Ist $F: V \rightarrow V$ linear und $P, Q \in K[t]$, so ist

$$P(F) \circ Q(F) = Q(F) \circ P(F) = (P \cdot Q)(F).$$

3. Sei $F: V \rightarrow V$ ein Endomorphismus des K -Vektorraums V . Dann gilt:

- Die Abbildung $\Phi_F: K[t] \rightarrow \text{End}(V)$, $P(t) \mapsto P(F)$ ist ein Homomorphismus von Ringen und von K -Vektorräumen.
- $K[F] = \{P(F) : P \in K[t]\}$ ist ein kommutativer Unterring von $\text{End}(V)$.
- Ist $\dim V = n < \infty$, so existiert ein normiertes Polynom $P \in K[t]$ vom Grad $\leq n^2$ mit $P(F) = 0$. (Hinweis: Betrachten Sie $\text{id}, F, F^2, \dots, F^{n^2}$.)

4. Beweisen Sie den Satz von CAYLEY-HAMILTON durch direkte Rechnung für Matrizen

$$A \in M(2 \times 2; K).$$

5. Beweisen Sie den Satz von CAYLEY-HAMILTON für einen diagonalisierbaren Endomorphismus.

6. Geben Sie noch einen anderen Beweis des Satzes von CAYLEY-HAMILTON durch Induktion nach $n = \dim V$ mit der folgenden Methode:

- Zeigen Sie, dass $F(W) \subset W$ und berechnen Sie $P_G(t)$ für $G := F|_W$ (siehe Aufgabe 4 in 4.2).
- Zeigen Sie $P_G(G) = 0 \in \text{End}(W)$.
- Folgern Sie daraus im Fall $k < n$ mit der Bemerkung aus 4.4.4 und der Induktionsannahme, dass $P_F(F) = 0$.

7. Seien $\lambda_1, \dots, \lambda_r \in K$ die paarweise verschiedenen Eigenwerte eines diagonalisierbaren Endomorphismus F über einem endlichdimensionalen Vektorraum. Zeigen Sie, dass $(t - \lambda_1) \cdots (t - \lambda_r) \in K[t]$ das Minimalpolynom von F ist.

4.6 Die Jordansche Normalform*

Gegenstand dieses Abschnittes sind Endomorphismen, deren charakteristisches Polynom in Linearfaktoren zerfällt. Bedingungen für die Diagonalisierbarkeit sind in 4.1.3 und 4.3.3 angegeben. Sind sie nicht erfüllt, so kann man entsprechend 4.4.3 wenigstens trigonalisieren. Unbefriedigend dabei ist, daß man gar nichts von den Einträgen über der Diagonale weiß. Das wird nun nachgeholt.

Wir benutzen dabei nur elementare Methoden der linearen Algebra und verzichten darauf, das Minimalpolynom und die Teilbarkeitstheorie im Polynomring einzusetzen. Damit geht manches etwas eleganter, aber insgesamt nicht kürzer (vgl. etwa [B2]).

4.6.1. Im folgenden sei stets K ein beliebiger Körper und V ein K -Vektorraum mit $\dim V \geq 1$. Ist $F \in \text{End}(V)$ und

$P_F = \pm(t - \lambda_1)^{r_1} \cdot \dots \cdot (t - \lambda_k)^{r_k}$ mit $\lambda_1, \dots, \lambda_k \in K$ paarweise verschieden, so ist V nach 4.3.3 genau dann direkte Summe der F -invarianten Eigenräume $\text{Eig}(F; \lambda_i)$, wenn

$$\dim \text{Eig}(F; \lambda_i) = \mu(P_F; \lambda_i) = r_i.$$

Ist die Dimension des Eigenraumes zu klein, so kann man ihn mit folgendem Trick passend vergrößern. Für einen Eigenwert λ der Vielfachheit $r \geq 1$ ist

$$\text{Eig}(F; \lambda) = \text{Ker}(F - \lambda \text{id}_V) \subset \text{Ker}(F - \lambda \text{id}_V)^r =: \text{Hau}(F; \lambda).$$

Man nennt $\text{Hau}(F; \lambda)$ den *Hauptraum* (oder *verallgemeinerten Eigenraum*) von F zum Eigenwert λ . Damit kann man schon das Ergebnis der ersten Etappe formulieren:

Satz über die Hauptraumzerlegung. Sei $F \in \text{End}_K(V)$ und

$$P_F = \pm(t - \lambda_1)^{r_1} \cdot \dots \cdot (t - \lambda_k)^{r_k}$$

mit paarweise verschiedenen $\lambda_1, \dots, \lambda_k \in K$. Es sei $V_i := \text{Hau}(F; \lambda_i) \subset V$ für jedes λ_i der Hauptraum. Dann gilt:

1) $F(V_i) \subset V_i$ und $\dim V_i = r_i$ für $i = 1, \dots, k$.

