Kapitel 3 Lineare Abbildungen

Ebenso wie bei den Gruppen und den Körpern wollen wir auch bei den Vektorräumen die strukturverträglichen Abbildungen studieren. Wir werden sehen, daß sich die endlich dimensionalen Vektorräume als "isomorph" zu einem der Standardräume $\mathbb{K}^n$ erweisen (mit $\mathbb{K}^0 := \{0\}$), so daß wir damit alle "Typen" endlich dimensionaler Vektorräume erfaßt haben.

Zur Vereinfachung und auch zur Präzisierung der Dimension werden wir im folgenden statt von endlich dimensionalen Vektorräumen meist von n–dimensionalen Vektorräumen V sprechen und dabei den Fall $n = 0$, also $V = \{o\}$, einschließen (wenn es nicht ausdrücklich anders gesagt wird). Um lästige Fallunterscheidungen zu vermeiden, werden wir uns aber bei den Beweisen auf Dimensionen $n \geq 1$ konzentrieren und dann mit Basisvektoren $x_1,\dots,x_n \in V$ arbeiten. Die Übertragung der Aussagen auf den Fall $V = \{o\}$ ist meist trivial und wird dem Leser überlassen.

§ 1 Definition und Eigenschaften linearer Abbildungen

Es seien V und W Vektorräume über demselben Körper $\mathbb{K}$. Eine Abbildung Φ von V nach W erhält die Vektorraumstruktur, wenn

$$\Phi\,(x + y) = \Phi\,(x) + \Phi\,(y)\,,$$
$$\Phi\,(a\,x) = a\,\Phi\,(x)$$

für alle x, $y \in V$ und alle $a \in \mathbb{K}$ gilt. Wir können beide Gleichungen zu einer einzigen zusammenfassen:

Definition. Es seien V und W $\mathbb{K}$–Vektorräume. Eine Abbildung $\Phi : V \longrightarrow W$ heißt *linear* oder (*Vektorraum–*) *Homomorphismus*, wenn

$$\Phi\,(a\,x + b\,y) = a\,\Phi\,(x) + b\,\Phi\,(y)$$

für alle x, $y \in V$ und alle $a,b \in \mathbb{K}$ gilt. Eine bijektive lineare Abbildung heißt (*Vektor-*

raum–) Isomorphismus. Zwei $\mathbb{K}$–Vektorräume V und W heißen *isomorph*, wenn es einen Isomorphismus $\Phi : V \longrightarrow W$ gibt, Schreibweise: $V \cong W$.

Offensichtlich ist $\cong$ eine reflexive Relation für $\mathbb{K}$–Vektorräume. Wie wir im nächsten Paragraphen sehen werden, ist $\cong$ auch symmetrisch und transitiv, also eine Äquivalenzrelation.

Weitere Bezeichnungen. Für spezielle lineare Abbildungen gibt es in der Literatur eine Reihe weiterer Bezeichnungen:

$\Phi : V \longrightarrow V$ linear : *Endomorphismus, lineare Selbstabbildung, lineare Transformation, linearer Operator* ,

$\Phi : V \longrightarrow V$ linear und bijektiv : *Automorphismus* ,

$\Phi : V \longrightarrow W$ linear und injektiv : *Monomorphismus* ,

$\Phi : V \longrightarrow W$ linear und surjektiv : *Epimorphismus* .

Bemerkungen. **(a)** Eine lineare Abbildung Φ ist insbesondere ein Homomorphismus der additiven Gruppen V und W. Damit gilt speziell $\Phi(o) = o$.

(b) Zu jeder linearen Abbildung $\Phi : V \longrightarrow W$ gehören in natürlicher Weise die Mengen

$$\text{Bild}\,\Phi := \Phi(V) \text{ und } \text{Kern}\,\Phi := \Phi^{-1}(\{o\}) .$$

Bild Φ ist ein Untervektorraum von W und Φ ist genau dann surjektiv, wenn Bild $\Phi = W$. Kern Φ ist ein Untervektorraum von V und Φ ist genau dann injektiv, wenn Kern $\Phi = \{o\}$.

Beispiele. **(a)** Es seien V, W $\mathbb{K}$–Vektorräume und $w_0 \in W$. Die konstante Abbildung

$$\begin{aligned} \Phi : V &\longrightarrow W \\ v &\longmapsto w_0 \end{aligned}$$

ist genau dann linear, wenn $w_0 = o$ gilt. Dann ist Bild $\Phi = \{o\}$ und Kern $\Phi = V$. Φ ist in diesem Fall die Nullabbildung, $\Phi = O$.

(b) Sei $A \in \mathbb{K}^{m\times n}$. Dann ist die Abbildung $\Phi : \mathbb{K}^n \longrightarrow \mathbb{K}^m$, $x \longmapsto A\,x$, linear. Für

$A = (s_1|\cdots|s_n)$ erhalten wir

$$\text{Bild } \Phi = \{A\, x \mid x \in \mathbb{K}^n\} = \{x_1\, s_1 + \cdots + x_n\, s_n \mid x_1,\ldots,x_n \in \mathbb{K}\} = [s_1,\ldots,s_n].$$

Kern Φ ist der Lösungsraum des homogenen linearen Gleichungssystems $A\, x = o$.

So ist zum Beispiel die Abbildung

$$\Phi : \mathbb{R}^4 \longrightarrow \mathbb{R}^2 \;,\; \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \longmapsto \begin{bmatrix} 1 & 2 & -1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 - x_3 \\ x_1 + x_4 \end{bmatrix}$$

linear, und es gilt

$$\text{Bild } \Phi = \left[\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right] = \mathbb{R}^2$$

sowie

$$\text{Kern } \Phi = \left[\begin{bmatrix} 0 \\ 1/2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1/2 \\ 0 \\ 1 \end{bmatrix}\right].$$

(c) Die Abbildungen $\Phi : \mathbb{R}^3 \longrightarrow \mathbb{R}^3 , (x_1, x_2, x_3) \longmapsto (x_1 x_2, x_2, x_1 x_3)$ und $\Phi : \mathbb{R} \longrightarrow \mathbb{R}^2 , x \longmapsto (x, x + 2)$ sind nicht linear.

(d) Die Abbildung

$$D : \mathbb{R}[X] \longrightarrow \mathbb{R}[X] \;,\; \sum_{i=0}^{n} a_i X^i \longmapsto \sum_{i=1}^{n} a_i \cdot i \cdot X^{i-1}$$

ist linear (*Ableitungsoperator*). Hier ist Bild $D = \mathbb{R}[X]$ und Kern D ist die Menge der konstanten Polynome, also Kern $D \cong \mathbb{R}$.

Schränken wir D ein auf den Untervektorraum $U = [X, X^2, \ldots]$ von $\mathbb{R}[X]$, so ist $D|_U : U \longrightarrow \mathbb{R}[X]$ sowohl surjektiv als auch injektiv, also sind U und $\mathbb{R}[X]$ isomorph.

(e) Es sei $V = \{f : [0,1] \longmapsto \mathbb{R} \mid f \text{ stetig}\}$. Die Abbildung

$$I : V \longrightarrow \mathbb{R} \;,\; f \longmapsto \int_0^1 f(t)\, dt$$

ist linear (*Integrationsoperator*). Hier ist Bild $I = \mathbb{R}$.

(f) Es sei $V = \{ (a_i) \in \mathbb{R}^{\mathbb{N}} \mid (a_i) \text{ konvergent}\}$. Die Abbildung

$$\Phi : V \longrightarrow \mathbb{R} \ , \ (a_i) \longmapsto \lim_{i \to \infty} a_i$$

ist linear. Hier ist Bild $\Phi = \mathbb{R}$ und Kern Φ ist die Menge der Nullfolgen.

Die Unterverktorräume Kern $\Phi \subset V$ bzw. Bild $\Phi \subset W$ sind insbesondere auch Untergruppen von V bzw. W und Φ ist ein Gruppenhomomorphismus. Somit gilt der Homomorphiesatz für Gruppen (S.46). Weil V und W Vektorräume sind und Φ als lineare Abbildung auch mit der Skalarmultiplikation verträglich ist, können wir diesen Satz sofort auf Vektorräume übertragen. Des besseren Verständnisses wegen werden wir ihn nochmals vollständig beweisen.

Satz 1 (Homomorphiesatz für Vektorräume). *Es seien V und W zwei $\mathbb{K}$-Vektorräume und $\Phi : V \longrightarrow W$ eine lineare Abbildung. Dann gilt:*

(a) *Die kanonische Abbildung $k : V \longrightarrow V/_{\text{Kern } \Phi}$ ist linear.*

(b) *Es gibt eine injektive lineare Abbildung $\overline{\Phi} : V/_{\text{Kern } \Phi} \longrightarrow W$ mit $\Phi = \overline{\Phi} \circ k$.*

(c) *Ist Φ surjektiv, so sind $V/_{\text{Kern } \Phi}$ und W isomorph.*

Beweis. (a) $k(x+y) = [x+y]_{\sim} = [x]_{\sim} + [y]_{\sim} = k(x) + k(y)$ nach Definition der Addition im Faktorraum $V/_{\text{Kern } \Phi}$ und $k(a\,x) = [a\,x]_{\sim} = a\,[x]_{\sim} = a\,k(x)$ nach Definition der Multiplikation mit Skalaren in $V/_{\text{Kern } \Phi}$.

(b) Wir betrachten das Diagramm

$$\begin{tikzcd} V \arrow[rr, "\Phi"] \arrow[dr, "k"'] & & W \\ & V/_{\text{Kern } \Phi} \arrow[ur, "\overline{\Phi}"'] & \end{tikzcd}$$

und erklären die gesuchte Abbildung $\overline{\Phi}$ durch $\overline{\Phi}([x]_{\sim}) := \Phi(x)$. Diese Definition ist repräsentantenunabhängig, da $x \sim y$ nach Definition $x - y \in \text{Kern } \Phi$ bedeutet und daraus $\Phi\,(x-y)=o$, also $\Phi\,(x)=\Phi\,(y)$ folgt.

$\overline{\Phi}$ ist linear: $\overline{\Phi}\,(a\,[x]_\sim + b\,[y]_\sim) = \overline{\Phi}\,([a\,x+b\,y]_\sim) = \Phi\,(a\,x+b\,y) = a\,\Phi\,(x)+b\,\Phi\,(y) = a\,\overline{\Phi}\,([x]_\sim)+b\,\overline{\Phi}\,([y]_\sim)$.

$\overline{\Phi}$ ist injektiv: Sei $\overline{\Phi}\,([x]_\sim)=\overline{\Phi}\,([y]_\sim)$. Dann gilt $\Phi\,(x)=\Phi\,(y)$ und $x-y\in$ Kern Φ, d.h. $x\sim y$ und somit $[x]_\sim = [y]_\sim$. Nach Definition von $\overline{\Phi}$ ist schließlich $\overline{\Phi}\circ k=\Phi$.

(c) Ist Φ surjektiv, so gibt es zu jedem $y\in W$ ein $x\in V$ mit $\Phi\,(x)=y$. Dann gilt $\overline{\Phi}\,([x]_\sim)=\Phi\,(x)=y$, d.h. auch $\overline{\Phi}$ ist surjektiv. Da $\overline{\Phi}$ nach (b) injektiv ist, folgt insgesamt, daß $\overline{\Phi}$ ein Isomorphismus ist. ∎

Korollar 2. *Es seien V, W $\mathbb{K}$-Vektorräume und $\Phi: V\longrightarrow W$ eine lineare Abbildung. Dann sind die Vektorräume $V/_{\text{Kern}\,\Phi}$ und* Bild Φ *isomorph.*

Beispiele. (a) $V/_V \cong \{o\}$ (Kor. 2 auf die Nullabbildung angewendet), $V/_{\{o\}} \cong V$ (Kor. 2 auf die Identität angewendet).

(b) Es sei V der Vektorraum der reellen konvergenten Folgen, und die lineare Abbildung $\Phi: V\longrightarrow \mathbb{R}$ sei gegeben durch $(x_i)\longmapsto \lim\limits_{i\to\infty} x_i$. Dann ist Bild $\Phi=\mathbb{R}$ und Kern Φ ist der Vektorraum der Nullfolgen in $\mathbb{R}$. Also gilt $V/_{\text{Kern}\,\Phi}\cong\mathbb{R}$.

Korollar 3. *Es seien V ein $\mathbb{K}$-Vektorraum und U, W Untervektorräume von V mit $V=U\oplus W$. Dann gilt*

$$V/_U \cong W \quad \textit{und} \quad V/_W \cong U.$$

Beweis. Wegen $V=U\oplus W$ besitzt jeder Vektor $x\in V$ eine eindeutige Darstellung $x=u+w$ mit $u\in U$ und $w\in W$. Die Abbildung $\pi: V\longrightarrow V$, $x\longmapsto w$ mit $x=u+w$ ist linear. Es ist Kern $\pi=U$, Bild $\pi=W$ und somit $V/_U = V/_{\text{Kern}\,\pi}$ isomorph zu W. Analog beweist man die andere Behauptung. ∎

Bemerkungen und Definitionen. (a) Die Abbildung π heißt *Projektion* von V auf den Untervektorraum W. Sie hat die Eigenschaft, daß $\pi^2=\pi$ gilt. Allgemein nennt man jede lineare Abbildung $\pi: V\longrightarrow V$ mit $\pi^2=\pi$ eine *Projektion.*

(b) Ist $\pi : V \longrightarrow V$ eine Projektion, so gilt $V = \text{Kern } \pi \oplus \text{Bild } \pi$.

Beweis. Sei $x \in V$. Wir setzen $x_1 := x - \pi(x)$ und $x_2 := \pi(x)$. Dann gilt $x_2 \in \text{Bild } \pi$ und $\pi(x_1) = \pi(x - \pi(x)) = \pi(x) - \pi^2(x) = o$, d.h. $x_1 \in \text{Kern } \pi$. Somit ist $x = x_1 + x_2$ und $x_1 \in \text{Kern } \pi$, $x_2 \in \text{Bild } \pi$.

Die Summe ist auch direkt: Sei $x \in \text{Kern } \pi \cap \text{Bild } \pi$. Dann gilt $\pi(x) = o$ und es existiert ein $y \in V$ mit $x = \pi(y)$. Daraus folgt $x = \pi(y) = \pi^2(y) = \pi(x) = o$. ■

Aus $V = \text{Kern } \Phi \oplus \text{Bild } \Phi$ muß aber umgekehrt nicht folgen, daß Φ eine Projektion ist. So sind beispielsweise die bijektiven linearen Abbildungen von V, die von der Identität verschieden sind, keine Projektionen.

Wir wollen nun zeigen, wie lineare Abbildungen auf Basen wirken. Zur Vorbereitung dient der folgende Satz.

Satz 4. *Es seien V, W $\mathbb{K}$-Vektorräume, B eine Basis von V sowie $\Phi' : B \longrightarrow W$ eine beliebige Abbildung. Dann kann Φ' auf genau eine Weise zu einer linearen Abbildung $\Phi : V \longrightarrow W$ fortgesetzt werden. Insbesondere ist jede lineare Abbildung $\Phi : V \longrightarrow W$ durch ihre Werte auf B eindeutig festgelegt.*

Beweis. Jedes $x \in V$ besitzt nach der Bemerkung von S. 104/105 eine eindeutige Darstellung $x = a_1 v_1 + \cdots + a_n v_n$, $v_i \in B$, $a_i \in \mathbb{K}$. Wir setzen

$$(*) \qquad \Phi(x) := a_1 \Phi'(v_1) + \cdots + a_n \Phi'(v_n).$$

Dann ist Φ linear und $\Phi = \Phi'$ auf B. Ist $\Psi : V \longrightarrow W$ linear und gilt $\Psi = \Phi'$ auf B, so folgt aus (*), daß $\Psi = \Phi$ auf ganz V gilt. ■

Satz 5. *Es seien V, W $\mathbb{K}$-Vektorräume,* dim $V = n$, *$B = \{v_1,\dots,v_n\}$ eine Basis von V und $\Phi : V \longrightarrow W$ linear. Dann gilt:*

(a) *Φ ist genau dann injektiv, wenn die Vektoren $\Phi(v_1),\dots,\Phi(v_n)$ linear unabhängig sind.*

(b) *Φ ist genau dann surjektiv, wenn die Vektoren $\Phi(v_1),\dots,\Phi(v_n)$ ein Erzeugendensystem von W bilden.*

(c) *Φ ist genau dann bijektiv, wenn die Vektoren $\Phi(v_1),\ldots,\Phi(v_n)$ linear unabhängig und ein Erzeugendensystem von W sind, also eine Basis von W bilden.*

Beweis. (a) Die Vektoren $\Phi(v_1),\ldots,\Phi(v_n)$ sind genau dann linear unabhängig, wenn aus $a_1\,\Phi(v_1)+\ldots+a_n\,\Phi(v_n)=o$ stets $a_1=\cdots=a_n=0$ folgt. Weil $v_1,\ldots,v_n$ linear unabhängig sind, ist dies äquivalent zu der Aussage, daß aus $\Phi\,(a_1\,v_1+\cdots+a_n\,v_n)=o$ stets $a_1\,v_1+\cdots+a_n\,v_n=o$ folgt, daß also Kern $\Phi=\{o\}$ gilt und Φ somit injektiv ist.

(b) $\Phi(v_1),\ldots,\Phi(v_n)$ bilden genau dann ein Erzeugendensystem von W, wenn jeder Vektor $y\in W$ sich als Linearkombination in der Form

$$y=a_1\,\Phi(v_1)+\ldots+a_n\,\Phi(v_n)=\Phi\,(a_1\,v_1+\cdots+a_n\,v_n)$$

darstellen läßt. Weil $v_1,\ldots,v_n$ ein Erzeugendensystem von V bilden, ist dies wiederum äquivalent dazu, daß es zu jedem $y\in W$ ein $x\in V$ gibt mit $\Phi(x)=y$, daß also Φ surjektiv ist.

(c) folgt aus (a) und (b). ∎

Ist $\dim V=\dim W=n$, so sind nach Korollar 2.9 und Korollar 2.11 die Aussagen in Satz 5 (a),(b),(c) auf der rechten Seite alle äquivalent. Also erhalten wir direkt das folgende Isomorphiekriterium.

Korollar 6. *Es seien V,W $\mathbb{K}$-Vektorräume mit* $\dim V=\dim W=n$ *und Φ eine lineare Abbildung von V nach W. Dann sind folgende Aussagen äquivalent:*

(a) *Φ ist bijektiv,*

(b) *Φ ist surjektiv,*

(c) *Φ ist injektiv.*

Bemerkung. Für $\dim V=\dim W=\infty$ ist dieser Satz nicht mehr richtig. So ist z.B. der Ableitungsoperator aus Beispiel (d) von S. 134 zwar surjektiv, aber nicht injektiv.

Nun können wir einen Zusammenhang zwischen dem Isomorphiebegriff und der Dimension herstellen, falls die betrachteten $\mathbb{K}$-Vektorräume endlich dimensional sind.

Satz 7. *Es seien V und W endlich dimensionale $\mathbb{K}$–Vektorräume. Dann gilt:*

Die Vektorräume V und W sind genau dann isomorph, wenn sie dieselbe Dimension haben.

Beweis. Es seien $\dim V = n$, $\{v_1,\dots,v_n\}$ eine Basis von V und $\Phi : V \longrightarrow W$ ein Isomorphismus. Nach Satz 5 bilden $\Phi(v_1),\dots,\Phi(v_n)$ eine Basis von W. Also gilt $\dim V = \dim W$.

Ist umgekehrt $\dim V = \dim W = n$, so wählen wir Basen $\{v_1,\dots,v_n\}$ von V und $\{w_1,\dots,w_n\}$ von W. Nach Satz 4 existiert eine lineare Abbildung $\Phi : V \longrightarrow W$ mit $\Phi(v_i) = w_i$, $i = 1,\dots,n$. Nach Satz 5 ist Φ ein Isomorphismus. ■

Jeder n–dimensionale $\mathbb{K}$–Vektorraum V ist also zu $\mathbb{K}^n$ isomorph. Wir werden deshalb in Zukunft oft den $\mathbb{K}^n$ als Prototyp des n–dimensionalen $\mathbb{K}$–Vektorraumes zugrunde legen.

Bemerkung. Für unendlich dimensionale Vektorräume gilt der obige Satz nur noch teilweise. Zwar besitzen isomorphe Vektorräume stets dieselbe Dimension, aber die unendlich dimensionalen Vektorräume $\mathbb{K}[X]$ und $\mathbb{K}^{\mathbb{N}}$ sind zum Beispiel nicht isomorph.

Satz 8 (Dimensionssatz für lineare Abbildungen). *Es seien V, W $\mathbb{K}$–Vektorräume und $\Phi : V \longrightarrow W$ linear. Dann gilt:*

$$\dim \operatorname{Kern} \Phi + \dim \operatorname{Bild} \Phi = \dim V .$$

Beweis. Nach Korollar 2 gilt

$$V/_{\operatorname{Kern} \Phi} \cong \operatorname{Bild} \Phi .$$

Daraus folgt

$$\dim V/_{\operatorname{Kern} \Phi} = \dim \operatorname{Bild} \Phi$$

und nach Satz 2.23 somit

$$\dim V = \dim V/_{\operatorname{Kern} \Phi} + \dim \operatorname{Kern} \Phi = \dim \operatorname{Bild} \Phi + \dim \operatorname{Kern} \Phi. \qquad ■$$

§ 2 Vektorräume linearer Abbildungen

Zu Beginn unserer Betrachtungen über Vektorräume haben wir schon Beispiele dieser Strukturen kennengelernt, bei denen die Elemente Abbildungen waren, etwa $W^A = \{f: A \longrightarrow W \mid f \text{ Abbildung}\}$, wobei A eine nichtleere Menge ist und W ein $\mathbb{K}$–Vektorraum. Von besonderem Interesse ist der Fall, daß die Menge A ebenfalls ein Vektorraum ist und $f: A \longrightarrow W$ linear.

Satz 9. *Es seien V, W, X $\mathbb{K}$–Vektorräume. Dann gilt:*

(a) *Sind $\Phi : V \longrightarrow W$ und $\Psi : W \longrightarrow X$ linear, so auch $\Psi \circ \Phi : V \longrightarrow X$.*

(b) *Ist $\Phi : V \longrightarrow W$ ein Isomorphismus, so auch $\Phi^{-1} : W \longrightarrow V$.*

(c) *Sind $\Phi, \Psi : V \longrightarrow W$ linear, so auch $\Phi + \Psi$ und $a\,\Phi$, $a \in \mathbb{K}$.*

Beweis. (a) und (c) sind trivial und werden als Übung überlassen.

(b) Es fehlt nur noch die Linearität von Φ^{-1}: Aus $\Phi\,(\Phi^{-1}(a\,x + b\,y)) = a\,x + b\,y = a\,\Phi\,(\Phi^{-1}(x)) + b\,\Phi\,(\Phi^{-1}(y)) = \Phi(a\,\Phi^{-1}(x) + b\,\Phi^{-1}(y))$ für alle $x, y \in W$ und alle $a, b \in \mathbb{K}$, folgt wegen der Injektivität $\Phi^{-1}(a\,x + b\,y) = a\,\Phi^{-1}(x) + b\,\Phi^{-1}(y)$. ∎

Bemerkungen und Bezeichnungen. **(a)** Wegen Satz 9 (a),(b) ist die Isomorphie von $\mathbb{K}$–Vektorräumen transitiv und symmetrisch.

(b) Nach Satz 9 (c) ist die Menge $\{\Phi : V \longrightarrow W \mid \Phi \text{ linear}\}$ ein Untervektorraum des $\mathbb{K}$–Vektorraumes W^V; Bezeichnung: $\operatorname{Hom}(V, W)$.

(c) Nach Satz 9 (a),(c) ist $\operatorname{Hom}(V, V)$ ein $\mathbb{K}$–Vektorraum und ein Ring mit Eins. Außerdem gilt in $\operatorname{Hom}(V, V)$:

$$a\,(\Phi \circ \Psi) = (a\,\Phi) \circ \Psi = \Phi \circ (a\,\Psi) \quad , a \in \mathbb{K} .$$

Eine solche Struktur heißt *Algebra* ($\mathbb{K}$–Algebra). In der Literatur wird auch oft $\operatorname{End}(V)$ statt $\operatorname{Hom}(V, V)$ geschrieben.

(d) Nach Satz 9 (a),(b) ist die Teilmenge der Automorphismen in $\operatorname{Hom}(V, V)$ bezüglich der Komposition "$\circ$" eine Gruppe; Bezeichnung: $\operatorname{Aut}(V)$.

Sind V und W endlich dimensional, so gilt dies aufgrund des folgenden Satzes auch für Hom (V, W).

Satz 10. *Es seien V und W endlich dimensionale $\mathbb{K}$-Vektorräume. Dann ist auch der Vektorraum* Hom (V, W) *endlich dimensional, und es gilt:*

$$\dim \operatorname{Hom}(V, W) = \dim V \cdot \dim W.$$

Beweis. Für $\dim V = 0$ oder $\dim W = 0$ ist der Satz trivial. Seien nun $\dim V = n$, $\dim W = m$ und $\{v_1,\ldots,v_n\}$ bzw. $\{w_1,\ldots,w_m\}$ Basen von V bzw. W. Durch

$$\Phi_{ij}(v_k) := \delta_{jk}\, w_i \qquad (i = 1,\ldots,m\,;\ j,k = 1,\ldots,n)$$

werden nach Satz 4 $m \cdot n$ lineare Abbildungen $\Phi_{ij} : V \longrightarrow W$ definiert, von denen wir nun zeigen wollen, daß sie eine Basis von Hom (V, W) bilden.

1. Schritt : $\Phi_{11},\ldots,\Phi_{mn}$ sind linear unabhängig (und damit auch paarweise verschieden): Aus

$$\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} \Phi_{ij} = O \quad , a_{ij} \in \mathbb{K} ,$$

erhalten wir für alle $k = 1,\ldots,n$:

$$o = O(v_k) = \Big(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} \Phi_{ij}\Big)(v_k) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} \Phi_{ij}(v_k)$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}\, \delta_{jk}\, w_i = \sum_{i=1}^{m} a_{ik}\, w_i \,.$$

Aus der linearen Unabhängigkeit der Vektoren $w_1,\ldots,w_m$ folgt somit $a_{ik} = 0$ für $i = 1,\ldots,m$ und $k = 1,\ldots,n$. Also sind die Abbildungen $\Phi_{11},\ldots,\Phi_{mn}$ linear unabhängig.

2. Schritt : $\{\Phi_{ij} \mid i = 1,\ldots,m\ ;\ j = 1,\ldots,n\}$ ist ein Erzeugendensystem des Vektorraumes Hom (V, W): Sei $\Phi \in$ Hom (V, W) beliebig. Dann gibt es für jedes $k = 1,\ldots,n$ Zahlen $a_{1k},\ldots,a_{mk} \in \mathbb{K}$ mit

$$\Phi(v_k) = \sum_{i=1}^{m} a_{ik} w_i .$$

Daraus folgt für $k = 1,\dots,n$

$$\Big(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} \Phi_{ij}\Big)(v_k) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} \Phi_{ij}(v_k)$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} \delta_{jk} w_i = \sum_{i=1}^{m} a_{ik} w_i = \Phi(v_k) ,$$

also nach Satz 4

$$\Phi = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} \Phi_{ij} .$$

Somit ist $\{\Phi_{ij} \mid i = 1,\dots,m \,;\, j = 1,\dots,n\}$ eine Basis von Hom (V,W), und es gilt außerdem $\dim \operatorname{Hom}(V,W) = n \cdot m = \dim V \cdot \dim W$. ■

Wir wenden uns nun einem wichtigen Spezialfall zu. Da der Körper $\mathbb{K}$ selbst ein $\mathbb{K}$-Vektorraum ist, können wir $W = \mathbb{K}$ setzen. Statt Hom $(V,\mathbb{K})$ schreiben wir V^*, und statt Φ schreiben wir häufig x^*.

V^* heißt der *Dualraum* von V, und $x^* \in V^*$ heißt *lineares Funktional* oder *Linearform* auf V.

Beispiele. **(a)** Die Beispiele (e) und (f) aus § 1 enthielten Linearformen, nämlich Integral und Grenzwert.

(b) Es seien A eine nichtleere Menge, $t_0 \in A$ und $V = \mathbb{K}^A$. Die durch

$$\Phi : \mathbb{K}^A \longrightarrow \mathbb{K},\ f \longmapsto \Phi(f) := f(t_0)$$

erklärte Abbildung ist eine Linearform (*Auswertungsfunktional*).

(c) Die Abbildung $\Phi : \mathbb{K}^{n\times n} \longrightarrow \mathbb{K}$, $A = (\!(a_{ij})\!) \longmapsto \sum_{i=1}^{n} a_{ii} =:$ Spur A ist eine Linearform.

Bemerkungen und Definitionen. **(a)** Gilt $\dim V = n$, so ist nach Satz 10 auch $\dim V^* = n$, also $V \cong V^*$.

Ist $B = \{v_1,\dots,v_n\}$ eine Basis von V, so wird, wenn wir in $\mathbb{K}$ als Basis $\{1\}$ wählen, nach dem Beweis von Satz 10 durch

$$v_j^*(v_k) = \delta_{jk}, \qquad j,k = 1,\dots,n$$

eine Basis $B^* = \{v_1^*,\dots,v_n^*\}$ von V^* festgelegt. B^* heißt die zur Basis B *duale Basis*. Sie ist nach Satz 4 eindeutig bestimmt. Der Basisvektor $v_j^* \in B^*$ heißt *j–tes Koordinatenfunktional* bezüglich der Basis B.

Dieser Name erklärt sich dadurch, daß für alle $x \in V$ und alle $x^* \in V^*$ gilt:

$$x = v_1^*(x)\, v_1 + \cdots + v_n^*(x)\, v_n$$

$$x^* = x^*(v_1)\, v_1^* + \cdots + x^*(v_n)\, v_n^*.$$

(b) Speziell für $V = \mathbb{K}^n$ und die Standardbasis $B = \{e_1,\dots,e_n\}$ erhalten wir, daß das j–te Koordinatenfunktional e_j^* durch das n–Tupel

$$(e_j^*(e_1),\dots,e_j^*(e_n)) = (0,\dots,0,\underbrace{1}_{j\text{–te Stelle}},0,\dots,0)$$

festgelegt ist und ein beliebiges $x^* \in (\mathbb{K}^n)^*$ wegen $x^* = \sum\limits_{j=1}^{n} x_j\, e_j^*$ und $x^*(e_i) = x_i$ durch das n–Tupel $(x_1,\dots,x_n)$. Weil wir n–Tupel und Spaltenvektoren identifizieren, können wir e_j^* als Standardbasisvektor e_j auffassen und x^* als Spaltenvektor $x = (x_1 \cdots x_n)^{\mathsf{T}}$. Damit wird $(\mathbb{K}^n)^*$ in natürlicher Weise mit $\mathbb{K}^n$ identifiziert und B^* mit B. Die Wirkung von x^* auf

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

wird dann durch

$$x^*(y) = x_1 y_1 + \cdots + x_n y_n = x^\top y =: <x,y>$$

gegeben. Das so definierte Produkt $<\cdot,\cdot>$ heißt das *Standardskalarprodukt* auf $\mathbb{K}^n$. Es wird in Kapitel 5 weiter behandelt.

(c) Unendlich dimensionale Vektorräume sind im allgemeinen nicht zu ihrem Dualraum isomorph:

Sei beispielsweise $V = \mathbb{R}[X]$, und sei $B = \{1, X, X^2, X^3, \ldots\}$ die kanonische Basis von V. Dann ist $x^* \in V^*$ eindeutig bestimmt durch die Folge $(x^*(1), x^*(X), x^*(X^2), \ldots) \in \mathbb{R}^{\mathbb{N}_0}$. Ist umgekehrt $(a_0, a_1, a_2, \ldots)$ eine Folge reeller Zahlen, so wird durch $x^*(X^i) = a_i$, $i = 0,1,2,\ldots$, eine Linearform auf V definiert. Die Abbildung

$$V^* \longrightarrow \mathbb{R}^{\mathbb{N}_0}, \quad x^* \longmapsto (x^*(1), x^*(X), x^*(X^2), \ldots)$$

ist dann ein Isomorphismus. Weil V und $\mathbb{R}^{\mathbb{N}_0}$ nicht isomorph sind, ist V damit auch nicht zu V^* isomorph.

Nun gehen wir einen Schritt weiter und betrachten den Dualraum von V^*. Statt $(V^*)^*$ schreiben wir V^{**} und nennen V^{**} den *Bidualraum* von V.

Es gibt einen natürlichen Zusammenhang zwischen dem Vektorraum V und seinem Bidualraum V^{**}. Wir erhalten ihn, indem wir $x^*(x)$, $x \in V$, $x^* \in V^*$, betrachten und bei festem x die Linearform $x^* \in V^*$ variieren lassen.

Bemerkung. Sei $x \in V$. Gilt $x^*(x) = 0$ für alle $x^* \in V^*$, so ist $x = o$.

Beweis. Ist $x \neq o$, so können wir x zu einer Basis B von V ergänzen und dann eine nichttriviale Linearform $x^* \in V^*$ definieren durch $x^*(x) = 1$ und $x^*(x') = 0$ für alle $x' \in B \setminus \{x\}$. ∎

Satz 11. *Es seien V ein $\mathbb{K}$-Vektorraum und V^{**} der Bidualraum. Dann ist die kanonische Abbildung*

$$\begin{aligned} \Psi : V &\longrightarrow V^{**} \\ x &\longmapsto x^{**}, \end{aligned}$$

wo $x^{**}(x^*) := x^*(x)$ *für alle* $x^* \in V^*$, *ein injektiver Homomorphismus. Ist* V *endlich dimensional, so ist* Ψ *ein Isomorphismus.*

Beweis. Ψ ist linear: Für alle x, $y \in V$ und alle $a,b \in \mathbb{K}$ ist $\Psi(a\,x + b\,y)(x^*) = x^*(a\,x + b\,y) = a\,x^*(x) + b\,x^*(y) = a\,\Psi(x)(x^*) + b\,\Psi(y)(x^*) = (a\,\Psi(x) + b\,\Psi(y))(x^*)$. Dies gilt für alle $x^* \in V^*$, also folgt $\Psi(a\,x + b\,y) = a\,\Psi(x) + b\,\Psi(y)$.

Ψ ist injektiv: Sei $x \in V$ und $\Psi(x)$ der Nullvektor in V^{**}. Nach Definition von Ψ erhalten wir $x^*(x) = \Psi(x)(x^*) = 0$ für alle $x^* \in V^*$. Nach der obigen Bemerkung folgt $x = o$, also ist Ψ injektiv.

Ist V endlich dimensional, so ist Ψ nach Korollar 6 dann auch bijektiv. ∎

Ist V endlich dimensional, so werden wir wegen des natürlichen Isomorphismus zwischen V und V^{**} in Zukunft V mit seinem Bidualraum V^{**} identifizieren.

Ist V unendlich dimensional, so ist Ψ eine Einbettung von V in V^{**}, und wir können deswegen V als Untervektorraum von V^{**} auffassen.

Zum Abschluß dieses Paragraphen wollen wir noch jeder linearen Abbildung $\Phi : V \longrightarrow W$ eine lineare Abbildung $\Phi^{\top} : W^* \longrightarrow V^*$ zuordnen.

Definition. Es seien V, W $\mathbb{K}$-Vektorräume und $\Phi : V \longrightarrow W$ linear. Dann heißt die durch

$$\begin{aligned} \Phi^{\top} : W^* &\longrightarrow V^* \\ y^* &\longmapsto y^* \circ \Phi \end{aligned}$$

erklärte lineare Abbildung $\Phi^{\top}$ die *transponierte* oder *duale Abbildung* von Φ.

Bemerkung. Der Name transponierte Abbildung erklärt sich von selbst, wenn wir im nächsten Paragraphen den Zusammenhang zwischen linearen Abbildungen und Matrizen behandeln.

Für transponierte Abbildungen gelten folgende Rechenregeln.

Satz 12. *Es seien V, W, X $\mathbb{K}$–Vektorräume. Dann gilt:*

(a) $(\mathrm{id}_V)^\top = \mathrm{id}_{V^*}$,

(b) $(\Phi + \Psi)^\top = \Phi^\top + \Psi^\top$, $(a\,\Phi)^\top = a\,\Phi^\top$ *für alle* $\Phi, \Psi \in \mathrm{Hom}\,(V, W)$ *und alle* $a \in \mathbb{K}$,

(c) $(\Psi \circ \Phi)^\top = \Phi^\top \circ \Psi^\top$ *für alle* $\Phi \in \mathrm{Hom}\,(V, W)$ *und alle* $\Psi \in \mathrm{Hom}\,(W, X)$,

(d) $\Phi \in \mathrm{Hom}\,(V, W)$ *ist genau dann Isomorphismus, wenn* $\Phi^\top \in \mathrm{Hom}\,(W^*, V^*)$ *Isomorphismus ist. In diesem Fall gilt* $(\Phi^{-1})^\top = (\Phi^\top)^{-1}$.

Beweis. (a) $(\mathrm{id}_V)^\top(x^*) = x^* \circ \mathrm{id}_V = x^*$ für alle $x^* \in V^*$.

(b) $(\Phi + \Psi)^\top(y^*) = y^* \circ (\Phi + \Psi) = y^* \circ \Phi + y^* \circ \Psi = \Phi^\top(y^*) + \Psi^\top(y^*) = (\Phi^\top + \Psi^\top)(y^*)$ für alle $y^* \in W^*$. Analog folgt $(a\,\Phi)^\top(y^*) = y^* \circ (a\,\Phi) = a\,(y^* \circ \Phi) = a\,(\Phi^\top(y^*)) = (a\,\Phi^\top)(y^*)$ für alle $y^* \in W^*$.

(c) Es ist $\Psi \circ \Phi \in \mathrm{Hom}\,(V, X)$ und $(\Psi \circ \Phi)^\top \in \mathrm{Hom}\,(X^*, V^*)$. Dann gilt für alle $z^* \in X^*$: $(\Psi \circ \Phi)^\top(z^*) = z^* \circ (\Psi \circ \Phi)$ und $(\Phi^\top \circ \Psi^\top)(z^*) = \Phi^\top(\Psi^\top(z^*)) = \Psi^\top(z^*) \circ \Phi = (z^* \circ \Psi) \circ \Phi = z^* \circ (\Psi \circ \Phi)$.

(d) Ist $\Phi \in \mathrm{Hom}(V, W)$ bijektiv, so existiert Φ^{-1} und es gilt $\Phi^{-1} \circ \Phi = \mathrm{id}_V$, $\Phi \circ \Phi^{-1} = \mathrm{id}_W$. Daraus folgt $\Phi^\top \circ (\Phi^{-1})^\top = (\Phi^{-1} \circ \Phi)^\top = (\mathrm{id}_V)^\top = \mathrm{id}_{V^*}$ und $(\Phi^{-1})^\top \circ \Phi^\top = (\Phi \circ \Phi^{-1})^\top = (\mathrm{id}_W)^\top = \mathrm{id}_{W^*}$. Also ist auch $\Phi^\top$ bijektiv, und es gilt $(\Phi^\top)^{-1} = (\Phi^{-1})^\top$.

Die Umkehrung ergibt sich aus folgenden zwei Behauptungen:

1. Beh.: Ist $\Phi^\top \in \mathrm{Hom}(W^*, V^*)$ surjektiv, so ist $\Phi \in \mathrm{Hom}(V, W)$ injektiv.

Beweis. Sei $\Phi(x) = o$. Weil $\Phi^\top$ surjektiv ist, gibt es zu jedem $x^* \in V^*$ ein $y^* \in W^*$ mit $\Phi^\top(y^*) = x^*$. Damit folgt $x^*(x) = \Phi^\top(y^*)(x) = y^*(\Phi(x)) = 0$ für alle $x^* \in V^*$. Nach der obigen Bemerkung ist dann $x = o$, also ist Φ injektiv.

2. Beh.: Ist $\Phi^\top \in \mathrm{Hom}(W^*, V^*)$ injektiv, so ist $\Phi \in \mathrm{Hom}(V, W)$ surjektiv.

Beweis. Wir führen einen Widerspruchsbeweis und nehmen an, daß Φ nicht surjektiv sei. Somit gilt $\Phi(V) \neq W$. Wir ergänzen nun eine Basis C von $\Phi(V)$ zu einer Basis C' von W, wählen ein festes $y \in C' \backslash C$ und definieren eine Linearform $y^* \in W^*$ durch $y^*(y) = 1$ und $y^*(y') = 0$ für $y' \in C' \backslash \{y\}$. Dann gilt $\Phi^\top(y^*)(x) = y^*(\Phi(x)) = 0$ für alle $x \in V$, also $\Phi^\top(y^*) = o^*$. Wegen der Injektivität von $\Phi^\top$ folgt hieraus, daß y^* die Nullform ist, ein Widerspruch zur Definition von y^*. Somit ist Φ surjektiv. ∎

§ 3 Darstellung linearer Abbildungen durch Matrizen

Ein n-dimensionaler $\mathbb{K}$-Vektorraum V ist nach Satz 7 isomorph zum $\mathbb{K}^n$, und wenn wir eine Basis $\{v_1,\dots,v_n\}$ von V gegeben haben, können wir einen Isomorphismus konkret angeben. Da es im folgenden auf die Reihenfolge der Basisvektoren ankommt, schreiben wir

$$B = (v_1,\dots,v_n)$$

und nennen dieses n-Tupel von Vektoren eine *geordnete Basis* von V.

Bemerkung. Bei einer geordneten Basis $(v_1,\dots,v_n)$ sind die Vektoren $v_1,\dots,v_n$ paarweise verschieden.