2) $V = V_1 \oplus \dots \oplus V_k$.

3) F hat eine Zerlegung $F = F_D + F_N$ mit

a) F_D diagonalisierbar,

b) F_N nilpotent,

c) $F_D \circ F_N = F_N \circ F_D$.

Übersetzt man das in die Sprache der Matrizen, und verwendet man noch, daß jede nilpotente Matrix in eine obere Dreiecksmatrix transformiert werden kann (4.5.7), so erhält man das

Korollar. Sei $A \in M(n \times n; K)$, so daß $P_A = \pm(t - \lambda_1)^{r_1} \cdot \dots \cdot (t - \lambda_k)^{r_k}$. Dann gibt es eine invertierbare Matrix $S \in GL(n; K)$ derart, daß

$$SAS^{-1} = \begin{pmatrix} \boxed{\lambda_1 E_{r_1} + N_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{\lambda_k E_{r_k} + N_k} \end{pmatrix} =: \tilde{A},$$

wobei für $i = 1, \dots, k$

$$\lambda_i E_{r_i} + N_i = \begin{pmatrix} \lambda_i & & * \\ & \ddots & \\ 0 & & \lambda_i \end{pmatrix} \in M(r_i \times r_i; K).$$

Insbesondere ist $\tilde{A} = D + N$, wobei D Diagonalmatrix, N nilpotent ist, und es gilt

$$D \cdot N = N \cdot D. \quad \square$$

Folgerungen aus der letzten Eigenschaft sind in Aufgabe 4 und Aufgabe 5 enthalten.

4.6.2. Zur Vorbereitung der Hauptraumzerlegung betrachten wir einen Eigenwert λ von F und Potenzen von

$$G := F - \lambda \text{id}_V.$$

Das können wir auch für ein ganz beliebiges $G \in \text{End } V$ tun. Offensichtlich hat man die Ketten

$$\begin{aligned} \{0\} &\subset \text{Ker } G \subset \text{Ker } G^2 \subset \dots \subset \text{Ker } G^l, \\ V &\supset \text{Im } G \supset \text{Im } G^2 \supset \dots \supset \text{Im } G^l, \end{aligned}$$

und für jedes l ist $\dim \text{Ker } G^l + \dim \text{Im } G^l = \dim V$ nach 2.2.4. Im allgemeinen ist jedoch

$$\text{Ker } G^l \cap \text{Im } G^l \neq \{0\},$$

man hat also keine direkte Summe (vgl. Beispiel 4.5.6). Da V endlichdimensional ist, können die beiden Ketten nicht endlos auf- bzw. absteigen. Das ist der Schlüssel zum entscheidenden

Lemma von FITTING. Zu einem $G \in \text{End}_K(V)$ betrachten wir die beiden Zahlen

$$d := \min\{l \in \mathbb{N} : \text{Ker } G^l = \text{Ker } G^{l+1}\} \quad \text{und} \quad r := \mu(P_G; 0),$$

wobei $G^0 := \text{id}_V$. Dann gilt:

- 1) $d = \min\{l : \text{Im } G^l = \text{Im } G^{l+1}\}$.
- 2) $\text{Ker } G^{d+i} = \text{Ker } G^d$, $\text{Im } G^{d+i} = \text{Im } G^d$ für alle $i \in \mathbb{N}$.
- 3) Die Räume $U := \text{Ker } G^d$ und $W := \text{Im } G^d$ sind G -invariant.
- 4) $(G|U)^d = 0$ und $G|W : W \rightarrow W$ ist ein Isomorphismus.
- 5) Für das Minimalpolynom von $G|U$ gilt $M_{G|U} = t^d$.
- 6) $V = U \oplus W$, $\dim U = r \geq d$, $\dim W = n - r$.

Insbesondere gibt es eine Basis B von V , so daß

$$M_B(G) = \begin{pmatrix} N & 0 \\ 0 & C \end{pmatrix} \quad \text{mit} \quad N^d = 0 \quad \text{und} \quad C \in \text{GL}(n-r; K).$$

Beweis. Wir betrachten das Diagramm

$$\begin{array}{ccccc} \text{Ker } G^l & \subset & V & \xrightarrow{G^l} & \text{Im } G^l \\ & \cap & & \parallel & \cup \\ \text{Ker } G^{l+1} & \subset & V & \xrightarrow{G^{l+1}} & \text{Im } G^{l+1} \end{array}$$

Unter Benutzung der Dimensionsformel 2.2.4 für G^l und G^{l+1} folgt

$$\begin{aligned} \text{Im } G^{l+1} = \text{Im } G^l & \Leftrightarrow \dim \text{Im } G^{l+1} = \dim \text{Im } G^l \\ & \Leftrightarrow \dim \text{Ker } G^{l+1} = \dim \text{Ker } G^l \\ & \Leftrightarrow \text{Ker } G^{l+1} = \text{Ker } G^l. \end{aligned}$$

Dies ist weiter gleichwertig damit, daß $G|_{\text{Im } G^l} : \text{Im } G^l \rightarrow \text{Im } G^{l+1}$ ein Isomorphismus ist. Daraus folgt sofort 1) und $G|_W$ Isomorphismus, also auch 2). Der Rest von 3) und 4) ist klar. Zu 5) genügt es zu zeigen, daß aus $(G|U)^{d-1} = 0$ folgen würde, daß

$$\text{Ker } G^d = U \subset \text{Ker } G^{d-1}.$$

Ist $v \in U \cap W$, so ist $G^d(v) = 0$ und $v = G^d(w)$ für ein $w \in V$. Daraus folgt $G^{2d}(w) = 0$, also

$$w \in \text{Ker } G^{2d} = \text{Ker } G^d$$

nach 2), und somit $0 = G^d(w) = v$. Damit hat man $V = U \oplus W$, und es bleibt die Berechnung der Dimension von U . Nach Definition von U ist $\dim U \geq d$. Wegen

$$t^r \cdot Q = P_G = P_{G|U} \cdot P_{G|W} \quad \text{mit } Q(0) \neq 0$$

und $P_{G|U} = \pm t^m$ mit $m = \dim U$ folgt $m = r$, denn für den Isomorphismus $G|W$ ist $P_{G|W}(0) \neq 0$. $\square$

Insbesondere folgt aus den obigen Argumenten die

Bemerkung. Ist λ Eigenwert von F und $G := F - \lambda \text{id}_V$, so ist $d = 1$ gleichbedeutend mit

$$\text{Eig}(F; \lambda) = \text{Hau}(F; \lambda).$$

$\square$

4.6.3. Nun beweisen wir den Satz über die Hauptraumzerlegung durch Induktion über die Zahl $k \geq 1$ der verschiedenen Eigenwerte. Zu λ_1 definieren wir $G := F - \lambda_1 \cdot \text{id}_V$. Dann ist

$$P_G(t - \lambda_1) = P_F(t), \quad \text{also} \quad \mu(P_G; 0) = \mu(P_F; \lambda_1) = r_1.$$

Nach Lemma 4.6.2 ist $V = \text{Hau}(F; \lambda_1) \oplus W$, und die beiden Summanden sind wegen $F = G + \lambda_1 \text{id}_V$ auch F -invariant. Weiter gilt

$$P_{F|W} = \pm(t - \lambda_2)^{r_2} \cdot \dots \cdot (t - \lambda_k)^{r_k},$$

also können wir auf $F|W$ die Induktionsannahme anwenden, daraus folgen 1) und 2).

Die Zerlegung 3) zeigt man am einfachsten mit Matrizen, die Bezeichnungen seien wie im Korollar gewählt. Die Existenz erhält man mit

$$D := \begin{pmatrix} \boxed{\lambda_1 E_{r_1}} & & 0 \\ & \ddots & \\ 0 & & \boxed{\lambda_k E_{r_k}} \end{pmatrix} \quad \text{und} \quad N := \begin{pmatrix} \boxed{N_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{N_k} \end{pmatrix},$$

eine einfache Rechnung zeigt

$$D \cdot N = \begin{pmatrix} \boxed{\lambda_1 N_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{\lambda_k N_k} \end{pmatrix} = N \cdot D.$$

$\square$

Zusatz. Die Zerlegung $F = F_D + F_N$ ist sogar eindeutig, wenn man a), b) und c) verlangt.

Der Beweis benutzt 4.3.6 und etwas Teilbarkeitstheorie von Polynomen (vgl. etwa [B2], §11).

4.6.4. Aus den obigen Beweisen kann man ein Rechenverfahren ableiten, wir geben dafür ein

Beispiel. Gegeben sei die Matrix

$$A = \begin{pmatrix} 25 & 34 & 18 \\ -14 & -19 & -10 \\ -4 & -6 & -1 \end{pmatrix}$$

mit $P_A = -t^3 + 5t^2 - 7t + 3 = -(t-1)^2 \cdot (t-3)$. In der Notation von 4.6.3 ist $k=2$, $\lambda_1=1$, $r_1=2$, $\lambda_2=3$, $r_2=1$. Wir betrachten die Matrizen

$$B_1 = A - \lambda_1 E_3 = \begin{pmatrix} 24 & 34 & 18 \\ -14 & -20 & -10 \\ -4 & -6 & -2 \end{pmatrix}, \quad B_2 = A - \lambda_2 E_3 = \begin{pmatrix} 22 & 34 & 18 \\ -14 & -22 & -10 \\ -4 & -6 & -4 \end{pmatrix},$$

sowie

$$B_1^{r_1} = B_1^2 = \begin{pmatrix} 28 & 28 & 56 \\ -16 & -16 & -32 \\ -4 & -4 & -8 \end{pmatrix}, \quad B_2^{r_2} = B_2.$$