Definition. Es seien V ein $\mathbb{K}$-Vektorraum und $B = (v_1,\dots,v_n)$ eine geordnete Basis von V. Sei $x \in V$ und

$$x = x_1 v_1 + \cdots + x_n v_n\,,\ x_i \in \mathbb{K}$$

die eindeutige Darstellung von x bezüglich B. Dann heißen die Zahlen $x_1,\dots,x_n$ die *Koordinaten* von x bezüglich B und der Vektor

$$\hat{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{K}^n$$

heißt *Koordinatenvektor* oder *Koordinatendarstellung* von x bezüglich B.

Bemerkung. Die Abbildung $V \longrightarrow \mathbb{K}^n$, die jedem $x \in V$ den zugehörigen Koordinatenvektor $\hat{x}$ bezüglich B zuordnet, ist linear und bildet insbesondere v_i auf den Standardbasisvektor $e_i \in \mathbb{K}^n$ ab, $i = 1,\dots,n$. Sie ist daher ein Isomorphismus.

Nun können wir einen Zusammenhang zwischen linearen Abbildungen endlich dimensionaler Vektorräume und Matrizen herstellen.

Definition. Es seien V, W $\mathbb{K}$-Vektorräume, $B = (v_1,\dots,v_n)$ bzw. $C = (w_1,\dots,w_m)$

geordnete Basen von V bzw. W und $\Phi : V \longrightarrow W$ linear. Für $j \in \{1,\dots,n\}$ sei

$$\Phi(v_j) = a_{1j}\, w_1 + \cdots + a_{mj}\, w_m$$

die eindeutige Darstellung von $\Phi(x_j)$ bezüglich C. Dann heißt die (m,n)-Matrix

$$A_\Phi := (\!(a_{ij})\!)$$

Abbildungsmatrix von Φ (bezüglich der geordneten Basen B von V und C von W).

Bemerkungen (a) Die Koordinaten von $\Phi(v_j)$ bilden gerade die j-te Spalte von A_Φ.

(b) Es ist $\mathrm{Rg}\, A_\Phi = \dim \mathrm{Bild}\, \Phi$. Aus diesem Grunde verwenden wir statt $\dim \mathrm{Bild}\, \Phi$ auch die Bezeichnung $\mathrm{Rg}\, \Phi$ und sprechen vom *Rang* der Abbildung Φ.

Satz 13. *Es seien V, W $\mathbb{K}$-Vektorräume*, $\dim V = n$, $\dim W = m$ *und* $B = (v_1,\dots,v_n)$ *bzw.* $C = (w_1,\dots,w_m)$ *geordnete Basen von V bzw. W. Dann ist die Abbildung*

$$\Phi \longmapsto A_\Phi ,$$

die jedem $\Phi \in$ Hom (V, W) *die Abbildungsmatrix A_Φ bezüglich der Basen B, C zuordnet, ein Isomorphismus von* Hom (V, W) *auf* $\mathbb{K}^{m\times n}$.

Beweis. Ist $A_\Phi = (\!(a_{ij})\!)$, $A_{\Phi'} = (\!(b_{ij})\!)$, so gilt für alle $a, a' \in \mathbb{K}$ und $j = 1,\dots,n$

$$(a\,\Phi + a'\,\Phi')(v_j) = a\,\Phi(v_j) + a'\,\Phi'(v_j)$$

$$= a \sum_{i=1}^{m} a_{ij}\, w_i + a' \sum_{i=1}^{m} b_{ij}\, w_i = \sum_{i=1}^{m} (a\, a_{ij} + a'\, b_{ij})\, w_i .$$

Also folgt $A_{a\,\Phi + a'\Phi'} = a\,A_\Phi + a' A_{\Phi'}$, und die Abbildung $\Phi \longmapsto A_\Phi$ ist daher linear.

Für jedes $A = (\!(a_{ij})\!) \in \mathbb{K}^{m\times n}$ wird durch

$$\Phi(v_j) = \sum_{i=1}^{m} a_{ij}\, w_i , \; j = 1,\dots,n,$$

eine lineare Abbildung $\Phi : V \longrightarrow W$ eindeutig festgelegt (Satz 4), die $A_\Phi = A$ erfüllt. Also ist die Abbildung $\Phi \longmapsto A_\Phi$ ein Isomorphismus. ■

Bemerkung. Aus diesem Satz folgt wieder das schon bekannte Ergebnis über die Dimension von $\mathrm{Hom}(V,W)$: $\dim \mathrm{Hom}\,(V,W) = \dim V \cdot \dim W$.

Die Wirkung der Abbildungsmatrix wird durch den folgenden Satz deutlich.

Satz 14. *Es seien V, W endlich dimensionale $\mathbb{K}$-Vektorräume mit geordneten Basen B bzw. C, $\Phi : V \longrightarrow W$ linear und A_Φ die Abbildungsmatrix von Φ. Weiterhin sei $\hat{x}$ der Koordinatenvektor von $x \in V$ und $\hat{y}$ der Koordinatenvektor von $y = \Phi(x)$. Dann gilt*

$$\hat{y} = A_\Phi \cdot \hat{x}.$$

Beweis. Seien $B = (v_1,\dots,v_n)$, $C = (w_1,\dots,w_m)$ und $A_\Phi = ((a_{ij}))$. Für $x = x_1 v_1 + \dots + x_n v_n$ folgt $\Phi(x) = y_1\, w_1 + \dots + y_m\, w_m$ sowie

$$\Phi(x) = \sum_{j=1}^{n} x_j\, \Phi(v_j) = \sum_{j=1}^{n} x_j \,(\sum_{i=1}^{m} a_{ij}\, w_i) = \sum_{i=1}^{m} (\sum_{j=1}^{n} a_{ij}\, x_j)\, w_i\,.$$

Also gilt

$$y_i = \sum_{j=1}^{n} a_{ij}\, x_j$$

für $i = 1,\dots,m$, d.h.

$$\hat{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = A_\Phi \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = A_\Phi\, \hat{x}\,.$$ ■

Bemerkung. Im Fall $V = \mathbb{K}^n$, $W = \mathbb{K}^m$ gilt bezüglich der Standardbasen

$$A_\Phi = (\Phi(e_1) \,|\cdots|\, \Phi(e_n)) \quad \text{und} \quad \Phi(x) = A_\Phi \cdot x \,,\quad x \in \mathbb{K}^n.$$

Beispiele. **(a)** Es seien V, W reelle Vektorräume , (v_1, v_2, v_3) eine geordnete Basis von V sowie (w_1, w_2, w_3, w_4) eine geordnete Basis von W. Weiter sei $\Phi \in \operatorname{Hom}(V, W)$ festgelegt durch

$$\begin{aligned} \Phi(v_1) &= w_1 - w_2 + 3\,w_3 - w_4 \\ \Phi(v_2) &= 2\,w_1 + w_2 + 7\,w_3 + 2\,w_4 \\ \Phi(v_3) &= 3\,w_2 + w_3 + 4\,w_4\,. \end{aligned}$$

Die Abbildungsmatrix von Φ bezüglich der gegebenen Basen ist

$$A_\Phi = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 1 & 3 \\ 3 & 7 & 1 \\ -1 & 2 & 4 \end{bmatrix}.$$

Wir wollen Kern Φ berechnen. Wegen Kern $\Phi = \{x \in V \mid A_\Phi\, \hat{x} = o\}$ müssen wir das homogene lineare Gleichungssystem $A_\Phi\, \hat{x} = o$ lösen. Es gilt

$$L_h = \{a \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \mid a \in \mathbb{R}\},$$

also folgt

$$\text{Kern}\, \Phi = \{a\,(2\,v_1 - v_2 + v_3) \mid a \in \mathbb{R}\}.$$

(b) Sei $V \subset \mathbb{R}[X]$ der Untervektorraum der Polynome vom Grad $\leq n-1$, $n \in \mathbb{N}$. Dann gilt $\dim V = n$. Der Ableitungsoperator $D: V \longrightarrow V$ hat bezüglich der Standardbasis $(1, X, \ldots, X^{n-1})$ von V die Abbildungsmatrix

$$A_D = \begin{bmatrix} 0 & 1 & 0 & \cdots & & 0 \\ \vdots & \cdot & 2 & \cdot & & \vdots \\ \vdots & & \cdot & \cdot & \cdot & \vdots \\ \vdots & & & \cdot & \cdot & 0 \\ \vdots & & & & \cdot & n-1 \\ 0 & \cdots & & & \cdots & 0 \end{bmatrix}$$

Wir wollen nun zeigen, wie sich das Verketten linearer Abbildungen, die Inversenbildung sowie der Übergang zur transponierten Abbildung bei den zugehörigen Abbildungsmatrizen auswirkt.

Satz 15. *Es seien V, W, X $\mathbb{K}$-Vektorräume mit* $\dim V = n$, $\dim W = m$, $\dim X = k$ *und mit geordneten Basen B_1, B_2, B_3. Weiter seien $\Phi : V \longrightarrow W$ und $\Phi' : W \longrightarrow X$ linear. Dann gilt:*

$$A_{\Phi' \circ \Phi} = A_{\Phi'} \cdot A_{\Phi} .$$

Beweis. Es seien $B_1 = (v_1, \ldots, v_n)$, $B_2 = (w_1, \ldots, w_m)$, $B_3 = (x_1, \ldots, x_k)$ sowie $A_\Phi = (\!(a_{ij})\!)$, $A_{\Phi'} = (\!(b_{ij})\!)$ und $A_{\Phi' \circ \Phi} = (\!(c_{ij})\!)$. Dann gilt

$$\Phi' \circ \Phi\,(v_j) = \Phi'(\Phi(v_j)) = \Phi'\Big(\sum_{i=1}^{m} a_{ij}\, w_i\Big) = \sum_{i=1}^{m} a_{ij}\, \Phi'(w_i)$$

$$= \sum_{i=1}^{m} a_{ij} \Big(\sum_{l=1}^{k} b_{li}\, x_l \Big) = \sum_{l=1}^{k} \Big(\sum_{i=1}^{m} b_{li}\, a_{ij} \Big) x_l .$$

Damit erhalten wir

$$c_{lj} = \sum_{i=1}^{m} b_{li}\, a_{ij} ,$$

also gilt die Behauptung. ∎

Satz 16. *Es seien V, W n-dimensionale $\mathbb{K}$-Vektorräume mit geordneten Basen B, C und $\Phi : V \longrightarrow W$ eine lineare Abbildung. Φ ist genau dann ein Isomorphismus, wenn die Abbildungsmatrix A_Φ von Φ bezüglich B, C regulär ist. In diesem Fall gilt für die Abbildungsmatrix von Φ^{-1} bezüglich C, B*

$$A_{\Phi^{-1}} = A_\Phi^{-1} .$$

Beweis. Ist Φ ein Isomorphismus, so folgt $\Phi^{-1} \circ \Phi = \mathrm{id}_V$, $\Phi \circ \Phi^{-1} = \mathrm{id}_W$, also nach Satz 15

$$A_{\Phi^{-1}} \cdot A_{\Phi} = A_{\mathrm{id}_V} = E_n ,$$

$$A_{\Phi} \cdot A_{\Phi^{-1}} = A_{\mathrm{id}_W} = E_n .$$

Somit gilt $A_{\Phi^{-1}} = A_{\Phi}^{-1}$. Nun sei umgekehrt A_{Φ} regulär. Wir betrachten die Abbildung $\Phi' : W \longrightarrow V$, die nach Satz 13 zu A_{Φ}^{-1} gehört. Dann folgt wieder nach Satz 15

$$E_n = A_{\Phi} \cdot A_{\Phi}^{-1} = A_{\Phi \circ \Phi'} ,$$

$$E_n = A_{\Phi}^{-1} \cdot A_{\Phi} = A_{\Phi' \circ \Phi} .$$

Nach Satz 13 ist folglich

$$\Phi \circ \Phi' = \mathrm{id}_W , \quad \Phi' \circ \Phi = \mathrm{id}_V ,$$

also $\Phi' = \Phi^{-1}$. ■

Satz 17. *Es seien V, W $\mathbb{K}$-Vektorräume mit* dim $V = n$, dim $W = m$ *und mit geordneten Basen $B = (v_1,...,v_n)$, $C = (w_1,...,w_m)$. Weiter seien $\Phi : V \longrightarrow W$ eine lineare Abbildung und A_{Φ} die Abbildungsmatrix von Φ bezüglich B,C. Ist $A_{\Phi^\top}$ die Abbildungsmatrix von $\Phi^\top$ bezüglich der dualen Basen C^* und B^* , so gilt*

$$A_{\Phi^\top} = A_{\Phi}^\top .$$

Beweis. Es gilt für $A_{\Phi^\top} = (\!(c_{kl})\!)$ nach Definition der Abbildungsmatrix

$$\Phi^\top(w_l^*) = \sum_{j=1}^{n} c_{jl}\, v_j^* , \text{ also } (\Phi^\top(w_l^*))(v_k) = \sum_{j=1}^{n} c_{jl}\, v_j^*\,(v_k) = \sum_{j=1}^{n} c_{jl}\, \delta_{jk} = c_{kl} .$$

Andererseits ist nach Definition $\Phi^\top(w_l^*) = w_l^* \circ \Phi$, also gilt mit $A_{\Phi} = (\!(a_{ij})\!)$

$$(\Phi^\top(w_l^*))(v_k) = w_l^*\,(\Phi(v_k)) = w_l^*(\sum_{i=1}^{m} a_{ik}\, w_i) = \sum_{i=1}^{m} a_{ik}\, w_l^*\,(w_i) = a_{lk} .$$

Damit folgt $c_{kl} = a_{lk}$ für alle k,l . ■

Zum Schluß dieses Abschnitts wollen wir uns noch überlegen, wie sich die Abbildungsmatrix einer linearen Abbildung $\Phi : V \longrightarrow W$ ändert, wenn man in den Vektorräumen V und W zu anderen Basen übergeht. Es seien

$$B = (v_1,\dots,v_n) \quad , \quad \tilde{B} = (\tilde{v}_1,\dots,\tilde{v}_n)$$

geordnete Basen von V und

$$C = (w_1,\dots,w_m) \quad , \quad \tilde{C} = (\tilde{w}_1,\dots,\tilde{w}_m)$$

geordnete Basen von W. Weiterhin sei $A_\Phi = (\!(a_{ij})\!) \in \mathbb{K}^{m\times n}$ die Abbildungsmatrix von $\Phi \in \operatorname{Hom}(V,W)$ bezüglich der Basen B,C und $\tilde{A}_\Phi = (\!(\tilde{a}_{ij})\!)$ sei die Abbildungsmatrix von Φ bezüglich der neuen Basen $\tilde{B},\tilde{C}$.

Die neuen Basisvektoren lassen sich mit Hilfe der alten linear kombinieren:

$$\begin{aligned} \tilde{v}_j &= s_{1j}\, v_1 + \cdots + s_{nj}\, v_n \quad , \quad j = 1,\dots,n\,, \\ \tilde{w}_k &= t_{1k}\, w_1 + \cdots + t_{mk}\, w_m \quad , \quad k = 1,\dots,m\,. \end{aligned}$$

Die Matrix $S = (\!(s_{ij})\!) \in \mathbb{K}^{n\times n}$ ist regulär, denn sie ist die Abbildungsmatrix von id_V bezüglich der Basen $\tilde{B},B$. Ebenso ist auch die Matrix $T = (\!(t_{ij})\!) \in \mathbb{K}^{m\times m}$ regulär als Abbildungsmatrix von id_W bezüglich der Basen $\tilde{C},C$.

Nun gilt für alle $j = 1,\dots,n$:

$$\Phi(\tilde{v}_j) = \sum_{k=1}^{m} \tilde{a}_{kj}\, \tilde{w}_k = \sum_{k=1}^{m} \tilde{a}_{kj} \sum_{i=1}^{m} t_{ik}\, w_i = \sum_{i=1}^{m} \Big(\sum_{k=1}^{m} t_{ik}\, \tilde{a}_{kj} \Big)\, w_i$$

bzw.

$$\Phi(\tilde{v}_j) = \sum_{k=1}^{n} s_{kj}\, \Phi(v_k) = \sum_{k=1}^{n} s_{kj} \sum_{i=1}^{m} a_{ik}\, w_i = \sum_{i=1}^{m} \Big(\sum_{k=1}^{n} a_{ik}\, s_{kj} \Big)\, w_i\,.$$

Daraus folgt für alle $j = 1,\dots,n$ und alle $i = 1,\dots,m$:

$$\sum_{k=1}^{m} t_{ik}\, \tilde{a}_{kj} = \sum_{k=1}^{n} a_{ik}\, s_{kj}$$

d.h.

$$T \tilde{A}_{\Phi} = A_{\Phi} S$$

und somit

$$\tilde{A}_{\Phi} = T^{-1} A_{\Phi} S .$$

Die Art und Weise wie A_{Φ} und $\tilde{A}_{\Phi}$ zusammenhängen, führt zu folgender Definition.

Definition. Zwei Matrizen $A, \tilde{A} \in \mathbb{K}^{m \times n}$ heißen *äquivalent*, wenn es reguläre Matrizen $S \in \mathbb{K}^{n \times n}$ und $T \in \mathbb{K}^{m \times m}$ gibt mit $\tilde{A} = T^{-1} A S$.

Zusammenfassung. Durch Basiswechsel in V (B wird ersetzt durch $\tilde{B}$) und in W (C wird ersetzt durch $\tilde{C}$) geht die Abbildungsmatrix A_{Φ} einer linearen Abbildung $\Phi : V \longrightarrow W$ in eine äquivalente Matrix $\tilde{A}_{\Phi}$ über,

$$\tilde{A}_{\Phi} = T^{-1} A_{\Phi} S ,$$

wobei die regulären Matrizen $S = (\!(s_{ik})\!)$ bzw. $T = (\!(t_{ik})\!)$ den Basiswechsel in V bzw. in W beschreiben und durch

$$\tilde{v}_j = \sum_{i=1}^{n} s_{ij} v_i ,$$

$j = 1,\ldots,n$, bzw.

$$\tilde{w}_k = \sum_{i=1}^{m} t_{ik} w_i ,$$

$k = 1,\ldots,m$, gegeben sind.

Beispiel. Die lineare Abbildung $\Phi : \mathbb{R}^3 \longrightarrow \mathbb{R}^4$ sei bezüglich der Standardbasen durch die Matrix

$$A_{\Phi} = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 1 & 3 \\ 3 & 7 & 1 \\ -1 & 2 & 4 \end{bmatrix}$$

beschrieben. Als neue Basen betrachten wir

$$\tilde{B} = \left(\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}, \begin{bmatrix}1\\0\\1\end{bmatrix}\right) \text{ in } \mathbb{R}^3 \text{ sowie } \tilde{C} = \left(\begin{bmatrix}1\\1\\0\\0\end{bmatrix}, \begin{bmatrix}1\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\0\\0\\1\end{bmatrix}\right) \text{ in } \mathbb{R}^4 .$$

Dann ist

$$S = \begin{bmatrix}1 & 0 & 1\\1 & 1 & 0\\0 & 1 & 1\end{bmatrix} , \quad T = \begin{bmatrix}1 & 1 & 0 & 1\\1 & 0 & 1 & 0\\0 & 1 & 1 & 0\\0 & 0 & 0 & 1\end{bmatrix}$$

und

$$\tilde{A}_\Phi = T^{-1} A_\Phi S = \begin{bmatrix}1/2 & 1/2 & -1/2 & -1/2\\1/2 & -1/2 & 1/2 & -1/2\\-1/2 & 1/2 & 1/2 & 1/2\\0 & 0 & 0 & 1\end{bmatrix} \begin{bmatrix}3 & 2 & 1\\0 & 4 & 2\\10 & 8 & 4\\1 & 6 & 3\end{bmatrix} = \begin{bmatrix}-4 & -4 & -2\\6 & 0 & 0\\4 & 8 & 4\\1 & 6 & 3\end{bmatrix}$$

Bemerkungen. **(a)** Die Äquivalenz von Matrizen ist eine Äquivalenzrelation auf der Menge $\mathbb{K}^{m\times n}$.

(b) Durch elementare Zeilen– oder Spaltenumformungen geht eine Matrix $A \in \mathbb{K}^{m\times n}$ in eine äquivalente Matrix A' über.

Beweis. A ist die Abbildungsmatrix der linearen Abbildung $\Phi : \mathbb{K}^n \longrightarrow \mathbb{K}^m$, $x \longmapsto A\, x$, bezüglich der Standardbasen. Elementare Zeilenumformungen entsprechen einem Basiswechsel in $\mathbb{K}^m$, elementare Spaltenumformungen einem Basiswechsel in $\mathbb{K}^n$. Daraus folgt die Behauptung. ■

(c) Jede Matrix $A \in \mathbb{K}^{m\times n}$ ist zu ihrer Gaußschen Normalform äquivalent.

Beweis. Die Aussage folgt direkt aus (b). ■

(d) Zwei Matrizen $A,B \in \mathbb{K}^{m\times n}$ sind genau dann äquivalent, wenn sie gleichen Rang haben.

Beweis. Durch elementare Zeilen– **und** Spaltenumformungen kann jede Matrix A auf die Form

$$A' = \begin{bmatrix} E_r & O \\ O & O \end{bmatrix}$$

mit $r = \operatorname{Rg} A$ gebracht werden. Nach (b) sind A und A' äquivalent. Gilt also $\operatorname{Rg} A = \operatorname{Rg} B$, so sind auch A und B äquivalent. Gilt umgekehrt $B = T^{-1} A\, S$ mit regulären Matrizen $S \in \mathbb{K}^{n\times n}$ und $T \in \mathbb{K}^{m\times m}$, so ist A Abbildungsmatrix von $\Phi : \mathbb{K}^n \longrightarrow \mathbb{K}^m$, $x \longmapsto A\, x$, bezüglich der Standardbasen in $\mathbb{K}^n$ und $\mathbb{K}^m$ und B ist die Abbildungsmatrix von Φ bezüglich der geordneten Basen $(S\, e_1,\dots,S\, e_n)$ in $\mathbb{K}^n$ und $(T\, e_1,\dots,T\, e_m)$ in $\mathbb{K}^m$. Damit folgt $\operatorname{Rg} A = \dim \operatorname{Bild} \Phi = \operatorname{Rg} B$. ∎

Zum Schluß betrachten wir noch den Fall, daß eine lineare Abbildung $\Phi : V \longrightarrow V$ bezüglich einer Basis B in V dargestellt ist. Wird die Basis B gewechselt, so geht A_Φ über in $\tilde{A}_\Phi = S^{-1} A_\Phi\, S$ (wegen $T = S$). Auch für diese Beziehung verwenden wir einen eigenen Namen:

Definition. Zwei Matrizen $A, \tilde{A} \in \mathbb{K}^{n\times n}$ heißen *ähnlich*, wenn es eine reguläre Matrix $S \in \mathbb{K}^{n\times n}$ gibt mit

$$\tilde{A} = S^{-1} A\, S.$$

Bemerkungen. **(a)** Die Ähnlichkeit von Matrizen ist ein Äquivalenzrelation auf der Menge $\mathbb{K}^{n\times n}$.

(b) Ähnliche Matrizen sind äquivalent. Umgekehrt gibt es aber äquivalente quadratische Matrizen, die nicht ähnlich sind. In jeder Klasse äquivalenter Matrizen in $\mathbb{K}^{n\times n}$ gibt es nämlich nach der obigen Bemerkung (d) eine Matrix, die Diagonalgestalt hat, aber es gibt nicht zu jeder (n,n)–Matrix eine ähnliche Diagonalmatrix. Wann dies der Fall ist, werden wir im nächsten Kapitel sehen.

§ 4 Affine Abbildungen eines Vektorraumes

In diesem Abschnitt wollen wir zunächst wieder die affine Struktur des Vektorraumes V in den Vordergrund stellen und die zugehörigen strukturverträglichen Abbildungen studieren. Es sind dies die affinen Abbildungen von V. In Kapitel 6.1 werden wir dann die gewonnenen Ergebnisse auf allgemeine affine Räume übertragen.

Der affine Unterraum $L = x_0 + U$ geht aus seinem Richtungsraum durch die Abbildung $\tau : x \longmapsto x + x_0$ hervor. Diese Abbildung von V in sich ist nur für $x_0 = o$ linear; τ ist eine bijektive affine Abbildung von V in sich. Sie heißt *Translation* mit *Translationsvektor* x_0.

Definition. Es seien V, W $\mathbb{K}$–Vektorräume, $\Phi \in \operatorname{Hom}(V,W)$ und $w \in W$. Dann heißt die Abbildung

$$\begin{aligned} \varphi : V &\longrightarrow W \\ x &\longmapsto \Phi(x) + w \end{aligned}$$

eine *affine Abbildung* von V in W. Der Vektor w heißt *Translationsvektor* von φ.

Jede bijektive affine Abbildung $\varphi : V \longrightarrow V$ heißt eine *Affinität* von V.

Bemerkungen. **(a)** Jede affine Abbildung $\varphi : V \longrightarrow W$ ist also Komposition einer linearen Abbildung $\Phi : V \longrightarrow W$ mit einer Translation τ in W, $\varphi = \tau \circ \Phi$. Die Abbildungen Φ und τ sind dabei eindeutig bestimmt.

φ ist genau dann bijektiv, wenn es die zugehörige lineare Abbildung Φ ist. In diesem Fall gilt $\varphi^{-1}(y) = \Phi^{-1}(y) - \Phi^{-1}(w)$ für alle $y \in W$, und die Umkehrfunktion $\varphi^{-1} : W \longrightarrow V$ ist daher ebenfalls affin.

(b) Die Komposition affiner Abbildungen $\varphi : V \longrightarrow W$, $\varphi' : W \longrightarrow X$ ist affin.

Beweis. Seien $\varphi(v) = \Phi(v) + w_0$, $\Phi \in \operatorname{Hom}(V,W)$, $w_0 \in W$ und $\varphi'(w) = \Phi'(w) + x_0$, $\Phi' \in \operatorname{Hom}(W,X)$, $x_0 \in X$. Dann gilt $\varphi' \circ \varphi\ (v) = \Phi' \circ \Phi\ (v) + \Phi'(w_0) + x_0$ mit $\Phi' \circ \Phi \in \operatorname{Hom}(V,X)$ und $\Phi'(w_0) + x_0 \in X$. ■

(c) Die Affinitäten von V bilden bezüglich der Komposition eine Gruppe, die *affine Gruppe* von V. Wir bezeichnen sie mit $\mathrm{Aff}(V)$.

Affine Abbildungen lassen die geometrische Struktur invariant. Genauer bedeutet dies:

Satz 18. *Es seien V, W endlich dimensionale $\mathbb{K}$-Vektorräume und $\varphi : V \longrightarrow W$ eine affine Abbildung. Dann gilt:*

(a) *Ist L ein affiner Unterraum von V, so ist $\varphi(L)$ ein affiner Unterraum von W und es ist* $\dim \varphi(L) \leq \dim L$.

(b) *Ist M ein affiner Unterraum von W, so ist $\varphi^{-1}(M)$ entweder die leere Menge oder ein affiner Unterraum von V.*

(c) *φ erhält die Parallelität, d.h. für alle affinen Unterräume L_1, L_2 von V folgt aus $L_1 \parallel L_2$ stets $\varphi(L_1) \parallel \varphi(L_2)$.*

Beweis. Es sei $\varphi(x) = \Phi(x) + w$ mit $\Phi \in \mathrm{Hom}\,(V, W)$ und $w \in W$.

(a) Mit $L = x_0 + U$ erhalten wir $\varphi(L) = \Phi(U) + \Phi(x_0) + w$, also ist $\varphi(L)$ ein affiner Unterraum von W, und es gilt $\dim \varphi(L) = \dim \Phi(U) \leq \dim U = \dim L$.

(b) Sei $M = y' + U'$ mit $U' \subset W$ und $y' \in W$. Ist $\varphi^{-1}(M) \neq \emptyset$, so gibt es ein $x_0 \in V$ mit $\varphi(x_0) \in M$. Damit folgt $\Phi(x_0) \in y' - w + U'$ und nach der Bemerkung von S.126 $\Phi(x_0) + U' = y' - w + U'$. Es gilt dann:

$$x \in \varphi^{-1}(M) \iff \varphi(x) \in M \iff \Phi(x) \in y' - w + U' = \Phi(x_0) + U'$$

$$\iff \Phi(x - x_0) \in U' \iff x \in x_0 + \Phi^{-1}(U'),$$

d.h. $\varphi^{-1}(M)$ ist ein affiner Unterraum von V.

(c) Es seien $L_1 = x_1 + U_1$, $L_2 = x_2 + U_2$ und o.B.d.A. sei $U_1 \subset U_2$. Dann folgt $\Phi(U_1) \subset \Phi(U_2)$ und somit $\varphi(L_1) \parallel \varphi(L_2)$. ■

Nun wenden wir uns den affinen Selbstabbildungen von V zu. Die einfachsten Beispiele neben der Identität sind die Translationen.

Weitere Beispiele sind die *Streckungen* mit *Zentrum* z und *Streckungsfaktor* c. Sie bilden z auf sich ab und ordnen jedem von z verschiedenen Punkt x denjenigen

Punkt y auf der Geraden durch x und z zu, für den $y - z = c\,(x - z)$ gilt. Somit hat eine solche Streckung $\sigma : V \longrightarrow V$ die Darstellung $\sigma(x) = y = c\,x + (1 - c)\,z$, ist also tatsächlich eine affine Abbildung.

Für $c = 0$ bildet σ alle Punkte auf z ab, für $c = 1$ ist σ die Identität. Für char $\mathbb{K} \neq 2$ und $c = -1$ schließlich gilt $z = \frac{1}{2}(x + \sigma(x))$, d.h. z ist der *Mittelpunkt* der Punkte x und $\sigma(x)$. In diesem Fall heißt σ *Punktspiegelung* an z.

Translationen und Streckungen mit Streckungsfaktor $c \neq 0$ sind Affinitäten.

Fixelemente affiner Selbstabbildungen

Für die Klassifikation der affinen Selbstabbildungen von V nach geometrischen Gesichtspunkten sind die Fixelemente (Fixpunkte, Fixräume, Fixrichtungen) einer solchen Abbildung wichtig.

Definition. Es seien V ein $\mathbb{K}$–Vektorraum und $\varphi : V \longrightarrow V$ eine affine Abbildung. Jeder Punkt x mit $\varphi(x) = x$ heißt *Fixpunkt* von φ. Allgemein heißt ein affiner Unterraum $L \subset V$ der Dimension k ein *k–dimensionaler Fixraum* von φ, falls $\varphi(L) = L$ gilt.

Bemerkungen. **(a)** Fixräume bleiben unter φ im allgemeinen nicht punktweise fest.

(b) $L = x_0 + U$ ist genau dann Fixraum der affinen Abbildung $\varphi : x \longmapsto \Phi(x) + v$, $x \in V$, wenn $\Phi(U) = U$ und $\varphi(x_0) \in L$ gilt.

(c) x ist genau dann Fixpunkt von φ, wenn x die Gleichung $(\Phi - \mathrm{id}_V)(x) = -v$ erfüllt. Die Menge aller Fixpunkte von φ ist also entweder leer oder ein affiner Unterraum L von V mit Richtungsraum $\mathrm{Kern}(\Phi - \mathrm{id}_V)$; L heißt dann *Fixpunktraum* von φ.

(d) Ist V endlich dimensional, so besitzt φ genau dann genau <u>einen</u> Fixpunkt, wenn $\Phi - \mathrm{id}_V$ injektiv ist. In diesem Fall enthält jeder Fixraum L von φ diesen Fixpunkt:

Sei $L = x_0 + U$. Aus $\Phi(U) = U$ folgt zunächst $(\Phi - \mathrm{id}_V)(U) \subset U$. Weil $\Phi - \mathrm{id}_V$ bijektiv ist, gilt sogar Gleichheit, also auch $(\Phi - \mathrm{id}_V)^{-1}(U) = U$. Aus $\varphi(x_0) \in L$ folgt $(\Phi - \mathrm{id}_V)(x_0) = -v + u$, $u \in U$, und aus $\varphi(x) = x$ folgt $(\Phi - \mathrm{id}_V)(x) = -v$. Durch Subtraktion erhalten wir $(\Phi - \mathrm{id}_V)(x - x_0) = -u \in U$, also $x \in x_0 + (\Phi - \mathrm{id}_V)^{-1}(U) = L$.

Beispiele. **(a)** Translationen, die von der Identität verschieden sind, besitzen keine Fixpunkte. Die k–dimensionalen Fixräume, $k \geq 1$, sind genau die k–dimensionalen affinen Unterräume, die zur Translationsrichtung parallel sind.

(b) Die Streckungen mit Streckungsfaktor $c \neq 0$ und $c \neq 1$ besitzen das Zentrum als einzigen Fixpunkt, und genau die affinen Unterräume, die das Zentrum enthalten, sind Fixräume.

(c) Sei $\varphi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, $\varphi(x) = A\,x + v$ mit $A = \begin{bmatrix} 1 & 3 \\ 1 & -1 \end{bmatrix}$, $v = \begin{bmatrix} 6 \\ -4 \end{bmatrix}$.

Wir bestimmen die Fixpunkte: $\varphi(x) = x \iff (A - E_2)\,x = -\,v$. Die einzige Lösung dieses LGS ist $\begin{bmatrix} 0 \\ -2 \end{bmatrix}$ ist. Somit ist $\begin{bmatrix} 0 \\ -2 \end{bmatrix}$ einziger Fixpunkt von φ.

Um die Fixgeraden von φ zu erhalten, müssen wir nach der obigen Bemerkung (b) diejenigen eindimensionalen Unterräume $[x]$ berechnen, für die $A\,([x]) = [x]$ gilt. Dies führt uns auf das Problem, alle $c \in \mathbb{R}$, $c \neq 0$, und alle $x \in \mathbb{R}^2$, $x \neq o$, zu bestimmen, für die $A\,x = c\,x$ gilt. Jedes solche $x \neq o$ heißt ein *Eigenvektor* zum *Eigenwert* c. Nun gilt

$$A\,x = c\,x\,,\ x \neq o\,,\ \text{ist lösbar} \iff (A - c\,E_2)\,x = o \ \text{ist nichttrivial lösbar}$$

$$\iff \operatorname{Rg}(A - c\,E_2) < 2 \iff c = \pm\,2.$$

Für $c = 2$ erhalten wir als Fixrichtung $[\begin{bmatrix} 3 \\ 1 \end{bmatrix}]$, für $c = -2$ erhalten wir als Fixrichtung $[\begin{bmatrix} -1 \\ 1 \end{bmatrix}]$. Da in diesem Fall jede Fixgerade den einzigen Fixpunkt enthalten muß, sind

$$\begin{bmatrix} 0 \\ -2 \end{bmatrix} + [\begin{bmatrix} 3 \\ 1 \end{bmatrix}] \quad \text{und} \quad \begin{bmatrix} 0 \\ -2 \end{bmatrix} + [\begin{bmatrix} -1 \\ 1 \end{bmatrix}]$$

die beiden einzigen Fixgeraden von φ.

(d) Sei $\varphi : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$, $\varphi(x) = A\,x + v$ mit $A = \begin{bmatrix} 7 & -6 & 11 \\ 0 & 1 & -1 \\ -4 & 4 & -7 \end{bmatrix}$, $v = \begin{bmatrix} 4 \\ -2 \\ -4 \end{bmatrix}$.

Zur Bestimmung der Fixpunkte lösen wir das LGS $(A - E_3)\,x = -\,v$ mit der erweiterten Matrix

$$\left[\begin{array}{rrr|r} 6 & -6 & 11 & -4 \\ 0 & 0 & -1 & 2 \\ -4 & 4 & -8 & 4 \end{array}\right]$$

Es hat als Lösungsmenge die Gerade

$$h = \begin{bmatrix} 3 \\ 0 \\ -2 \end{bmatrix} + [\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}],$$

d.h. φ besitzt eine Fixpunktgerade.

Die Bestimmung der Fixgeraden $g = x_0 + [x]$ erfolgt in zwei Schritten.
1.Schritt: Wir berechnen wieder alle eindimensionalen Fixrichtungen $[x]$, d.h. alle Eigenvektoren zu dem Eigenwert $c \neq 0$:

$$(A - c\,E_3)\,x = 0 \text{ ist nichttrivial lösbar} \iff \mathrm{Rg}\,(A - c\,E_3) < 3 \iff c = \pm 1.$$

Für $c = -1$ bzw. $c = 1$ erhalten wir als zugehörige Fixrichtungen

$$[x_1] = [\begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}] \text{ bzw. } [x_2] = [\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}].$$

2. Schritt: Zu den Fixrichtungen $[x_1]$ bzw. $[x_2]$ suchen wir alle Punkte x_0 mit $\varphi(x_0) \in x_0 + [x_1]$ bzw. alle Punkte y_0 mit $\varphi(y_0) \in y_0 + [x_2]$:

$$\varphi(x_0) \in x_0 + [x_1] \iff A\,x_0 + v = x_0 + a\,x_1\,,\ a \in \mathbb{R} \iff (A - E_3)\,x_0 - a\,x_1 = -\,v\,.$$

Dies ist ein inhomogenes lineares Gleichungssystem mit den Unbekannten x_0 und a und der zugehörigen erweiterten Matrix

$$\left[\begin{array}{rrrr|r} 6 & -6 & 11 & 2 & -4 \\ 0 & 0 & -1 & -1 & 2 \\ -4 & 4 & -8 & -2 & 4 \end{array}\right]$$

Als Lösungen erhalten wir

$$x_0 = \begin{bmatrix} 3+ \quad b + 3/2\ a \\ b \\ -2-a \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ -2 \end{bmatrix} + b \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + a \begin{bmatrix} 3/2 \\ 0 \\ -1 \end{bmatrix} = \tilde{a} \begin{bmatrix} 3 \\ 0 \\ -2 \end{bmatrix} + b \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \tilde{a}, b \in \mathbb{R}\,.$$

Dies ist eine Ebene L_1 durch den Nullpunkt o, die die Fixpunktgerade h enthält. Weil x_1 Linearkombination der Vektoren

$$\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ -2 \end{bmatrix}$$

ist, liegen alle Fixgeraden mit Richtung $[x_1]$ in L_1. Sie schneiden dann alle die Fixpunktgerade, weil x_1 und $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ linear unabhängig sind. Somit erhalten wir als Fixgeraden mit Richtung $[x_1]$ die in der Ebene L_1 liegende Geradenschar

$$\begin{bmatrix} 3 \\ 0 \\ -2 \end{bmatrix} + b \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + [x_1]\,, \quad b \in \mathbb{R}.$$

Auf die gleiche Art und Weise ergeben sich als Fixgeraden mit Richtung $[x_2]$ die zu der Fixpunktgeraden h parallelen Geraden

$$\begin{bmatrix} 3 \\ 0 \\ -2 \end{bmatrix} + b \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix} + [x_2]\,, \quad b \in \mathbb{R},$$

die alle in einer von L_1 verschiedenen Ebene

$$L_2 = \begin{bmatrix} 3 \\ 0 \\ -2 \end{bmatrix} + \left[\begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right]$$

durch h liegen.

Kapitel 4 Determinanten und Eigenwerte

§ 1 Determinanten

Wir wollen bei der Einführung der Determinante einer quadratischen Matrix $A \in \mathbb{K}^{n \times n}$ nicht den historischen Weg wählen, sondern vielmehr eine Definition geben, die es erlaubt, theoretische Überlegungen einfach durchzuführen. Zu diesem Zweck müssen wir uns nochmals kurz mit Permutationen beschäftigen.

Es sei

$$\pi = \begin{bmatrix} 1 & \cdots & i & \cdots & j & \cdots & m \\ \pi(1) & \cdots & \pi(i) & \cdots & \pi(j) & \cdots & \pi(m) \end{bmatrix} \in S_m .$$

Dann heißt jedes Paar $(i,j) \in \{1,\dots,m\} \times \{1,\dots,m\}$ mit $i < j$ und $\pi(i) > \pi(j)$ ein *Fehlstand* von π . Die Anzahl $F(\pi)$ der Fehlstände von π heißt die *Fehlstandszahl* von π.

Beispiel. Die Permutation

$$\pi = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{bmatrix}$$

besitzt die Fehlstände (1,2), (1,3), (1,4) und (2,3), also ist $F(\pi) = 4$.

Eigenschaften der Fehlstandszahl

(a) Für alle $\pi \in S_m$ gilt

$$\prod_{1 \le i < j \le m} \frac{\pi(j) - \pi(i)}{j - i} = (-1)^{F(\pi)}.$$

Beweis. Die Gleichung folgt, weil links im Zähler und im Nenner bis auf Vorzeichen dieselben Faktoren stehen und $\pi(j) - \pi(i)$ genau dann negativ ist, wenn (i,j) ein Fehlstand ist. ■

(b) Für alle σ, $\pi \in S_m$ gilt: $(-1)^{F(\sigma\circ\pi)} = (-1)^{F(\sigma)} \cdot (-1)^{F(\pi)}.$

Beweis.

$$(-1)^{F(\sigma\circ\pi)} = \prod_{1 \le i < j \le m} \frac{\sigma(\pi(j)) - \sigma(\pi(i))}{j - i}$$

$$= \prod_{1 \le i < j \le m} \frac{\sigma(\pi(j)) - \sigma(\pi(i))}{\pi(j) - \pi(i)} \cdot \frac{\pi(j) - \pi(i)}{j - i}$$

$$= \prod_{1 \le \pi(i) < \pi(j) \le m} \frac{\sigma(\pi(j)) - \sigma(\pi(i))}{\pi(j) - \pi(i)} \prod_{1 \le i < j \le m} \frac{\pi(j) - \pi(i)}{j - i}$$

$$= (-1)^{F(\sigma)} \cdot (-1)^{F(\pi)}.$$ ■

(c) $(-1)^{F(\pi)} = (-1)^{F(\pi^{-1})}$.