Es ist $\text{rang } B_1 = 2 > 1 = \text{rang } B_1^2$. Daraus folgt

$$\text{Eig}(A; \lambda_1) = \text{Ker } B_1 \subsetneq \text{Ker } B_1^2 = \text{Hau}(A; \lambda_1),$$

und somit ist A nicht diagonalisierbar. Indem man die zu B_1^2 und B_2 gehörenden homogenen Gleichungssysteme löst, erhält man Basen

$$({}'(2, 0, -1), {}'(0, 2, -1)) \quad \text{von } \text{Hau}(A; \lambda_1)$$

und

$${}'(-7, 4, 1) \quad \text{von } \text{Hau}(A; \lambda_2) = \text{Eig}(A; \lambda_2).$$

Trägt man die erhaltenen Basisvektoren als Spalten in eine Matrix ein, so erhält man

$$S^{-1} = \begin{pmatrix} 2 & 0 & -7 \\ 0 & 2 & 4 \\ -1 & -1 & 1 \end{pmatrix} \quad \text{und} \quad S = \frac{1}{2} \begin{pmatrix} -6 & -7 & -14 \\ 4 & 5 & 8 \\ -2 & -2 & -4 \end{pmatrix},$$

sowie

$$SAS^{-1} = \begin{pmatrix} 16 & 25 & 0 \\ -9 & -14 & 0 \\ 0 & 0 & 3 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 16 & 25 \\ -9 & -14 \end{pmatrix} = E_2 + N \quad \text{mit } N^2 = 0.$$

Nun muß man noch die Basis von $\text{Hau}(A; \lambda_1)$ so transformieren, daß ein Basisvektor v Eigenvektor zu λ_1 wird (vgl. 4.4.5). Er hat die Form

$$v = \alpha \cdot {}'(2, 0, -1) + \beta \cdot {}'(0, 2, -1).$$

Die Bedingung $Av = v$ ergibt $3\alpha + 5\beta = 0$, also kann man $v = {}'(5, -3, -1)$ wählen. Aus der neuen Basis von $\mathbb{R}^3$ erhält man die Transformationsmatrix

$$T^{-1} = \begin{pmatrix} 5 & 2 & -7 \\ -3 & 0 & 4 \\ -1 & -1 & 1 \end{pmatrix}, \quad T = \frac{1}{3} \begin{pmatrix} -4 & -5 & -8 \\ 1 & 2 & -1 \\ -3 & -3 & -6 \end{pmatrix},$$

sowie

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

4.6.5. Auf der zweiten Etappe des Weges zur JORDANSchen Normalform wird zu einem nilpotenten Endomorphismus eine Basis maßgeschneidert. Um das Ergebnis einfach formulieren zu können, definieren wir die *Jordanmatrix*

$$J_k := \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix} \in M(k \times k; K),$$

für die $J_k^k = 0$ gilt, und k die minimale Potenz mit dieser Eigenschaft ist. Man beachte, daß $J_1 = (0)$.

Theorem. Sei G ein nilpotenter Endomorphismus eines K -Vektorraumes V und $d := \min\{l : G^l = 0\}$. Dann gibt es eindeutig bestimmte Zahlen $s_1, \dots, s_d \in \mathbb{N}$ mit

$$d \cdot s_d + (d-1)s_{d-1} + \dots + s_1 = r = \dim V$$

und eine Basis $\mathcal{B}$ von V , so daß

$$M_B(G) = \begin{pmatrix} J_d & & & & & \\ & \ddots & & & & \\ & & J_d & & & \\ & & & J_{d-1} & & \\ & & & & \ddots & \\ & & & & & 0 \\ & & & & & & \ddots \\ & & & & & & & J_{d-1} \\ & & & & & & & & \ddots \\ & & & & & & & & & J_1 \\ & & & & & & & & & & \ddots \\ & & & & & & & & & & & J_1 \end{pmatrix} \begin{matrix} \left. \begin{matrix} \\ \\ \\ \\ \end{matrix} \right\} s_d\text{-mal} \\ \left. \begin{matrix} \\ \\ \\ \end{matrix} \right\} s_{d-1}\text{-mal} \\ \vdots \\ \left. \begin{matrix} \\ \\ \end{matrix} \right\} s_1\text{-mal} \end{matrix}$$

Beweis. Wir definieren $U_l := \text{Ker } G^l$ und betrachten die Kette

$$\{0\} = U_0 \subset U_1 \subset \dots \subset U_{d-1} \subset U_d = V.$$

Da d minimal ist, sind alle Inklusionen echt. Wir vermerken zwei weitere Eigenschaften:

(1) Für $1 \leq l \leq d$ ist $G^{-1}(U_{l-1}) = U_l$, insbesondere $G(U_l) \subset U_{l-1}$.