(d) $F(\tau^{(i,j)}) = 2\,(j-i) - 1$.

Beweis. Die Fehlstände der Transposition $\tau^{(i,j)}$ sind gerade die Paare $(i,i+1),\ldots,$ $(i,j-1)$, (i,j) und $(i+1,j),\ldots,(j-1,j)$. Dies ergibt zusammen $(j-i) + (j-1-i) =$ $2\,(j-i) - 1$ Fehlstände. ■

Bemerkung. Aus diesen Eigenschaften folgt direkt der Satz : *Eine Permutation ist genau dann gerade, wenn ihre Fehlstandszahl gerade ist.*

Beweis. Ist π eine gerade Permutation, so ist π Produkt einer geraden Anzahl von Transpositionen $\tau_1,\ldots,\tau_{2k}$. Mit (b) folgt dann $(-1)^{F(\pi)} = (-1)^{F(\tau_1)} \cdots (-1)^{F(\tau_{2k})} =$ $(-1)^{2k} = 1$, also ist $F(\pi)$ gerade. Ebenso zeigt man, daß für ungerades π auch $F(\pi)$ ungerade ist. ■

Nun können wir die angekündigte Definition geben.

Definition. Es sei $A = (\!(a_{ij})\!) \in \mathbb{K}^{n\times n}$. Dann heißt

$$\det A := \sum_{\pi \in S_n} (-1)^{F(\pi)}\, a_{\pi(1),1} \cdots a_{\pi(n),n}$$

Determinante der Matrix A. Statt $\det A$ schreiben wir auch $|A|$ bzw.

$$\det A = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

Bemerkungen. **(a)** Für $n = 1$ gilt $\det A = a_{11}$.

Für $n = 2$ erhalten wir $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}\, a_{22} - a_{21}\, a_{12}$.

Für $n = 3$ müssen wir über die Elemente der Gruppe S_3 summieren. Dies ergibt die folgende Regel (*Regel von Sarrus*):

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23}$$

$$- a_{31}a_{22}a_{13} - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33} \,.$$

(b) Ist A eine obere bzw. untere Dreiecksmatrix, so folgt aus der Definition der Determinante sofort $\det A = a_{11} \cdots a_{nn}$.

(c) $\det E_n = 1$.

Satz 1. *Es sei* $A \in \mathbb{K}^{n\times n}$. *Dann gilt* $\det A = \det A^{\top}$.

Beweis. Es seien $A = (\!(a_{ij})\!)$ und $A^{\top} = (\!(\tilde{a}_{ij})\!)$. Dann ist

$$\det A^{\top} = \sum_{\pi \in S_n} (-1)^{F(\pi)}\, \tilde{a}_{\pi(1),1} \cdots \tilde{a}_{\pi(n),n}$$

$$= \sum_{\pi \in S_n} (-1)^{F(\pi)}\, a_{1,\pi(1)} \cdots a_{n,\pi(n)}$$

$$= \sum_{\pi \in S_n} (-1)^{F(\pi)} a_{\pi^{-1}\circ\pi(1),\pi(1)} \cdots a_{\pi^{-1}\circ\pi(n),\pi(n)}$$

$$= \sum_{\pi^{-1} \in S_n} (-1)^{F(\pi^{-1})} a_{\pi^{-1}(1),1} \cdots a_{\pi^{-1}(n),n} = \det A. \qquad \blacksquare$$

Bemerkung. Aus dem voranstehenden Beweis folgt unmittelbar

$$\det A = \sum_{\pi \in S_n} (-1)^{F(\pi)} a_{1,\pi(1)} \cdots a_{n,\pi(n)} .$$

Als nächstes wollen wir einige Eigenschaften der Abbildung $A \longmapsto \det A$ beweisen. Dabei wird insbesondere die Abhängigkeit von $\det A$ von den Spalten (Zeilen) von A eine Rolle spielen. Um diese Abhängigkeit besser überblicken zu können, definieren wir folgende Funktionen auf $(\mathbb{K}^n)^n$:

$$\Delta : (\mathbb{K}^n)^n \longrightarrow \mathbb{K} \ , \ \Delta(x_1,\dots,x_n) := \det\,(x_1 \mid \cdots \mid x_n) \, .$$

$$\tilde{\Delta} : (\mathbb{K}^n)^n \longrightarrow \mathbb{K} \ , \ \tilde{\Delta}(y_1,\dots,y_n) := \det \begin{bmatrix} y_1^{\mathsf{T}} \\ \cdots \\ \cdots \\ y_n^{\mathsf{T}} \end{bmatrix}$$

Wegen Satz 1 gilt aber $\Delta = \tilde{\Delta}$, es genügt daher nur Δ zu untersuchen.

Satz 2. *Die Abbildung* $\Delta : (\mathbb{K}^n)^n \longrightarrow \mathbb{K}$ *hat folgende Eigenschaften:*

(a) Δ *ist* n*–fach multilinear, d.h. es gilt für alle* $j \in \{1, \dots ,n\}$, *alle* $a,b \in \mathbb{K}$ *und alle* $x_j\, , \, \tilde{x}_j \in \mathbb{K}^n$:

$$\Delta(\dots,x_{j-1},a\,x_j + b\,\tilde{x}_j,x_{j+1},\dots) = a\,\Delta(\dots,x_{j-1},x_j,x_{j+1},\dots) + b\,\Delta(\dots,x_{j-1},\tilde{x}_j,x_{j+1},\dots) \, .$$

(b) *Für alle* $\pi \in S_n$ *gilt:* $\Delta(x_{\pi(1)},\dots,x_{\pi(n)}) = (-1)^{F(\pi)}\,\Delta(x_1,\dots,x_n)$.

Insbesondere ist Δ *alternierend, d.h. für alle* $i,j \in \{1,\dots,n\}$, $i \neq j$, *gilt:*

$$\Delta(\dots,\underbrace{x_i}_{\substack{i-te\\ Stelle}},\dots,\underbrace{x_j}_{\substack{j-te\\ Stelle}},\dots) = -\,\Delta(\dots,\underbrace{x_j}_{\substack{i-te\\ Stelle}},\dots,\underbrace{x_i}_{\substack{j-te\\ Stelle}},\dots)$$

(c) *Δ ist normiert, d.h.* $\Delta(e_1,\dots,e_n) = 1$.

(d) *Die Vektoren $x_1,\dots,x_n$ sind genau dann linear abhängig, wenn* $\Delta(x_1,\dots,x_n) = 0$.

Beweis. (a) folgt direkt aus der Definition der Determinante.

(b) Nach Definition von Δ und der Determinante gilt für alle $\pi \in S_n$

$$\Delta(x_{\pi(1)},\dots,x_{\pi(n)}) = \sum_{\sigma \in S_n} (-1)^{F(\sigma)}\, x_{\sigma(1),\pi(1)} \cdots x_{\sigma(n),\pi(n)}$$

$$\underset{\sigma = \tau\circ\pi}{=} \sum_{\tau \in S_n} (-1)^{F(\tau\circ\pi)}\, x_{\tau(1),1} \cdots x_{\tau(n),n} = (-1)^{F(\pi)}\, \Delta(x_1,\dots,x_n)\,.$$

Der zweite Teil der Behauptung ergibt sich nun als Spezialfall, wenn man $\pi = \tau^{(i,j)}$ wählt und beachtet, daß die Fehlstandszahl einer Transposition ungerade ist.

(c) $\Delta(e_1,\dots,e_n) = \det E_n = 1$.

(d) Die Vektoren $x_1,\dots,x_n$ seien linear abhängig. Wir beweisen zunächst den folgenden *Spezialfall:*

Ist $x_i = x_j$ *für* $i \neq j$, *so ist* $\Delta(x_1,\dots,x_n) = 0$.

Beweis. In $\Delta(x_1,\dots,x_n)$ tritt mit jedem Summanden

$$(-1)^{F(\pi)}\, x_{\pi(1),1} \cdots x_{\pi(i),i} \cdots x_{\pi(j),j} \cdots x_{\pi(n),n}$$

auch der Summand

$$(-1)^{F(\pi\circ\tau^{(i,j)})}\, x_{\pi(1),1} \cdots x_{\pi(j),i} \cdots x_{\pi(i),j} \cdots x_{\pi(n),n}$$

auf, der wegen $x_i = x_j$ die Form

$$(-1)^{F(\pi)}\,(-1)^{F(\tau^{(i,j)})}\, x_{\pi(1),1} \cdots x_{\pi(i),i} \cdots x_{\pi(j),j} \cdots x_{\pi(n),n}$$

hat, also das (-1) – fache des ersteren ist. Somit gilt $\Delta(x_1,\dots,x_n) = 0$. (Für Körper $\mathbb{K}$ der Charakteristik $\neq 2$ folgt der Spezialfall auch direkt aus (b).)

Nun können wir den allgemeinen Fall beweisen. Sind die Vektoren $x_1,\dots,x_n$

linear abhängig, so gibt es ein $i \in \{1,\dots,n\}$ mit $x_i = \sum_{\substack{j=1\\ j\neq i}}^{n} a_j x_j$, und wir erhalten mit Hilfe von (a) und dem Spezialfall, daß

$$\Delta(x_1,\dots,x_n) = \sum_{\substack{j=1\\ j\neq i}}^{n} a_j \, \Delta(x_1,\dots,x_{i-1},x_j,x_{i+1},\dots,x_n) = 0.$$

Es sei nun umgekehrt $\Delta(x_1,\dots,x_n) = 0$. Sind die Vektoren $x_1,\dots,x_n$ linear unabhängig, so bilden sie eine Basis von $\mathbb{K}^n$ und die Standardbasisvektoren $e_1,\dots,e_n$ lassen sich in der Form $e_j = \sum_{k=1}^{n} a_{kj} x_k$ darstellen, $j = 1,\dots,n$. Mit (a), dem Spezialfall, (b) und der Voraussetzung folgt dann

$$\Delta(e_1,\dots,e_n) = \sum_{k_1,\dots,k_n=1}^{n} a_{k_1 1} \cdots a_{k_n n} \, \Delta(x_{k_1},\dots,x_{k_n})$$

$$= \sum_{\pi \in S_n} a_{\pi(1),1} \cdots a_{\pi(n),n} \, \Delta(x_{\pi(1)},\dots,x_{\pi(n)}) = 0,$$

ein Widerspruch zu (c). Also sind die Vektoren $x_1,\dots,x_n$ linear abhängig. ■

Für das praktische Rechnen mit Determinanten ergeben sich aus Satz 2 zusammen mit Satz 1 folgende nützliche Regeln:

Satz 3. *Es sei $A \in \mathbb{K}^{n\times n}$. Dann gilt:*

(a) *Addition des Vielfachen einer Spalte (Zeile) zu einer anderen ändert* det A *nicht.*

(b) *Multiplikation einer Spalte (Zeile) mit $a \in \mathbb{K}$ vervielfacht* det A *um den Faktor a.*

(c) *Vertauschen von Spalten (Zeilen) ändert das Vorzeichen von* det A.

(d) *A ist genau dann regulär, wenn* det A *von Null verschieden ist.*

Wegen Satz 3 läßt sich das Prinzip des Gaußschen Algorithmus auch zur Berechnung von det A einsetzen.

Beispiel.

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

$$= \begin{vmatrix} 2 & 0 & 1 & 2 & 0 \\ 0 & -1 & -1 & -1 & 1 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & -1 & 4 \end{vmatrix} = \begin{vmatrix} 2 & 0 & 1 & 2 & 0 \\ 0 & -1 & -1 & -1 & 1 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & 0 & -3 \end{vmatrix} = 6.$$

Die Berechnung von det A läßt sich dadurch vereinfachen, daß man zuerst geeignet umformt und dann nach einer Spalte (Zeile) mit möglichst vielen Nullen "entwickelt". Diese Möglichkeit ergibt sich aus dem nachstehenden Satz, zu dessen Formulierung wir folgende Abkürzung verwenden:

Es seien $A \in \mathbb{K}^{n\times n}$ und $i,j \in \{1,...,n\}$. Dann bezeichnen wir mit $A_{i,j}$ diejenige Matrix, die aus A durch Streichen der i–ten Zeile und der j–ten Spalte hervorgeht.

Satz 4. *Es sei $A = (\!(a_{ij})\!) \in \mathbb{K}^{n\times n}$. Dann gilt für $j = 1,...,n$:*

(a) $\det A = \sum_{k=1}^{n} (-1)^{k+j} a_{kj} \det A_{k,j}$. *("Entwicklung nach der j–ten Spalte")*

(b) $\det A = \sum_{k=1}^{n} (-1)^{k+j} a_{jk} \det A_{j,k}$. *("Entwicklung nach der j–ten Zeile")*

Beweis. Zunächst beweisen wir wieder einen Spezialfall:

$$(*) \qquad \begin{vmatrix} b_{11} & \cdots & b_{1,n-1} & 0 \\ \vdots & & \vdots & \vdots \\ b_{n-1,1} & \cdots & b_{n-1,n-1} & 0 \\ b_{n,1} & \cdots & b_{n,n-1} & 1 \end{vmatrix} = \begin{vmatrix} b_{11} & \cdots & b_{1,n-1} \\ \vdots & & \vdots \\ b_{n-1,1} & \cdots & b_{n-1,n-1} \end{vmatrix}$$

Beweis. Für die linke Determinante gilt

$$\sum_{\pi \in S_n} (-1)^{F(\pi)} b_{\pi(1),1} \cdots b_{\pi(n),n} = \sum_{\substack{\pi \in S_n \\ \pi(n) = n}} (-1)^{F(\pi)} b_{\pi(1),1} \cdots b_{\pi(n-1),n-1}$$

$$= \sum_{\sigma \in S_{n-1}} (-1)^{F(\sigma)} b_{\sigma(1),1} \cdots b_{\sigma(n-1),n-1} .$$

Die letzte Summe ist aber nach Definition gerade die rechte Determinante.

(a) Wegen Satz 3 gilt für $j = 1,\dots,n$

$$\det A = a_{1j} \begin{vmatrix} a_{11} & \cdots & 1 & \cdots & a_{1n} \\ a_{21} & \cdots & 0 & \cdots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix} + \cdots + a_{nj} \begin{vmatrix} a_{11} & \cdots\cdots & 0 & \cdots\cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{n-1,1} & \cdots & 0 & \cdots & a_{n-1,n} \\ a_{n1} & \cdots\cdots & 1 & \cdots\cdots & a_{nn} \end{vmatrix} .$$

Alle Determinanten in dieser Summe können mit Hilfe von Satz 3 so umgeformt werden, daß wir (*) anwenden können. Wir erhalten durch Vertauschen von Zeilen und Spalten

$$\begin{array}{c} \\ \\ k\,\{ \\ \\ \\ \end{array} \begin{vmatrix} a_{11} & \cdots & 0 & \cdots & a_{1n} \\ \vdots & & 0 & & \vdots \\ a_{k1} & \cdots & 1 & \cdots & a_{kn} \\ \vdots & & 0 & & \vdots \\ a_{n1} & \cdots & \underbrace{0}_{j} & \cdots & a_{nn} \end{vmatrix} = (-1)^{n-j}(-1)^{n-k} \det A_{k,j} = (-1)^{k+j} \det A_{k,j}$$

und somit für $j = 1,\dots,n$: $\det A = \sum_{k=1}^{n} (-1)^{k+j} a_{kj} \det A_{k,j}$.

(b) Die Behauptung kann analog bewiesen werden, folgt aber auch sofort aus Satz 1 und Teil (a). ∎

Beispiele. **(a)** Wir rechnen das obige Beispiel nochmals:

$$\begin{vmatrix} 2 & 0 & 1 & 2 & 0 \\ 2 & -1 & 0 & 1 & 1 \\ 0 & 1 & 2 & 1 & 2 \\ -2 & 0 & 2 & -1 & 2 \\ 2 & 0 & 0 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 2 & 0 & 1 & 2 & 0 \\ 0 & -1 & -1 & -1 & 1 \\ 0 & 1 & 2 & 1 & 2 \\ 0 & 0 & 3 & 1 & 2 \\ 0 & 0 & -1 & -1 & 1 \end{vmatrix} \underset{\substack{\text{Entw.} \\ \text{1. Sp.}}}{=} 2\cdot \begin{array}{c} \begin{vmatrix} -1 & -1 & -1 & 1 \\ 1 & 2 & 1 & 2 \\ 0 & 3 & 1 & 2 \\ 0 & -1 & -1 & 1 \end{vmatrix} \\ -2 \end{array} -1$$

$$= 2\cdot \begin{vmatrix} -1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & -1 & -1 & 3 \end{vmatrix} \underset{\substack{\text{Entw.} \\ \text{1. Ze.}}}{=} (-2)\cdot \begin{vmatrix} 2 & 1 & 0 \\ 3 & 1 & 0 \\ -1 & -1 & 3 \end{vmatrix} \underset{\substack{\text{Entw.} \\ \text{3. Sp.}}}{=} (-2)\cdot 3 \cdot \begin{vmatrix} 2 & 1 \\ 3 & 1 \end{vmatrix} = 6 \,.$$

(b) *Vandermondesche Determinante.* Für alle $x_1,\dots,x_k \in \mathbb{K}$ gilt:

$$\begin{vmatrix} 1 & 1 & 1 & \cdots\cdots & 1 \\ x_1 & x_2 & x_3 & \cdots\cdots & x_n \\ x_1^2 & x_2^2 & x_3^2 & \cdots\cdots & x_n^2 \\ \vdots & \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \cdots\cdots & x_n^{n-1} \end{vmatrix} \begin{array}{l} \\ -x_1 \\ -x_1 \\ \\ -x_1 \end{array} = \begin{vmatrix} 1 & 1 & \cdots\cdots & 1 \\ 0 & (x_2-x_1) & \cdots\cdots & (x_n-x_1) \\ 0 & x_2(x_2-x_1) & \cdots\cdots & x_n(x_n-x_1) \\ \vdots & \vdots & & \vdots \\ 0 & x_2^{n-2}(x_2-x_1) & \cdots\cdots & x_n^{n-2}(x_n-x_1) \end{vmatrix}$$

$$= (x_2-x_1)\cdots(x_n-x_1) \begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_2 & x_3 & \cdots & x_n \\ \vdots & & & \vdots \\ x_2^{n-2} & x_3^{n-2} & \cdots & x_n^{n-2} \end{vmatrix}$$

$$= \cdots = (x_2-x_1)\cdots(x_n-x_1)(x_3-x_2)\cdots(x_n-x_2)\cdots(x_n-x_{n-1}) = \prod_{1\le i<j\le n} (x_j - x_i)\,.$$

Für Produkte von Matrizen gilt der folgende Determinantenmultiplikationssatz.

Satz 5. *Es seien $A,B \in \mathbb{K}^{n\times n}$. Dann gilt:* $\det(A\,B) = \det A \cdot \det B$.

Beweis. Seien $a_1,\dots,a_n$ die Spalten von A und $b_1,\dots,b_n$ die von $B = (\!(b_{ij})\!)$. Dann ist

$$A\,B = (Ab_1 \mid \cdots \mid Ab_n) = (b_{11}a_1 + \cdots + b_{n1}a_n \mid \cdots \mid b_{1n}a_1 + \cdots + b_{nn}a_n),$$

woraus mit Hilfe von Satz 2 (a),(b),(d) folgt:

$$\det(A\,B) = \sum_{i_1=1}^{n} \cdots \sum_{i_n=1}^{n} b_{i_1 1} \cdots b_{i_n n}\, \Delta(a_{i_1},\dots,a_{i_n})$$

$$= \sum_{\pi \in S_n} b_{\pi(1),1} \cdots b_{\pi(n),n}\, \Delta(a_{\pi(1)},\dots,a_{\pi(n)})$$

$$= \sum_{\pi \in S_n} (-1)^{F(\pi)}\, b_{\pi(1),1} \cdots b_{\pi(n),n}\, \Delta(a_1,\dots,a_n)$$

$$= \det A \cdot \det B\,. \qquad \blacksquare$$

Korollar 6 (***Kästchenmultiplikationssatz***).

(a) *Ist A von der Form*

$$A = \begin{bmatrix} B & O \\ C & D \end{bmatrix}$$

mit $B \in \mathbb{K}^{m\times m}$, $C \in \mathbb{K}^{(n-m)\times m}$, $O \in \mathbb{K}^{m\times(n-m)}$ *und* $D \in \mathbb{K}^{(n-m)\times(n-m)}$, *so gilt*

$$\det A = \det B \cdot \det D.$$

(b) *Ist A von der Form*

$$A = \begin{bmatrix} B & C \\ O & D \end{bmatrix}$$

mit $B \in \mathbb{K}^{m\times m}$, $O \in \mathbb{K}^{(n-m)\times m}$, $C \in \mathbb{K}^{m\times(n-m)}$ *und* $D \in \mathbb{K}^{(n-m)\times(n-m)}$, *so gilt*

$$\det A = \det B \cdot \det D.$$

Beweis. (a) Es ist

$$A = \begin{bmatrix} B & O \\ C & E_{n-m} \end{bmatrix} \cdot \begin{bmatrix} E_m & O \\ O & D \end{bmatrix},$$

woraus mit den Sätzen 5 und 4 sofort die Behauptung folgt.