Das ist klar, denn

$$v \in G^{-1}(U_{l-1}) \Leftrightarrow G(v) \in U_{l-1} \Leftrightarrow 0 = G^{l-1}(G(v)) = G^l(v) \Leftrightarrow v \in U_l.$$

Durch Anwendung von G auf die erste Gleichung folgt

$$G(U_l) = G(G^{-1}(U_{l-1})) \subset U_{l-1}.$$

Daß hier keine Gleichheit mehr gelten muß, sieht man etwa an der durch

$$G(e_1) = G(e_2) = 0, \quad G(e_3) = e_2$$

gegebenen Abbildung des K^3 mit $d = 2$.

(2) Ist $W \subset V$ ein Untervektorraum mit $W \cap U_l = \{0\}$ für irgend ein $l > 0$, so ist $G|_W$ injektiv.

Denn da $\text{Ker } G = U_1 \subset U_l$, gilt sogar $W \cap \text{Ker } G = \{0\}$.

Damit konstruieren wir nun schrittweise eine direkte Summenzerlegung von V . Zunächst wählen wir $W_d \subset V$ mit

$$V = U_d = U_{d-1} \oplus W_d.$$

Aus (1) folgt $G(W_d) \subset U_{d-1}$ und $G(W_d) \cap U_{d-2} = \{0\}$. Also gibt es eine Zerlegung

$$U_{d-1} = U_{d-2} \oplus W_{d-1} \quad \text{mit} \quad G(W_d) \subset W_{d-1}.$$

Die Iteration dieses Verfahrens kann man schematisch darstellen, wobei die Pfeile die Wirkung von G zeigen:

$$\begin{array}{ccccccc}
 U_d & & & & & & \\
 \downarrow & & & & & & \\
 U_{d-1} \oplus & W_d & & & & & \\
 \downarrow & & \downarrow & & & & \\
 U_{d-2} \oplus & W_{d-1} \oplus & W_d & & & & \\
 \downarrow & & \downarrow & & \downarrow & & \\
 \vdots & & \vdots & & \vdots & & \\
 U_1 \oplus & W_2 \oplus & W_3 \oplus & \dots \oplus & W_d & & \\
 \downarrow & & \downarrow & & \downarrow & & \downarrow \\
 U_0 \oplus & W_1 \oplus & W_2 \oplus & \dots \oplus & W_{d-1} \oplus & W_d &
 \end{array}$$

Jede Zeile ist dabei eine Zerlegung von V , wegen $U_0 = 0$ ist insbesondere

$$V = W_1 \oplus W_2 \oplus \dots \oplus W_d.$$

Da nach (2) all die Beschränkungen von G in der Kette

$$W_d \rightarrow W_{d-1} \rightarrow \dots \rightarrow W_1$$

injektiv sind, können wir nun bei W_d beginnend durch schrittweise Ergänzung Basen aller W_i und damit insgesamt eine Basis von V nach folgendem Schema erhalten:

$$\begin{array}{ccccccc}
 w_1^{(d)}, \dots, w_{s_d}^{(d)} & & & & & & \\
 G(w_1^{(d)}), \dots, G(w_{s_d}^{(d)}), & w_1^{(d-1)}, \dots, w_{s_{d-1}}^{(d-1)} & & & & & \\
 \vdots & \vdots & \vdots & & \vdots & & \\
 G^{d-1}(w_1^{(d)}), \dots, G^{d-1}(w_{s_d}^{(d)}), & G^{d-2}(w_1^{(d-1)}), \dots, G^{d-2}(w_{s_{d-1}}^{(d-1)}), \dots, & w_1^{(1)}, \dots, w_{s_1}^{(1)} & & & &
 \end{array}$$

Dabei ist die erste Zeile eine Basis von W_d , die zweite von W_{d-1} , und schließlich die letzte von

$$W_1 = U_1 = \text{Ker } G.$$

Die Matrix $M_{\mathcal{B}}(G)$ hat die versprochene Form, wenn man $\mathcal{B}$ in folgender Weise anordnet: Man läuft in jeder Spalte von unten nach oben und liest die Spalten von links nach rechts.

Um zu zeigen, daß die Zahlen $s_1, \dots, s_d$ allein durch G bestimmt sind, benutzen wir die Existenz von Zerlegungen

$$U_l = U_{l-1} \oplus W_l = U_{l-1} \oplus G(W_{l+1}) \oplus \tilde{W}_l$$

für $l = 1, \dots, d$ mit $W_{d+1} = 0$. Danach ist

$$s_l = \dim \tilde{W}_l = \dim U_l - \dim U_{l-1} - \dim W_{l+1},$$

also sind diese Zahlen rekursiv aus den Dimensionen der Kerne von G^l berechenbar (vgl. Aufgabe 6). $\square$

4.6.6. Wir berechnen ein einfaches

Beispiel. Gegeben sei die Matrix

$$B = \begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{mit} \quad B^2 = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{und} \quad B^3 = 0.$$