(b) Satz 1 und Teil (a). ■

Bemerkungen und Definitionen. **(a)** Ist die Matrix $A \in \mathbb{K}^{n\times n}$ regulär, so gilt $\det(A^{-1}) = (\det A)^{-1}$.

(b) Ähnliche Matrizen besitzen dieselbe Determinante.

(c) Sind V ein n–dimensionaler $\mathbb{K}$–Vektorraum und $\Phi \in \mathrm{Hom}(V, V)$, so haben wegen der Aussage (b) alle Abbildungsmatrizen von Φ dieselbe Determinante. Wir setzen daher $\det \Phi := \det A_\Phi$ und nennen dies die *Determinante* der linearen Abbildung Φ.

Wir wollen nun Determinanten zur Lösung eines linearen Gleichungssystems bzw. zur Bestimmung der Inversen einer regulären Matrix heranziehen.

Satz 7 (***Cramersche Regel***). *Es seien $A \in \mathbb{K}^{n\times n}$ eine reguläre Matrix mit den Spalten $a_1,\dots,a_n$ und $b \in \mathbb{K}^n$. Dann ist die eindeutige Lösung des linearen Gleichungssystems $A\,x = b$ gegeben durch*

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad mit \quad x_i = \frac{\det\ (a_1 \,|\cdots|\, a_{i-1} \,|\, b \,|\, a_{i+1} \,|\cdots|\, a_n)}{\det\ A}.$$

Beweis. Nach Satz 3 gilt $x_i \cdot \det A = \det(a_1 \,|\cdots|\, a_{i-1} \,|\, x_i\, a_i \,|\, a_{i+1} \,|\cdots|\, a_n)$

$$= \det(a_1 \,|\cdots|\, a_{i-1} \,|\, \sum_{j=1}^{n} x_j\, a_j \,|\, a_{i+1} \,|\cdots|\, a_n) = \det(a_1 \,|\cdots|\, a_{i-1} \,|\, b \,|\, a_{i+1} \,|\cdots|\, a_n). \ ■$$

Satz 8. *Es sei $A \in \mathbb{K}^{n\times n}$ eine reguläre Matrix mit der Inversen $A^{-1} = (\!(b_{ij})\!)$. Dann gilt für alle $i,j = 1,\dots,n$:*

$$b_{ij} = (-1)^{i+j}\,(\det A)^{-1} \cdot \det A_{j,i}\,.$$

Beweis. Für $A^{-1} = (b_1 \,|\cdots|\, b_n)$ gilt $A \cdot (b_1 \,|\cdots|\, b_n) = E_n = (e_1 \,|\cdots|\, e_n)$, also $A\, b_j = e_j$ für $j = 1,\dots,n$. Mit den Sätzen 7 und 4 erhalten wir daher

$$b_{ij} = (\det A)^{-1} \cdot \det (a_1 | \cdots | a_{i-1} | e_j | a_{i+1} | \cdots | a_n) = (-1)^{i+j} (\det A)^{-1} \cdot \det A_{j,i} \,. \blacksquare$$

Neben dem hier gewählten direkten Zugang zur Determinante findet man in vielen Büchern auch den Zugang über eine Funktion (*Determinantenfunktion, Volumenfunktion*) auf $\mathbb{K}^{n\times n}$ bzw. $(\mathbb{K}^n)^n$ bzw. V^n mit gewissen Eigenschaften. Es wird dann gezeigt, daß es bis auf Normierung genau eine derartige Funktion Δ gibt und daß sie auf $(\mathbb{K}^n)^n$ die Form hat, die wir auf S. 166 als Definition benutzt haben. Wir wollen deshalb zum Schluß die entsprechende Aussage anführen.

Definition. Jede n-fach multilineare Abbildung $\Delta : V^n \longrightarrow \mathbb{K}$ eines n-dimensionalen $\mathbb{K}$-Vektorraums V, die für beliebige Vektoren $x_1,\dots,x_n \in V$

$$\Delta(x_1,\dots,x_n) = 0 \iff x_1,\dots,x_n \text{ sind linear abhängig}$$

erfüllt, heißt eine *Determinantenform* auf V.

Bemerkungen. **(a)** Sind Δ_1 und Δ_2 Determinantenformen auf V, so existiert ein $c \in \mathbb{K}$ mit $\Delta_1 = c\,\Delta_2$. Insbesondere gibt es auf V genau eine Determinantenform, die auf einer gegebenen Basis den Wert 1 besitzt.

(b) Die Funktion Δ aus Satz 2 ist eine (normierte) Determinantenform auf $\mathbb{K}^n$. Sie ist durch die Eigenschaften (a), (c), (d) in Satz 2 eindeutig festgelegt. Die Eigenschaft alternierend zu sein folgt aus (a) und (d). Umgekehrt kann man (d) für alternierende Multilinearformen nur dann folgern, wenn die Charakteristik des Körpers $\mathbb{K}$ von 2 verschieden ist.

Zum Schluß wollen wir noch an die Bemerkung am Ende des Abschnittes über Matrizen in § I.4 anknüpfen. Die Definition der Determinante einer Matrix ist auch für Matrizen $A = (\!(a_{ij})\!)$ über einem kommutativen Ring mit 1 sinnvoll. Sind die Ringelemente a_{ij} speziell Polynome, $a_{ij} \in \mathbb{K}[X]$, so ist $\det A$ ebenfalls ein Polynom. Alle Ergebnisse dieses Paragraphen bleiben weiterhin richtig, solange sie nicht von der Division Gebrauch machen, insbesondere gelten die Sätze 4, 5 und 6 auch in diesem Fall.

§ 2 Eigenwerte und Diagonalisierbarkeit

Bei der geometrischen Beschreibung affiner Abbildungen in § 3.4 spielten die Fixpunkte und Fixrichtungen eine wesentliche Rolle. Die Bestimmung solcher Fixelemente führte auf Eigenwerte und Eigenvektoren der zugehörigen linearen Abbildung Φ. Wir werden nun allgemeiner untersuchen, welche Informationen über lineare Abbildungen Φ mit den Eigenwerten von Φ verbunden sind. In bestimmten Fällen lassen sich lineare Abbildungen vollständig über ihre Eigenwerte beschreiben.

Definition. Es seien V ein $\mathbb{K}$-Vektorraum und $\Phi : V \longrightarrow V$ eine lineare Abbildung. Der Skalar $c \in \mathbb{K}$ heißt *Eigenwert* von Φ, falls ein Vektor $v \in V$, $v \neq o$, existiert mit

$$\Phi(v) = c\, v.$$

Der Vektor v heißt *Eigenvektor* von Φ zum Eigenwert c.

Bemerkungen und Bezeichnungen. **(a)** Die Menge aller Eigenvektoren von Φ zu einem festen Eigenwert c bildet zusammen mit dem Nullvektor einen Untervektorraum von V. Er heißt der zu c gehörige *Eigenraum* von Φ und wird mit E_c bezeichnet. Offensichtlich ist $E_c = \mathrm{Kern}(\Phi - c \cdot \mathrm{id}_V)$. Die Menge aller Eigenwerte von Φ heißt das *Spektrum* von Φ.

(b) Es sei $A \in \mathbb{K}^{n\times n}$. Die Matrix A legt in eindeutiger Weise die lineare Abbildung $\Phi : \mathbb{K}^n \longrightarrow \mathbb{K}^n$, $x \longmapsto A\cdot x$, fest. Wir können deshalb auch von *Eigenwerten, Eigenvektoren* und *Eigenräumen* quadratischer Matrizen reden.

(c) Ähnliche Matrizen besitzen dieselben Eigenwerte.

Satz 9. *Es seien V ein $\mathbb{K}$-Vektorraum, $\Phi : V \longrightarrow V$ eine lineare Abbildung und $c_1,\dots,c_k$ paarweise verschiedene Eigenwerte von Φ mit zugehörigen Eigenvektoren $v_1,\dots,v_k$. Dann sind die Vektoren $v_1,\dots,v_k$ linear unabhängig.*

Beweis. Wir führen den Beweis durch vollständige Induktion nach k :

$k = 1$: Definition des Eigenvektors.

$k-1 \longrightarrow k$: Sei $a_1 v_1 + \cdots + a_k v_k = o$. Wir wenden Φ darauf an und erhalten

$$o = \Phi(o) = a_1 c_1 v_1 + \cdots + a_k c_k v_k .$$

Andererseits gilt

$$o = c_1 o = a_1 c_1 v_1 + \cdots + a_k c_1 v_k ,$$

woraus $(c_2 - c_1) a_2 v_2 + \cdots + (c_k - c_1) a_k v_k = o$ folgt und somit nach Induktionsvoraussetzung $(c_2 - c_1) a_2 = \cdots = (c_k - c_1) a_k = 0$. Damit ergibt sich der Reihe nach $a_2 = \cdots = a_k = 0$, $a_1 v_1 = o$ und schließlich $a_1 = 0$. ∎

Korollar 10. *Jeder Endomorphismus eines n–dimensionalen $\mathbb{K}$–Vektorraums besitzt höchstens n Eigenwerte. Jede Matrix $A \in \mathbb{K}^{n\times n}$ besitzt höchstens n Eigenwerte.*

Wir wollen uns nun mit der Frage beschäftigen, wie man bei einem Endomorphismus Φ das Spektrum und die Eigenräume bestimmen kann. Im allgemeinen ist dies ein schwieriges Problem, im endlich dimensionalen Fall läßt es sich jedoch mit Hilfe des im vorigen Paragraphen erklärten Begriffs der Determinante vereinfachen und in vielen Fällen auch lösen.

Seien also V ein n–dimensionaler Vektorraum, $\Phi \in \mathrm{Hom}\,(V, V)$, A_Φ die Abbildungsmatrix von Φ bezüglich einer (geordneten) Basis B und $\hat{v}$ die Koordinatendarstellung von $v \in V$ bezüglich derselben Basis B. Dann ist $\Phi(v) = c\, v$, $c \in \mathbb{K}$, gleichwertig mit $A_\Phi\, \hat{v} = c\, \hat{v}$. Es genügt daher, Eigenwerte von Matrizen zu betrachten. Nun gilt:

$$\begin{aligned} & c \text{ ist Eigenwert von } A \in \mathbb{K}^{n\times n} \\ \Longleftrightarrow\ & \text{es existiert } v \in \mathbb{K}^n,\ v \neq o, \text{ mit } A\,v = c\,v \text{ bzw. mit } (A - c\,E_n)\,v = o \\ \Longleftrightarrow\ & (A - c\,E_n)\,v = o \text{ ist nichttrivial lösbar} \\ \Longleftrightarrow\ & \mathrm{Rang}\,(A - c\,E_n) < n \Longleftrightarrow \det\,(A - c\,E_n) = 0 . \end{aligned}$$

Nach Definition der Determinante erhalten wir für $A = (\!(a_{ij})\!)$

$$\det\,(A - c\,E_n) = \sum_{\pi \in S_n} (-1)^{F(\pi)} (a_{\pi(1),1} - c\,\delta_{\pi(1),1}) \cdots (a_{\pi(n),n} - c\,\delta_{\pi(n),n}) .$$

Gemäß der Bemerkung von S. 174 können wir das Polynom

$$p = \det(A - X E_n) = \sum_{\pi \in S_n} (-1)^{F(\pi)} (a_{\pi(1),1} - X \delta_{\pi(1),1}) \cdots (a_{\pi(n),n} - X \delta_{\pi(n),n})$$

definieren. Es heißt *charakteristisches Polynom* (oder *Hauptpolynom*) von A.

Wir ordnen nach den Potenzen von X und erhalten

$$p = a_0 + a_1 X + a_2 X^2 + \cdots + a_{n-1} X^{n-1} + (-1)^n X^n$$

mit $a_0, \ldots, a_n \in \mathbb{K}$. Dabei ist speziell

$$a_0 = \det A ,$$

$$a_{n-1} = (-1)^{n-1}(a_{11} + \cdots + a_{nn}) = (-1)^{n-1} \operatorname{Spur} A ,$$

wobei Spur A definiert ist als Summe der Diagonalelemente von A.

Wir fassen zusammen:

Satz 11. *Genau dann ist $c \in \mathbb{K}$ ein Eigenwert der Matrix $A \in \mathbb{K}^{n \times n}$, wenn c Nullstelle des charakteristischen Polynoms p von A ist.*

Ist c ein Eigenwert von A, so ist der Eigenraum E_c gleich dem Lösungsraum des homogenen linearen Gleichungssystems $(A - c E_n) x = o$.

Bemerkung. Ähnliche Matrizen besitzen dasselbe charakteristische Polynom.

Beweis. Aus $A, B \in \mathbb{K}^{n \times n}$, $B = S^{-1} A S$, wobei $S \in \mathbb{K}^{n \times n}$ regulär ist, folgt

$$\det (S^{-1} A S - X E_n) = \det S^{-1} \cdot \det (A - X E_n) \cdot \det S = \det (A - X E_n). \qquad \blacksquare$$

Damit können wir auch vom *charakteristischen Polynom* eines Endomorphismus Φ reden, und Satz 11 gilt entsprechend.

Beispiele. **(a)** Sei $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$, $\mathbb{K} = \mathbb{R}$.

Es ist $p = \begin{vmatrix} 1 - X & 0 \\ 1 & 1 - X \end{vmatrix} = (1 - X)^2$, also ist $c = 1$ einziger Eigenwert von A. Für den Eigenraum E_1 erhalten wir $E_1 = [\begin{bmatrix} 0 \\ 1 \end{bmatrix}]$.

(b) Sei $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \in \mathbb{K}^{2\times 2}$.

Dann ist $p = \begin{vmatrix} -X & 1 \\ -1 & -X \end{vmatrix} = X^2 + 1.$

An diesem einfachen Beispiel sehen wir, daß die Existenz von Eigenwerten von dem zugrunde gelegten Körper abhängt.

Für $\mathbb{K} = \mathbb{R}$ existiert keine Nullstelle von p, also auch kein Eigenwert von A.

Für $\mathbb{K} = \mathbb{C}$ ergeben sich die Eigenwerte $c_1 = i$ und $c_2 = -i$. Die zugehörigen Eigenräume sind

$$E_{c_1} = [\begin{bmatrix} 1 \\ i \end{bmatrix}] \text{ und } E_{c_2} = [\begin{bmatrix} 1 \\ -i \end{bmatrix}].$$

Für $\mathbb{K} = \mathbb{F}_2$ ist $c = 1$ einziger Eigenwert mit zugehörigem Eigenraum $[\begin{bmatrix} 1 \\ 1 \end{bmatrix}]$.

(c) Gegeben sei die reelle Matrix

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

Es gilt:

$$p = \begin{vmatrix} -X & -1 & 1 & 1 \\ -1 & 1-X & -2 & 3 \\ 2 & -1 & -X & 0 \\ 1 & -1 & 1 & -X \end{vmatrix} = \begin{vmatrix} -X & -1 & 1 & 1 \\ 0 & -X & -1 & 3-X \\ 0 & 1 & -2-X & 2X \\ 1 & -1 & 1 & -X \end{vmatrix}$$

$$= \begin{vmatrix} -X & -1-X & 0 & 1 \\ 0 & -X & -1-X & 3-X \\ 0 & 1 & -1-X & 2X \\ 1 & 0 & 0 & -X \end{vmatrix} = (-X)(-X) \begin{vmatrix} -X & -1-X \\ 1 & -1-X \end{vmatrix} - \begin{vmatrix} -1-X & 0 & 1 \\ -X & -1-X & 3-X \\ 1 & -1-X & 2X \end{vmatrix}$$

$$= X^2(1+X)^2 + (1+X)[2X(-1-X) - (3-X)(-1-X)] - (X+1)^2$$

$$= (1+X)^2\,(X^2-3\,X+2) = (1+X)^2\,(1-X)(2-X).$$

Also sind $c_1 = -1$, $c_2 = 1$ und $c_3 = 2$ die Eigenwerte von A. Die zugehörigen Eigenräume sind

$$E_{c_1} = [\begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}]\ ,\ E_{c_2} = [\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}]\ ,\ E_{c_3} = [\begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}]\ .$$

Zum Schluß dieses Paragraphen wollen wir diejenigen Endomorphismen eines endlich dimensionalen Vektorraums untersuchen, die bezüglich einer geeigneten Basis eine Abbildungsmatrix besitzen, die Diagonalgestalt hat, bzw. solche quadratischen Matrizen, die ähnlich zu einer Diagonalmatrix sind.

Definition. Eine Matrix $A \in \mathbb{K}^{n\times n}$ heißt *diagonalisierbar*, wenn sie zu einer Diagonalmatrix

$$\begin{bmatrix} c_1 & 0 & \cdots\cdots & 0 \\ 0 & c_2 & 0 \cdots\cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots\ 0 & c_{n-1} & 0 \\ 0 & \cdots\cdots & 0 & c_n \end{bmatrix}$$

ähnlich ist. Ein Endomorphismus Φ heißt *diagonalisierbar*, wenn es eine Abbildungsmatrix von Φ gibt, die Diagonalgestalt hat.

Bemerkung. Da alle Abbildungsmatrizen eines Endomorphismus Φ ähnlich sind, ist im Falle der Diagonalisierbarkeit von Φ jede Abbildungsmatrix von Φ diagonalisierbar und außerdem zur *gleichen* Diagonalmatrix ähnlich. Es kann aber mehrere ähnliche Diagonalmatrizen geben

Im folgenden Satz stellen wir einfache äquivalente Bedingungen für Diagonalisierbarkeit zusammen.

Satz 12. *Für einen Endomorphismus Φ eines n-dimensionalen $\mathbb{K}$-Vektorraumes V sind folgende Aussagen äquivalent :*

(a) *Φ ist diagonalisierbar.*

(b) *In V gibt es eine Basis aus Eigenvektoren von Φ.*

(c) *V ist die direkte Summe der Eigenräume von Φ.*

(d) *Die Summe der Dimensionen der Eigenräume von Φ ist n.*

Beweis. (a) $\Longrightarrow$ (b): Nach Definition der Diagonalisierbarkeit gibt es eine Basis $B = (v_1,\dots,v_n)$ von V, bezüglich der die Abbildungsmatrix von Φ Diagonalgestalt

$$\begin{bmatrix} c_1 & 0 & \cdots & \cdots & 0 \\ 0 & c_2 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & c_{n-1} & 0 \\ 0 & \cdots & \cdots & 0 & c_n \end{bmatrix}$$

(mit nicht notwendig verschiedenen $c_1,\dots,c_n$) hat. Also gilt $\Phi(v_i) = c_i\, v_i$ für $i = 1,\dots,n$, d.h. die Vektoren v_i sind Eigenvektoren von Φ.

(b) $\Longrightarrow$ (c): Es seien $c_1,\dots,c_k$ die verschiedenen Eigenwerte von Φ. Dann ist nach Satz 9 die Summe der Eigenräume von Φ direkt, und es gilt $E_{c1} \oplus \cdots \oplus E_{ck} \subset V$. Sei nun x ein beliebiger Vektor aus V. Bezüglich der Basis $(v_1,\dots,v_n)$ aus Eigenvektoren von Φ gilt dann $x = a_1\, v_1 + \cdots + a_n\, v_n$. Fassen wir alle Eigenvektoren zum gleichen Eigenwert c_i zusammen, so erhalten wir $x = \tilde{v}_1 + \cdots + \tilde{v}_k$ mit $\tilde{v}_i \in E_{c_i}$, $i = 1,\dots,k$.

(c) $\Longrightarrow$ (d) ist trivial.

(d) $\Longrightarrow$ (a): Es seien $E_{c1},\dots,E_{ck}$ die Eigenräume von Φ und $\dim E_{c_i} = n_i$ für $i = 1,\dots,k$. In jedem Eigenraum E_{c_i} wählen wir eine Basis B_i. Dann ist $B = B_1 \cup \cdots \cup B_k$ nach Satz 9 linear unabhängig, und wegen $n_1 + \cdots + n_k = n$ ist B sogar Basis von V. Bezüglich dieser Basis hat die Abbildungsmatrix von Φ Diagonalgestalt. ∎

Bemerkungen. **(a)** Satz 12 gilt analog, wenn wir Φ durch $A \in \mathbb{K}^{n\times n}$ und V durch $\mathbb{K}^n$ ersetzen.