Es ist $d = 3$, und für $U_l := \text{Ker } B^l$ gilt

$$\{0\} = U_0 \subset U_1 = \text{span } e_1 \subset U_2 = \text{span } (e_1, e_2) \subset U_3 = \mathbb{R}^3.$$

Aus der Bedingung

$$\mathbb{R}^3 = U_2 \oplus W_3$$

folgt, daß wir $W_3 = \text{span } e_3$ wählen können. Somit ist $s_3 = 1$. Aus

$$\mathbb{R}^3 = U_1 \oplus W_2 \oplus W_3$$

folgt $\dim W_2 = 1$, also $s_2 = 0$, und $B \cdot e_3 = {}^t(3, 2, 0)$ ist der richtige Basisvektor von W_2 . Schließlich ist

$$\mathbb{R}^3 = U_0 \oplus W_1 \oplus W_2 \oplus W_3.$$

Der gesuchte Basisvektor von W_1 ist $B^2 \cdot e_3 = B \cdot {}^t(3, 2, 0) = 2e_1$, somit ist auch $s_1 = 0$. Trägt man die gefundenen Basisvektoren in der richtigen Reihenfolge als Spalten in eine Matrix ein, so erhält man

$$T^{-1} = \begin{pmatrix} 2 & 3 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{und daraus} \quad T = \frac{1}{4} \begin{pmatrix} 2 & -3 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

Das Ergebnis ist die Jordanmatrix

$$TBT^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

mit d Einträgen λ_i und $d - 1$ Einsen heißt *Jordanblock* der Länge d zu λ_i . Der oberste und größte Jordanblock in $\lambda_i E_{r_i} + N_i$ hat die Größe d_i mit

$$1 \leq d_i = \min \{l: N_i^l = 0\} \leq r_i,$$

das ist die Vielfachheit der Nullstelle λ_i im Minimalpolynom von F (Aufgabe 8).

Für $1 \leq j \leq d_i$ seien $s_j^{(i)} \geq 0$ die Anzahlen der Jordanblöcke der Größe j zu λ_i in $\lambda_i E_{r_i} + N_i$. Es ist $s_{d_i}^{(i)} \geq 1$, und durch Abzählung der Längen folgt

$$d_i s_{d_i}^{(i)} + (d_i - 1) s_{d_i-1}^{(i)} + \dots + s_1^{(i)} = r_i.$$

Die Elemente $\lambda_i \in K$ und $r_i, d_i, s_j^{(i)} \in \mathbb{N}$ sind allein durch den Endomorphismus F bestimmt (für die $s_j^{(i)}$ folgt das aus Aufgabe 6). Sie heißen daher *Invarianten* von F . Die Eigenwerte λ_i kann man im Fall $K = \mathbb{R}$ oder $\mathbb{C}$ *kontinuierliche* Invarianten nennen. Die anderen sind *diskret* und durch die oben angegebenen Bezeichnungen, sowie

$$r_1 + \dots + r_k = n$$

aneinander gekoppelt. Alle Invarianten zusammen beschreiben in komprimierter Form die „Geometrie von F “. Umgekehrt bedeutet dies, daß es so viele Möglichkeiten gibt, wesentlich verschiedene Endomorphismen zu konstruieren, wie man „Sätze“ von solchen Invarianten zusammenstellen kann. Nur die Reihenfolge der Jordanblöcke ist unwesentlich, sie kann durch eine passende Permutation der Basisvektoren beliebig verändert werden. Das ist die Aussage dieses wohl schwierigsten Theorems der elementaren linearen Algebra.

Das Verfahren, die Normalform zu berechnen, hat praktische Anwendungen, etwa bei der Lösung von Systemen linearer Differentialgleichungen mit konstanten Koeffizienten. An sich genügt es, die Systeme nach verschiedenen Eigenwerten zu entkoppeln (das geht mit der Hauptraumzerlegung) und in jedem Hauptraum zu trigonalisieren (dann kann man die Teilsysteme von unten nach oben lösen, vgl. [Fo 2], §14). Aber der Rechenaufwand für die verbesserte Trigonalisierung mit der JORDAN-Methode ist kaum größer als bei der Methode aus 4.4.5.

Nach all den schon geleisteten Vorbereitungen ist der *Beweis* nunmehr ganz kurz: Für $i = 1, \dots, k$ setzen wir

$$V_i := \text{Hau}(F; \lambda_i) \quad \text{und} \quad G_i := (F - \lambda_i \text{id}_V)|_{V_i}.$$

Anwendung von 4.6.5 auf G_i ergibt eine Basis $\mathcal{B}_i$ von V_i , und aus $\mathcal{B}_1, \dots, \mathcal{B}_k$ baut man sich das gesuchte $\mathcal{B}$ auf (vgl. 1.6.4). $\square$

Für die Minimalpolynome von G_i und F gilt (Aufgabe 8)

$$M_{G_i} = t^{d_i} \quad \text{und} \quad M_F = (t - \lambda_1)^{d_1} \cdot \dots \cdot (t - \lambda_k)^{d_k}.$$

Ist speziell $d_1 = \dots = d_k = 1$, so gibt es keine Einsen neben der Diagonale, und es folgt zusammen mit 4.6.1 und 4.6.7 das

Korollar. Für ein $F \in \text{End}_K(V)$ sind folgende Bedingungen gleichwertig:

- i) F ist diagonalisierbar.
- ii) $M_F = (t - \lambda_1) \cdot \dots \cdot (t - \lambda_k)$, wobei $\lambda_1, \dots, \lambda_k$ die verschiedenen Eigenwerte von F bezeichnen.
- iii) Es gibt paarweise verschiedene $\lambda_1, \dots, \lambda_m \in K$, so daß

$$(F - \lambda_1 \text{id}_V) \circ \dots \circ (F - \lambda_m \text{id}_V) = 0 \in \text{End}_K(V).$$

□

4.6.8. Für die praktische Berechnung der Jordanschen Normalform verwenden wir wieder das

Beispiel.

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

mit dem charakteristischen Polynom $P_A = -(t-2)^3$ (vgl. 4.4.5). Wir betrachten

$$B := A - 2E_3 = \begin{pmatrix} 1 & 4 & 3 \\ -1 & -2 & -1 \\ 1 & 2 & 1 \end{pmatrix} \quad \text{mit} \quad B^2 = \begin{pmatrix} 0 & 2 & 2 \\ 0 & -2 & -2 \\ 0 & 2 & 2 \end{pmatrix}$$

und $B^3 = 0$. Dann ist

$$U_1 := \text{Ker } B = \text{span}'(1, -1, 1)$$

$$U_2 := \text{Ker } B^2 = \text{span}'(1, -1, 1), (0, -1, 1).$$

Entsprechend 4.6.5 haben wir die Zerlegungen

$$\mathbb{R}^3 = U_2 \oplus W_3 = U_1 \oplus W_2 \oplus W_3 = U_0 \oplus W_1 \oplus W_2 \oplus W_3,$$

das ergibt die Basisvektoren

$$e_3 \in W_3, \quad B \cdot e_3 = (3, -1, 1) \in W_2 \quad \text{und} \quad B^2 \cdot e_3 = (2, -2, 2) \in W_1.$$

In der richtigen Reihenfolge als Spalten in eine Matrix eingetragen erhält man

$$T^{-1} = \begin{pmatrix} 2 & 3 & 0 \\ -2 & -1 & 0 \\ 2 & 1 & 1 \end{pmatrix} \quad \text{mit} \quad T = \frac{1}{4} \begin{pmatrix} -1 & -3 & 0 \\ 2 & 2 & 0 \\ 0 & 4 & 4 \end{pmatrix},$$

und als Ergebnis

$$TAT^{-1} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Das kann man vergleichen mit den Ergebnissen der Trigonalisierung in 4.4.5

$$SAS^{-1} = D = \begin{pmatrix} 2 & 1 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{pmatrix} \quad \text{mit} \quad S = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix},$$

und der Transformation der nilpotenten Matrix $D - 2E_3$ zu

$$\tilde{T}(D - 2E_2)\tilde{T}^{-1} = J_3 \quad \text{mit} \quad \tilde{T} = \frac{1}{4} \begin{pmatrix} 2 & -3 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

Es ist $\tilde{T} \cdot S = T$, also kann man T auch durch Kombination der vorher durchgeführten Rechnungen erhalten. Man beachte, daß dabei nicht notwendig dieselbe Transformationsmatrix entstehen muß, weil sie nicht eindeutig bestimmt ist.

Hat A mehr als einen Eigenwert, so muß man zunächst die Zerlegung in Haupträume berechnen. Dann kann man in jedem Hauptraum wie oben vorgehen. Da der gesamte Raum nach 4.6.1 direkte Summe der Haupträume ist, ergeben die unabhängig voneinander konstruierten Basen der Haupträume zusammen eine Basis der gewünschten Art des gesamten Raumes.

Aufgaben zu 4.6

1. Bestimmen Sie die Haupträume der folgenden Matrizen:

$$\begin{pmatrix} 1 & 4 & 2 & 1 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} 2 & 3 & 3 & 1 & 8 \\ 0 & 2 & 7 & 2 & 8 \\ 0 & 0 & 2 & 5 & 4 \\ 0 & 0 & 0 & -1 & -4 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix}.$$

2. Bestimmen Sie Basen, bezüglich derer die folgenden nilpotenten Matrizen Jordansche Normalform haben, und geben Sie jeweils das Minimalpolynom an:

$$\begin{pmatrix} 0 & 2 & 2 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & -2 & 0 & -1 & 2 \\ 1 & -3 & -1 & 0 & 3 \\ 0 & 2 & 1 & -1 & -3 \\ 1 & 0 & 0 & -1 & -2 \\ 0 & -1 & 0 & 0 & 2 \end{pmatrix}.$$