(b) Ist $A \in \mathbb{K}^{n\times n}$ diagonalisierbar, und ist $(v_1,\dots,v_n)$ eine Basis von $\mathbb{K}^n$ aus Eigenvektoren von A mit $A\, v_i = c_i\, v_i$ für $i = 1,\dots,n$, so gilt für die reguläre Matrix $S = (v_1 \mid \cdots \mid v_n)$

$$S^{-1} A\, S = \begin{bmatrix} c_1 & 0 & \cdots\cdots & 0 \\ 0 & c_2 & 0 \cdots\cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 \cdots & 0 & c_{n-1} & 0 \\ 0 & \cdots\cdots & 0 & c_n \end{bmatrix}$$

Beweis. Die obige Diagonalmatrix ist gerade die Abbildungsmatrix von $x \longmapsto A\, x$ bezüglich der neuen Basis $(v_1,\dots,v_n)$. ∎

Korollar 13. *Ein Endomorphismus Φ eines n-dimensionalen $\mathbb{K}$-Vektorraumes bzw. eine (n,n)-Matrix A mit n verschiedenen Eigenwerten ist diagonalisierbar.*

Es stellt sich nun die Frage, ob sich die Diagonalisierbarkeit von Φ bzw. A auch in anderen Fällen am charakteristischen Polynom p ablesen läßt. Wir zeigen dazu das folgende, für die Anwendung wichtigste Diagonalisierbarkeitskriterium.

Satz 14. *Es seien V ein n-dimensionaler $\mathbb{K}$-Vektorraum und $\Phi : V \longrightarrow V$ eine lineare Abbildung. Dann gilt:*

Φ ist genau dann diagonalisierbar, wenn sein charakteristisches Polynom p in der Form

$$(*) \qquad p = (-1)^n\, (X - c_1)^{r_1} \cdots (X - c_k)^{r_k}$$

darstellbar ist mit $r_i \in \mathbb{N}$ und paarweise verschiedenen $c_i \in \mathbb{K}$ und wenn für $i = 1,\dots,k$ gilt:

$$\dim \operatorname{Bild}\, (\Phi - c_i \cdot \mathrm{id}_V) = n - r_i\,.$$

Bezeichnung. Besitzt ein Polynom p die Darstellung $(*)$, so sagen wir auch, daß p in *Linearfaktoren zerfällt* und nennen r_i die *Vielfachheit* der Nullstelle c_i.

Bemerkungen. **(a)** Die zweite Forderung in Satz 14 besagt, daß die Dimension des Eigenraumes E_{c_i} mit der Vielfachheit r_i übereinstimmen muß $(i = 1,\dots,k)$.

(b) Satz 14 gilt entsprechend für $A \in \mathbb{K}^{n\times n}$, wenn wir $\dim \operatorname{Bild}\, (\Phi - c_i \cdot \mathrm{id}_V)$ durch $\operatorname{Rang}\, (A - c_i\, E_n)$ ersetzen.

Beweis von Satz 14. Es sei Φ diagonalisierbar. Dann besitzt Φ eine Abbildungsmatrix A_Φ der Form

$$A_\Phi = \begin{bmatrix} c_1 & 0 & \cdots & & & & 0 \\ 0 & \ddots & & & & & \vdots \\ \vdots & & c_1 & & & & \vdots \\ \vdots & & & \ddots & & & \vdots \\ \vdots & & & & c_k & & \vdots \\ \vdots & & & & & \ddots & 0 \\ 0 & \cdots & & & & 0 & c_k \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} r_1 \\ \\ \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} r_k \end{matrix}$$

wobei die c_i paarweise verschieden sind. Somit gilt

$$\det(A_\Phi - X E_n) = (c_1 - X)^{r1} \cdots (c_k - X)^{rk} = (-1)^n (X - c_1)^{r1} \cdots (X - c_k)^{rk}$$

und

$$\dim \operatorname{Bild}(\Phi - c_i \operatorname{id}_V) = \operatorname{Rang}(A_\Phi - c_i E_n)$$

$$= \operatorname{Rang} \begin{bmatrix} c_1 - c_i & 0 & \cdots & & & & 0 \\ 0 & \ddots & & & & & \vdots \\ \vdots & & c_1 - c_i & & & & \vdots \\ \vdots & & & \ddots & & & \vdots \\ \vdots & & & & c_k - c_i & & \vdots \\ \vdots & & & & & \ddots & 0 \\ 0 & \cdots & & & & 0 & c_k - c_i \end{bmatrix} = n - r_i .$$

Ist umgekehrt $p = (-1)^n (X - c_1)^{r1} \cdots (X - c_k)^{rk}$, so sind $c_1, \ldots c_k$ gerade die Eigenwerte von Φ. Wegen $\dim \operatorname{Bild}(\Phi - c_i \cdot \operatorname{id}_V) = n - r_i$ gilt $\dim E_{ci} = r_i$ und somit $\dim E_{c1} + \cdots + \dim E_{ck} = r_1 + \cdots + r_k = n$. Nach Satz 12 ist daher der Endomorphismus Φ diagonalisierbar. ∎

Die folgenden Beispiele setzen die Beispiele aus Abschnitt 4.2. fort.

Beispiele. **(a)** Wir betrachten die reelle Matrix

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

mit dem charakteristischen Polynom $p = (1 - X)^2$. Wegen Rang $(A - E_2) = 1 \neq n-2 = 0$ ist A nicht diagonalisierbar.

(b) Die Matrix

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

ist über $\mathbb{R}$ nicht diagonalisierbar. Über $\mathbb{C}$ gilt dagegen $p = (i - X)(-i - X)$, also hat A zwei verschiedene Eigenwerte und ist somit diagonalisierbar. Über $\mathbb{K} = \mathbb{F}_2$ ist A nicht diagonalisierbar, da 1 einziger Eigenwert von A ist und $\dim E_1 = 1$ gilt.

(c) Die reelle Matrix

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

hat das charakteristische Polynom $p = (1 + X)^2 (1 - X)(2 - X)$ und die Eigenwerte $c_1 = -1$, $c_2 = 1$ und $c_3 = 2$ mit den Vielfachheiten $r_1 = 2$, $r_2 = 1$ und $r_3 = 1$. Wie wir schon wissen, ist $\dim E_{c_1} = 1$. Wegen $1 \neq r_1$ ist A nicht diagonalisierbar.

(d) Die reelle Matrix

$$A = \begin{bmatrix} 3 & 2 & -1 \\ 2 & 6 & -2 \\ 0 & 0 & 2 \end{bmatrix}$$

hat das charakteristische Polynom $p = (2 - X)^2 (7 - X)$. Wegen

$$c_1 = 2,\ r_1 = 2,\ \text{Rang}\,(A - c_1 E_3) = 1 = 3 - 2\,,$$
$$c_2 = 7,\ r_2 = 1,\ \text{Rang}\,(A - c_2 E_3) = 2 = 3 - 1\,,$$

ist A diagonalisierbar. Ferner ist

$$E_{c_1} = [\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}]\ ,\ E_{c_2} = [\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}]\,.$$

Wir bilden aus den Basisvektoren von E_{c_1} und E_{c_2} die Transformationsmatrix

$$S = \begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & 2 \\ 1 & 0 & 0 \end{bmatrix}$$

und erhalten damit

$$S^{-1} A\, S = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 7 \end{bmatrix}$$

§ 3 Der Satz von Cayley – Hamilton

Wie wir im vorigen Paragraphen gesehen haben, gibt es zu jedem diagonalisierbaren Endomorphismus Φ eines endlich dimensionalen $\mathbb{K}$–Vektorraums V eine Darstellung von V als direkte Summe der Eigenräume $E_{c_1},\dots,E_{c_k}$ von Φ.

Wegen $E_{c_i} = \operatorname{Kern}(\Phi - c_i \cdot \mathrm{id}_V)$ gilt also

$$V = \operatorname{Kern}(\Phi - c_1 \cdot \mathrm{id}_V) \oplus \cdots \oplus \operatorname{Kern}(\Phi - c_k \cdot \mathrm{id}_V).$$

Wir betrachten nun den speziellen Endomorphismus

$$\Psi = (\Phi - c_1 \cdot \mathrm{id}_V) \circ \cdots \circ (\Phi - c_k \cdot \mathrm{id}_V).$$

Die einzelnen Abbildungen, aus denen Ψ zusammengesetzt ist, sind miteinander vertauschbar, denn es ist

$$\begin{aligned}(\Phi - c_i \cdot \mathrm{id}_V) \circ (\Phi - c_j \cdot \mathrm{id}_V) &= \Phi^2 - (c_i + c_j)\,\Phi + c_i\, c_j \cdot \mathrm{id}_V \\ &= (\Phi - c_j \cdot \mathrm{id}_V) \circ (\Phi - c_i \cdot \mathrm{id}_V).\end{aligned}$$

Es sei nun $x \in V$ ein beliebiger Vektor. Dieser besitzt dann eine Darstellung $x = x_1 + \cdots + x_k$ mit $x_i \in E_{c_i}$. Wenden wir Ψ darauf an, so erhalten wir

$$\begin{aligned}\Psi(x) &= \Psi(x_1) + \cdots + \Psi(x_k) \\ &= (\Phi - c_2 \cdot \mathrm{id}_V) \circ \cdots \circ (\Phi - c_k \cdot \mathrm{id}_V) \circ (\Phi - c_1 \cdot \mathrm{id}_V)(x_1) \\ &\quad + \cdots + (\Phi - c_1 \cdot \mathrm{id}_V) \circ \cdots \circ (\Phi - c_{k-1} \cdot \mathrm{id}_V) \circ (\Phi - c_k \cdot \mathrm{id}_V)(x_k) \\ &= o + \cdots + o = o\,,\end{aligned}$$

d.h. Ψ ist die Nullabbildung auf V.

Den Endomorphismus Ψ können wir uns auch so entstanden denken, daß wir in das Polynom $m = (X - c_1) \cdots (X - c_k)$ den Endomorphismus Φ eingesetzt haben.

Das Polynom m ist ein *annullierendes Polynom*, weil $m(\Phi)$ die Nullabbildung in V ist. Wie wir später sehen werden, ist in diesem speziellen Fall m das normierte Polynom kleinsten Grades, für das $m(\Phi) = O$ gilt. Es wird deshalb auch *Minimalpolynom* von Φ genannt.

Es stellt sich die Frage, ob es zu jedem Endomorphismus Φ von V bzw. zu jeder Matrix $A \in \mathbb{K}^{n\times n}$ ein solches Polynom gibt. Bevor wir diese Frage beantworten, wollen wir noch einmal präzisieren, was unter dem Einsetzen eines Endomorphismus bzw. einer quadratischen Matrix in ein Polynom verstanden werden soll.

Definition. Es seien V ein $\mathbb{K}$–Vektorraum, Φ ein Endomorphismus von V sowie

$$q = a_0 + a_1 X + a_2 X^2 + \cdots + a_n X^n \in \mathbb{K}[X]$$

ein Polynom. Dann sei $q(\Phi)$ der Endomorphismus

$$q(\Phi) := a_0 \cdot \mathrm{id}_V + a_1 \cdot \Phi + a_2 \cdot \Phi^2 + \ldots + a_n \cdot \Phi^n,$$

wobei $\Phi^k = \underbrace{\Phi \circ \Phi \circ \cdots \circ \Phi}_{k \text{ Faktoren}}$ gilt.

Analog gilt für $A \in \mathbb{K}^{n\times n}$

$$q(A) := a_0 \cdot E_n + a_1 \cdot A + a_2 \cdot A^2 + \ldots + a_n \cdot A^n$$

mit $A^k = \underbrace{A \cdot A \cdots A}_{k \text{ Faktoren}}$.

Beispiel. Wir setzen eine Matrix A bzw. einen Endomorphismus Φ in das Polynom $q = 3 + X + X^4$ ein und erhalten $q(A) = 3\,E_n + A + A^4$ bzw. $q(\Phi) = 3\,\mathrm{id}_V + \Phi + \Phi^4$.

Bemerkung. Sei $\Phi \in \mathrm{Hom}(V, V)$ fest gewählt. Dann ist die Abbildung

$$f_\Phi : \mathbb{K}[X] \longrightarrow \mathrm{Hom}(V, V)$$

$$q \longmapsto q(\Phi)$$

ein Homomorphismus bezüglich der Vektorraumstruktur und der Ringstruktur, der sogenannte *Einsetzungshomomorphismus*, denn es gilt für alle $p,q \in \mathbb{K}[X]$ und alle $a \in \mathbb{K}$

$$(p+q)(\Phi) = p(\Phi) + q(\Phi),$$

$$(a \cdot p)(\Phi) = a \cdot p(\Phi),$$

$$(p \cdot q)(\Phi) = p(\Phi) \circ q(\Phi).$$

Die ersten beiden Eigenschaften sind trivial; f_Φ ist also linear. Es genügt deswegen, die letzte Eigenschaft auf der Basis $\{1, X, X^2, \ldots\}$ von $\mathbb{K}[X]$ nachzuprüfen. Dort ist sie aber unmittelbar klar.

Entsprechend wird bei festem $A \in \mathbb{K}^{n \times n}$ der Einsetzungshomomorphismus von $\mathbb{K}[X]$ in $\mathbb{K}^{n \times n}$ erklärt.

Man beachte, daß $f_\Phi(\mathbb{K}[X]) \subset \operatorname{Hom}(V,V)$ als Bild des kommutativen Rings $\mathbb{K}[X]$ ebenfalls kommutativ ist, der Ring $\operatorname{Hom}(V,V)$ dagegen nicht.

Wir wollen nun den Endomorphismus Φ in ein spezielles Polynom einsetzen, nämlich in sein charakteristisches Polynom.

Satz 15 (Cayley–Hamilton). *Es seien V ein n-dimensionaler $\mathbb{K}$-Vektorraum, Φ ein Endomorphismus von V und p das charakteristische Polynom von Φ. Dann ist $p(\Phi) = O$.*

Bemerkung. Man beachte, daß in der Behauptung rechts vom Gleichheitszeichen die Nullabbildung steht. Es ist also schon daher unsinnig, einen "Beweis" folgender Art führen zu wollen: $p(\Phi) = \det(\Phi - \Phi \circ \mathrm{id}_V) = \det(O) = 0$.

Beweis. Es ist zu zeigen, daß $p(\Phi)(v) = o$ für alle $v \in V$ gilt. Für $v = o$ ist dies klar. Sei also $v \neq o$. Wir betrachten nun für jedes $m \in \mathbb{N}_0$ die Vektoren

$$v, \Phi(v), \Phi^2(v), \ldots, \Phi^m(v),$$

wobei wir die Konvention $\Phi^0 := \mathrm{id}_V$ benutzen. Für $m = 0$ sind diese Vektoren linear unabhängig, für $m \geq n$ linear abhängig. Also existiert ein kleinstes $m \in \mathbb{N}$, für das die Vektoren linear abhängig sind. Dann ist

$$\tilde{B} = \{v, \Phi(v), \Phi^2(v), \dots, \Phi^{m-1}(v)\}$$

linear unabhängig und

$$\tilde{B} \cup \{\Phi^m(v)\}$$

linear abhängig. Nach Kapitel 2 existieren also Skalare $a_0, \dots, a_{m-1} \in \mathbb{K}$ mit

$$\Phi^m(v) = a_0 v + a_1 \Phi(v) + \dots + a_{m-1} \Phi^{m-1}(v) .$$

Sei $q = a_0 + a_1 X + \dots + a_{m-1} X^{m-1} + (-1) X^m \in \mathbb{K}[X]$. Dann ist $q(\Phi) \in \operatorname{Hom}(V, V)$, und es gilt $q(\Phi)(v) = o$.

Wir wollen jetzt zeigen, daß auch $p(\Phi)(v) = o$ gilt. Dazu setzen wir $U := [\tilde{B}]$. Wegen $\Phi(\tilde{B}) \subset U$ ist $\Phi(U) \subset U$. Somit ist $\tilde{\Phi} := \Phi|_U$ ein Endomorphismus von U. Bezüglich der (geordneten) Basis $\tilde{B}$ besitzt $\tilde{\Phi}$ die Abbildungsmatrix

$$\tilde{A} = \begin{bmatrix} 0 & 0 & \cdots & 0 & a_0 \\ 1 & 0 & \cdots & 0 & a_1 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & a_{m-2} \\ 0 & \cdots & 0 & 1 & a_{m-1} \end{bmatrix} \in \mathbb{K}^{m \times m} .$$

Wir bestimmen das charakteristische Polynom $\tilde{p}$ von $\tilde{\Phi}$.

$$\det(\tilde{A} - X E_m) = \begin{vmatrix} -X & 0 & \cdots & 0 & a_0 \\ 1 & -X & \cdots & 0 & a_1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & -X & a_{m-2} \\ 0 & \cdots & 0 & 1 & a_{m-1} - X \end{vmatrix}$$

$$= \cdots = \begin{vmatrix} 0 & \cdots\cdots & 0 & a_0 + a_1 X + \cdots + a_{m-1} X^{m-1} - X^m \\ 1 & \ddots & \vdots & \vdots \\ 0 & \ddots \ddots & \vdots & \vdots \\ \vdots & \ddots \ddots & 0 & a_{m-2} + a_{m-1} X - X^2 \\ 0 & \cdots\; 0 & 1 & a_{m-1} - X \end{vmatrix} = (-1)^{m+1}\, q\,.$$

Also ist $\tilde{p} = (-1)^{m+1}\, q$. Wir ergänzen nun $\tilde{B}$ zu einer Basis B von V. Der Endomorphismus Φ hat bezüglich B eine Abbildungsmatrix A_Φ der Form

$$A_\Phi = \begin{bmatrix} \tilde{A} & C \\ O & D \end{bmatrix}$$

mit geeigneten Matrizen C und D. Für das charakteristische Polynom p von Φ folgt nach dem Kästchenmultiplikationssatz für Determinanten $p = \tilde{p}\cdot\overline{p}$, wobei $\overline{p}$ das charakteristische Polynom von D ist. Setzen wir $r = (-1)^{m+1}\, \overline{p}$, so gilt $p = q\cdot r = r\cdot q$ und $p\,(\Phi) = r\,(\Phi) \circ q\,(\Phi)$. Daraus folgt $p\,(\Phi)(v) = r\,(\Phi)(q\,(\Phi)(v)) = o$. Da v beliebig war, gilt $p\,(\Phi) = O$. ■

Bemerkungen und Definition. **(a)** Satz 15 gilt analog für $A \in \mathbb{K}^{n\times n}$.

(b) Die Bedeutung dieses Satzes liegt nicht darin, daß es überhaupt ein Polynom $q \neq o$ gibt mit $q\,(\Phi) = O$, sondern daß das charakteristische Polynom von Φ diese Eigenschaft hat. Ein Polynom $q \neq o$ mit $q\,(\Phi) = O$ findet man nämlich leicht:

Wegen $\dim \operatorname{Hom}(V, V) = n^2$, sind die $n^2 + 1$ Vektoren $\mathrm{id}_V, \Phi, \Phi^2, \ldots, \Phi^{n^2}$ aus $\operatorname{Hom}(V, V)$ linear abhängig. Also gibt es Skalare $a_0, a_1, \ldots, a_{n^2}$, die nicht alle Null sind, mit

$$a_0\, \mathrm{id}_V + a_1\, \Phi + \cdots + a_{n^2}\, \Phi^{n^2} = O\,.$$

Für das Polynom $q = a_0 + a_1\, X + \cdots + a_{n^2}\, X^{n^2}$ gilt dann $q \neq o$ und $q\,(\Phi) = O$.

(c) Das normierte Polynom $m \in \mathbb{K}[X]$ vom kleinsten Grad, das $m\,(\Phi) = O$ erfüllt, heißt *Minimalpolynom* von Φ. Es gibt ein und nur ein derartiges Polynom:

Seien nämlich m und $\tilde{m}$ normierte Polynome vom kleinsten Grad r mit $m\,(\Phi) = \tilde{m}\,(\Phi) = O$. Ist $m \neq \tilde{m}$, so ist $m - \tilde{m}$ vom Nullpolynom verschieden und hat einen Grad, der kleiner als r ist. Normieren wir $m - \tilde{m}$ noch, so erhalten wir ein annullierendes Polynom, dessen Grad kleiner als r ist. Dies ist ein Widerspruch zur Wahl von r. Also ist $m = \tilde{m}$. ∎

Analog wird das *Minimalpolynom* einer Matrix $A \in \mathbb{K}^{n \times n}$ erklärt.

Wegen $\Phi^0 = \mathrm{id}_V \neq O$ ist der Grad von m immer mindestens eins.

(d) Ähnliche Matrizen haben das gleiche Minimalpolynom, denn sie können als Abbildungsmatrizen eines festen Endomorphismus aufgefaßt werden.