3. Bestimmen Sie Basen, bezüglich derer die folgenden Matrizen Jordansche Normalform haben, und geben Sie jeweils das charakteristische und das Minimalpolynom an:

$$\begin{pmatrix} 3 & 4 & 3 \\ -1 & 0 & -1 \\ 1 & 2 & 3 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 & 1 & 0 & -2 \\ 1 & 1 & 1 & 0 & -1 \\ 1 & 0 & 2 & 0 & -1 \\ 1 & 0 & 1 & 2 & -2 \\ 1 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

4. Mit Hilfe des Satzes über die Jordansche Normalform kann man recht einfach hohe Potenzen von Matrizen berechnen. Zeigen Sie:

- a) Ist $A \in M(n \times n; K)$, $S \in GL(n; K)$ und $m \in \mathbb{N}$, so gilt $(SAS^{-1})^m = SA^mS^{-1}$.
 b) Sind $A, B \in M(n \times n; K)$ mit $AB = BA$ und $m \in \mathbb{N}$, so gilt

$$(A + B)^m = \sum_{k=0}^m \binom{m}{k} A^k B^{m-k}.$$

- c) Bestimmen Sie für die Matrix

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

eine Matrix $S \in GL(3; \mathbb{R})$, so dass $A = S(D + N)S^{-1}$, wobei D Diagonalmatrix, N nilpotent und $DN = ND$ ist. Berechnen Sie mit Hilfe von a) und b) (und ohne Computer) A^{50} .

5. Betrachten Sie die Verallgemeinerung der Exponentialfunktion für Matrizen; für jede Matrix $A \in M(n \times n; \mathbb{R})$ existiert

$$\exp(A) := \lim_{m \rightarrow \infty} \sum_{k=0}^m \frac{1}{k!} A^k.$$

- a) Bestimmen Sie $\exp(D)$ für eine Diagonalmatrix D .

- b) Ist $A \in M(n \times n; \mathbb{R})$ und $S \in GL(n; \mathbb{R})$, so folgt $\exp(SAS^{-1}) = S \cdot \exp(A) \cdot S^{-1}$.
 c) Sind $A, B \in M(n \times n; \mathbb{R})$ mit $AB = BA$, so gilt $\exp(A + B) = \exp(A)\exp(B)$.
 d) Bestimmen Sie für die Matrix

$$A = \begin{pmatrix} 3 & 0 & -2 \\ -2 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix}$$

eine Matrix $S \in GL(3; \mathbb{R})$, so dass $A = S(D + N)S^{-1}$, wobei D Diagonalmatrix, N nilpotent und $DN = ND$ ist, und berechnen Sie $\exp(A)$.

6. Zeigen Sie, dass für die Zahlen $s_1, \dots, s_d$ in 4.6.5 gilt:

$$s_l = \dim(U_l/U_{l-1}) - \dim(U_{l+1}/U_l).$$

7. Gegeben sei ein endlichdimensionaler $\mathbb{C}$ -Vektorraum V . Zwei Endomorphismen F und G von V heißen *ähnlich*, wenn es einen Isomorphismus H von V gibt mit $G = H \circ F \circ H^{-1}$.

- a) Zeigen Sie, dass dadurch eine Äquivalenzrelation auf der Menge der Endomorphismen von V gegeben ist.
 b) Für $F, G \in \text{End}(V)$ sind folgende Bedingungen gleichwertig:
 i) F und G sind ähnlich.
 ii) Für jede Basis $\mathcal{B}$ von V sind $M_{\mathcal{B}}(F)$ und $M_{\mathcal{B}}(G)$ ähnlich.
 iii) Die Jordanschen Normalformen von F und G haben (bis auf die Reihenfolge) die gleichen Invarianten.

8. Sei $F \in \text{End}_K(V)$ ein Endomorphismus, dessen charakteristisches Polynom in Linearfaktoren zerfällt. Beweisen Sie, dass man das Minimalpolynom M_F aus den Invarianten von F berechnen kann: Mit den Bezeichnungen von 4.6.7 gilt

$$M_F(t) = (t - \lambda_1)^{d_1} \cdot \dots \cdot (t - \lambda_k)^{d_k}.$$

9. Sei V ein 6-dimensionaler $\mathbb{R}$ -Vektorraum und F ein Endomorphismus von V mit $P_F(t) = (t - 1)(t + 2)^5$, $M_F(t) = (t - 1)(t + 2)^3$. Bestimmen Sie alle möglichen Jordanschen Normalformen von F .

10. Sei V ein endlichdimensionaler $\mathbb{R}$ -Vektorraum und F ein Endomorphismus von V mit $F^3 = F$. Zeigen Sie, dass F diagonalisierbar ist.