Zwischen dem Minimalpolynom eines Endomorphismus Φ und den Polynomen $q \in \mathbb{K}[X]$ mit $q\,(\Phi) = O$ besteht folgender Zusammenhang:

Satz 16. *Das Minimalpolynom m des Endomorphismus Φ teilt jedes Polynom $q \in \mathbb{K}[X]$, für das $q\,(\Phi) = O$ gilt.*

Beweis. Wegen $m \neq o$ können wir nach Satz 1.14 q durch m dividieren: $q = s \cdot m + r$ mit $s,r \in \mathbb{K}[X]$. Dabei ist der Grad von r kleiner als der von m. Wir setzen Φ ein und erhalten $r\,(\Phi) = q\,(\Phi) - s\,(\Phi) \circ m\,(\Phi) = O$. Also muß r das Nullpolynom sein. ∎

Korollar 17. **(a)** *m teilt das charakteristische Polynom p von Φ.*

(b) *Die Nullstellen von m sind genau die Nullstellen von p, also gerade die Eigenwerte von Φ.*

Beweis von (b). Aus $m\,(c) = 0$ folgt wegen (a) $p\,(c) = 0$, und c ist Eigenwert von Φ. Ist umgekehrt c Nullstelle von p, also Eigenwert von Φ, so existiert ein Vektor $v \neq o$ mit $\Phi(v) = c\,v$. Für $i = 1,2,\ldots$ gilt dann $\Phi^i(v) = c^i\,v$, woraus $m\,(\Phi)(v) = m\,(c)\,v$ folgt. Also ist $m\,(c)\,v = o$ und wegen $v \neq o$ somit $m\,(c) = 0$. ∎

Bemerkungen. **(a)** Ist p von der Form

$$p = (-1)^n (X - c_1)^{r_1} \cdots (X - c_k)^{r_k}, \qquad r_i \geq 1,$$

mit paarweise verschiedenen $c_1, \ldots, c_k \in \mathbb{K}$, so kann m nur die Form

$$m = (X - c_1)^{s_1} \cdots (X - c_k)^{s_k}$$

haben mit $1 \leq s_i \leq r_i$.

(b) Ist Φ diagonalisierbar und sind $c_1, \ldots, c_k$ die Eigenwerte von Φ, so ist

$$m = (X - c_1) \cdots (X - c_k)$$

ein annullierendes Polynom und somit wegen Korollar 17 das Minimalpolynom von Φ. Die Umkehrung gilt auch, wie wir im nächsten Abschnitt sehen werden.

Beispiele. **(a)** Wir betrachten den Endomorphismus $\Phi : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$, $x \longmapsto A\,x$, mit

$$A = \begin{bmatrix} 7 & -6 & 11 \\ 0 & 1 & -1 \\ -4 & 4 & -7 \end{bmatrix}$$

Es ist

$$p = \begin{vmatrix} 7-X & -6 & 11 \\ 0 & 1-X & -1 \\ -4 & 4 & -7-X \end{vmatrix} = (7-X) \begin{vmatrix} 1-X & -1 \\ 4 & -7-X \end{vmatrix} - 4 \begin{vmatrix} -6 & 11 \\ 1-X & -1 \end{vmatrix}$$

$$= -(X-1)^2 (X+1).$$

Für das Minimalpolynom m gibt es nach Korollar 17 somit nur die Möglichkeiten $m = (X-1)(X+1)$ oder $m = (X-1)^2 (X+1)$. Wegen

$$(A - E_3)(A + E_3) = \begin{bmatrix} 6 & -6 & 11 \\ 0 & 0 & -1 \\ -4 & 4 & -8 \end{bmatrix} \begin{bmatrix} 8 & -6 & 11 \\ 0 & 2 & -1 \\ -4 & 4 & -6 \end{bmatrix} \neq O$$

ist $m = (X-1)^2\,(X+1) = -p$ das Minimalpolynom von Φ.

Im nächsten Paragraphen werden wir für den Fall, daß das charakteristische Polynom in Linearfaktoren zerfällt ein allgemeines Verfahren zur Bestimmung des Minimalpolynoms herleiten.

(b) Der Satz von Cayley–Hamilton ist auch bei dem Problem nützlich, für ein Polynom q und eine quadratische Matrix A die Matrix $q(A)$ zu bestimmen. Ist z. B.

$$q \;=\; -1 + 2\,X + 6\,X^2 - 4\,X^3 - X^4$$

und A die Matrix aus Beispiel (a), so verfahren wir folgendermaßen:

Wir dividieren q durch das charakteristische Polynom $p = -\,(X-1)^2\,(X+1)$ von A und erhalten

$$q \;=\; (X+5)\,p - 2\,X + 4.$$

Also gilt

$$q\,(A) \;=\; (A + 5\,E_3)\,p\,(A) - 2\,A + 4\,E_3 \;=\; -2\,A + 4\,E_3$$

und daher

$$q\,(A) \;=\; \begin{bmatrix} -10 & 12 & -22 \\ 0 & 2 & 2 \\ 8 & -8 & 18 \end{bmatrix}.$$

Ist das Minimalpolynom m bekannt und Grad $m <$ Grad p, so arbeitet man vorteilhafter mit m.

§ 4 Jordansche Normalform

Wir haben in § 2 gesehen, daß die diagonalisierbaren Endomorphismen bei geeigneter Basiswahl eine Diagonalmatrix als Abbildungsmatrix besitzen. In diesem Abschnitt wollen wir versuchen, auch für nichtdiagonalisierbare Endomorphismen eine möglichst einfache Abbildungsmatrix zu erhalten.

Wie wir wissen, läßt sich bei einem diagonalisierbaren Endomorphismus Φ der Vektorraum V in eine direkte Summe von Untervektorräumen zerlegen:

$$V = E_{c_1} \oplus \cdots \oplus E_{c_k}.$$

Es gilt $\Phi(E_{c_i}) \subset E_{c_i}$. Untervektorräume $U \subset V$, die $\Phi(U) \subset U$ erfüllen, heißen Φ*-invariant*.

Unser nächstes Ziel ist es, auch für nichtdiagonalisierbares $\Phi \in \operatorname{End}(V)$ Φ-invariante Untervektorräume $V_1, \ldots, V_k$ zu finden, so daß

$$V = V_1 \oplus \cdots \oplus V_k$$

gilt und die Einschränkungen $\Phi|_{V_1}, \ldots, \Phi|_{V_k}$ möglichst einfache Abbildungsmatrizen besitzen. Analog zu den Eigenräumen E_{c_i} versuchen wir, die Untervektorräume V_i in der Form $V_i = \operatorname{Kern} q_i(\Phi)$ mit geeigneten Polynomen $q_i \in \mathbb{K}[X]$ darzustellen. Dazu betrachten wir zunächst die speziellen Untervektorräume $\operatorname{Kern} q(\Phi)$ mit $q \in \mathbb{K}[X]$.

Bemerkungen. **(a)** Für jedes Polynom $q \in \mathbb{K}[X]$ ist $\operatorname{Kern} q(\Phi)$ ein Φ-invarianter Untervektorraum von V.

Beweis. Aus $v \in \operatorname{Kern} q(\Phi)$ folgt $q(\Phi)(v) = o$ und somit $q(\Phi)(\Phi(v)) = (\Phi \circ q(\Phi))(v) = o$. Also ist $\Phi(v) \in \operatorname{Kern} q(\Phi)$. ■

(b) Ist das Polynom $q \in \mathbb{K}[X]$ ein Teiler von $r \in \mathbb{K}[X]$, so folgt $\operatorname{Kern} q(\Phi) \subset \operatorname{Kern} r(\Phi)$.

Beweis. Es ist $r = s \cdot q$ mit $s \in \mathbb{K}[X]$. Aus $v \in \operatorname{Kern} q(\Phi)$ folgt dann $r(\Phi)(v) = s(\Phi)(q(\Phi)(v)) = o$ und somit $v \in \operatorname{Kern} r(\Phi)$. ■

(c) Sind $q, r \in \mathbb{K}[X]$ teilerfremde Polynome, so gilt

$$\text{Kern } (q \cdot r)\,(\Phi) \;=\; \text{Kern } q\,(\Phi) \oplus \text{Kern } r\,(\Phi).$$

Beweis. Sind q und r teilerfremd, so gibt es nach Satz 1.16 Polynome s und t mit $s \cdot q + t \cdot r = 1$. Setzen wir Φ ein, so erhalten wir $s\,(\Phi) \circ q\,(\Phi) + t\,(\Phi) \circ r\,(\Phi) = \mathrm{id}_V$ und somit für $v \in V$

$$v \;=\; \underbrace{s(\Phi)(q(\Phi)(v))}_{y} \;+\; \underbrace{t(\Phi)(r(\Phi)(v))}_{z}.$$

Aus $v \in \text{Kern } (q \cdot r)(\Phi)$ folgt

$$r\,(\Phi)(y) = (r\,(\Phi) \circ s\,(\Phi) \circ q\,(\Phi))(v) = (s\,(\Phi) \circ q\,(\Phi) \circ r\,(\Phi))(v) = s\,(\Phi)(o) = o\,,$$

also $y \in \text{Kern } r\,(\Phi)$. Ebenso erhalten wir $z \in \text{Kern } q\,(\Phi)$. Mit (b) ergibt sich daher

$$\text{Kern } (q \cdot r)\,(\Phi) \;=\; \text{Kern } q\,(\Phi) + \text{Kern } r\,(\Phi).$$

Schließlich folgt aus $v \in \text{Kern } r\,(\Phi) \cap \text{Kern } q\,(\Phi)$ direkt $z = y = o$ und somit $v = o$. Also ist die Summe direkt. ■

(d) Sind $q_1, \ldots, q_k$ paarweise teilerfremde Polynome, so gilt

$$\text{Kern } (q_1 \cdots q_k)(\Phi) \;=\; \text{Kern } q_1\,(\Phi) \oplus \cdots \oplus \text{Kern } q_k\,(\Phi).$$

Beweis. Nach Bemerkung (a) von S. 67 sind die Polynome $q_1 \cdots q_{k-1}$ und q_k teilerfremd. Durch vollständige Induktion und mit (c) folgt die Behauptung. ■

Die Bemerkungen (a),(b),(c) und (d) lassen sich insbesondere auf Polynome $q \in \mathbb{K}[X]$ anwenden, für die $q(\Phi) = O$ gilt. Wir erhalten so die folgende Aussage:

Satz 18. *Es seien V ein $\mathbb{K}$-Vektorraum,* dim $V = n$ *und $\Phi : V \longrightarrow V$ eine lineare Abbildung. Weiterhin sei $q \in \mathbb{K}[X]$ ein Polynom mit $q(\Phi) = O$ und $q = q_1 \cdots q_k$ eine Zerlegung von q in paarweise teilerfremde Faktoren $q_i \in \mathbb{K}[X]$, $i = 1, \ldots, k$. Dann sind die Unterevektorräume $V_i =$* Kern $q_i\,(\Phi)$ *Φ-invariant und es gilt:*

$$V \;=\; V_1 \oplus \cdots \oplus V_k\,.$$

Bei geeigneter Basiswahl hat die zugehörige Abbildungsmatrix A von Φ die Form

$$A = \begin{bmatrix} \boxed{A_1} & & O \\ & \ddots & \\ O & & \boxed{A_k} \end{bmatrix},$$

wobei A_i die Abbildungsmatrix von $\Phi|_{V_i}$ ist.

Bemerkung. Speziell für das Minimalpolynom m von Φ gilt über Satz 18 hinaus:

Ist $m = m_1 \cdots m_k$ eine Zerlegung von m in teilerfremde normierte Faktoren $m_i \in \mathbb{K}[X]$, $i = 1,\dots,k$, so ist m_i das Minimalpolynom von $\Phi|_{\text{Kern } m_i(\Phi)}$.

Beweis. Nach Satz 18 läßt sich jeder Vektor $v \in V$ in der Form $v = x_1 + \cdots + x_k$ mit $x_i \in \text{Kern } m_i(\Phi)$ darstellen. Wäre nun $\tilde{m}_i$ Minimalpolynom von $\Phi|_{\text{Kern } m_i(\Phi)}$ und $\tilde{m}_i \neq m_i$, so wäre Grad $\tilde{m}_i <$ Grad m_i und für das Polynom

$$\tilde{m} = m_1 \cdots m_{i-1} \cdot \tilde{m}_i \cdot m_{i+1} \cdots m_k$$

würde gelten:

$$\tilde{m}(\Phi)(v) = \tilde{m}(\Phi)(x_1) + \cdots + \tilde{m}(\Phi)(x_k) = m_2(\Phi) \circ \dots \circ m_k(\Phi) \circ m_1(\Phi)(x_1) + \cdots +$$
$$m_1(\Phi) \circ \dots \circ m_{i-1}(\Phi) \circ m_{i+1}(\Phi) \circ \dots \circ m_k(\Phi) \circ \tilde{m}_i(\Phi)(x_i) + \cdots +$$
$$m_1(\Phi) \circ \dots \circ m_k(\Phi)(x_k) = o$$

Also wäre $\tilde{m}(\Phi) = O$ und Grad $\tilde{m} <$ Grad m. Das ist ein Widerspruch, und somit ist m_i Minimalpolynom von $\Phi|_{\text{Kern } m_i(\Phi)}$. ■

Als Folgerung aus Satz 18 erhalten wir ein weiteres Diagonalisierbarkeitskriterium.

Satz 19. *Es seien V ein $\mathbb{K}$-Vektorraum,* dim $V = n$ *und* $\Phi \in$ Hom (V, V). *Dann gilt:*

Φ ist genau dann diagonalisierbar, wenn das Minimalpolynom m in einfache

Linearfaktoren zerfällt, d.h. die Form

$$m = (X-c_1)\cdots(X-c_k)$$

hat mit paarweise verschiedenen $c_1,\dots,c_k \in \mathbb{K}$.

Beweis. Wegen Bemerkung (b) von S. 191 müssen wir nur noch eine Richtung beweisen. Hat m die obige Form, so folgt nach Satz 18, daß $V = V_1 \oplus\cdots\oplus V_k$ ist mit $V_i = \text{Kern}\,(\Phi - c_i\,\mathrm{id}_V) = E_{c_i}$. Nach Satz 12 ist Φ dann diagonalisierbar. ■

Bemerkung. Ist $A \in \mathbb{K}^{n\times n}$, so gelten für die zugehörige lineare Abbildung $\Phi : \mathbb{K}^n \longrightarrow \mathbb{K}^n$, $\Phi(x) = A\,x$, die Sätze 18 und 19 sowie die Bemerkung nach Satz 18.

Insbesondere ist also A genau dann zu einer Diagonalmatrix ähnlich, wenn sein Minimalpolynom in einfache Linearfaktoren zerfällt.

Es sei nun c ein Eigenwert des Endomorphismus Φ mit der Vielfachheit r. Für das charakteristische Polynom von Φ gilt also

$$p = (X-c)^r\cdot \overline{p}$$

mit $\overline{p}\,(c) \neq 0$, und für das Minimalpolynom m von Φ gilt dann wegen Korollar 17

$$m = (X-c)^s\cdot \overline{m}$$

mit $1 \leq s \leq r$, $\overline{m}(c) \neq 0$ und $\overline{m} \mid \overline{p}$. Nach Satz 18 ist dann

$$V = \text{Kern}\,(\Phi - c\,\mathrm{id}_V)^r \oplus \text{Kern}\,\overline{p}\,(\Phi) = \text{Kern}\,(\Phi - c\,\mathrm{id}_V)^s \oplus \text{Kern}\,\overline{m}\,(\Phi),$$

woraus wegen

$$\text{Kern}\,(\Phi - c\,\mathrm{id}_V)^s \subset \text{Kern}\,(\Phi - c\,\mathrm{id}_V)^r$$

und

$$\text{Kern}\,\overline{m}\,(\Phi) \subset \text{Kern}\,\overline{p}(\Phi).$$

aus Dimensionsgründen

$$\text{Kern}\,(\Phi - c\,\mathrm{id}_V)^s = \text{Kern}\,(\Phi - c\,\mathrm{id}_V)^r$$

und

$$\operatorname{Kern} \overline{m}(\Phi) = \operatorname{Kern} \overline{p}(\Phi)$$

folgen.

Definition. Der Untervektorraum $H_c := \operatorname{Kern}(\Phi - c\,\mathrm{id}_V)^r$ heißt *Hauptraum* zum Eigenwert c von Φ, die Vielfachheit s der Nullstelle c des Minimalpolynoms m von Φ heißt *Index* des Hauptraums H_c.

Wir werden später zeigen, daß $\dim H_c = r$ gilt und daß $(-1)^r (X - c)^r$ gerade das charakteristische Polynom von $\Phi|_{H_c}$ ist.

Der Index des Hauptraumes H_c kann folgendermaßen berechnet werden.

Satz 20. *Es seien V ein $\mathbb{K}$-Vektorraum,* $\dim V = n$*, $\Phi \in \operatorname{Hom}(V,V)$ und c ein Eigenwert von Φ. Dann ist der Index des Hauptraums H_c zum Eigenwert c die kleinste Zahl $s \in \mathbb{N}$, für die*

$$\operatorname{Kern}(\Phi - c\,\mathrm{id}_V)^s = \operatorname{Kern}(\Phi - c\,\mathrm{id}_V)^{s+1}$$

gilt.

Beweis. Für den Index s von H_c und alle $k \in \mathbb{N}$ gilt

$$\operatorname{Kern}(\Phi - c\,\mathrm{id}_V)^s = \operatorname{Kern}(\Phi - c\,\mathrm{id}_V)^{s+k},$$

denn für das Polynom $\tilde{m} = (X - c)^{s+k}\,\overline{m}$ folgt wegen $\tilde{m}(\Phi) = (\Phi - c\,\mathrm{id}_V)^k \circ m(\Phi) = O$ nach Satz 18

$$V = \operatorname{Kern}(\Phi - c\,\mathrm{id}_V)^{s+k} \oplus \operatorname{Kern} \overline{m}(\Phi),$$

woraus wegen

$$\operatorname{Kern}(\Phi - c\,\mathrm{id}_V)^s \subset \operatorname{Kern}(\Phi - c\,\mathrm{id}_V)^{s+k}$$

aus Dimensionsgründen die Gleichheit folgt.

Ist nun $\overline{s}$ die kleinste natürliche Zahl mit

$$\operatorname{Kern}(\Phi - c\,\mathrm{id}_V)^{\overline{s}} = \operatorname{Kern}(\Phi - c\,\mathrm{id}_V)^{\overline{s}+1},$$

so gilt $\overline{s} \leq s$, und durch vollständige Induktion erhalten wir für alle $k \in \mathbb{N}$

$$\operatorname{Kern}\,(\Phi - c\,\mathrm{id}_V)^{\overline{s}} = \operatorname{Kern}\,(\Phi - c\,\mathrm{id}_V)^{\overline{s}+k}.$$

Sei nämlich $v \in \operatorname{Kern}\,(\Phi - c\,\mathrm{id}_V)^{\overline{s}+k}$. Dann ist

$$(\Phi - c\,\mathrm{id}_V)(v) \in \operatorname{Kern}\,(\Phi - c\,\mathrm{id}_V)^{\overline{s}+k-1} = \operatorname{Kern}\,(\Phi - c\,\mathrm{id}_V)^{\overline{s}},$$

also

$$v \in \operatorname{Kern}\,(\Phi - c\,\mathrm{id}_V)^{\overline{s}+1} = \operatorname{Kern}\,(\Phi - c\,\mathrm{id}_V)^{\overline{s}}.$$

Aus $\overline{s} < s$ würde nun

$$\operatorname{Kern}\,(\Phi - c\,\mathrm{id}_V)^{\overline{s}} = \operatorname{Kern}\,(\Phi - c\,\mathrm{id}_V)^{s} = H_c$$

folgen, und somit wäre $(X - c)^{\overline{s}}\,\overline{m}$ annullierendes Polynom von Φ im Widerspruch dazu, daß $m = (X - c)^s\,\overline{m}$ das Minimalpolynom von Φ ist. Also folgt $\overline{s} = s$. ■

Bemerkungen. **(a)** Für den Index s des Hauptraums H_c gibt es folgende weitere äquivalente Charakterisierungen:

$s \in \mathbb{N}$ ist die kleinste Zahl mit $\dim \operatorname{Kern}\,(\Phi - c\,\mathrm{id}_V)^s = \dim \operatorname{Kern}\,(\Phi - c\,\mathrm{id}_V)^{s+1}$,

$s \in \mathbb{N}$ ist die kleinste Zahl mit $\dim \operatorname{Bild}\,(\Phi - c\,\mathrm{id}_V)^s = \dim \operatorname{Bild}\,(\Phi - c\,\mathrm{id}_V)^{s+1}$,

$s \in \mathbb{N}$ ist die kleinste Zahl mit $\operatorname{Rang}\,(A_\Phi - c\,E_n)^s = \operatorname{Rang}\,(A_\Phi - c\,E_n)^{s+1}$.

(b) Für Matrizen $A \in \mathbb{K}^{n\times n}$ werden der Hauptraum H_c zum Eigenwert c und der Index von H_c entsprechend erklärt. Es ist

$$H_c = \{v \in \mathbb{K}^n \mid (A - c\,E_n)^r\, v = o\},$$

wobei r die Vielfachheit der Nullstelle c im charakteristischen Polynom von A ist, und entsprechend zu Satz 20 folgt dann für den Index des Hauptraumes H_c, daß er die kleinste natürliche Zahl s ist mit

$$\operatorname{Rang}\,(A - c\,E_n)^s = \operatorname{Rang}\,(A - c\,E_n)^{s+1}.$$

(c) Falls das charakteristische Polynom p von Φ bzw. von A vollständig in Linearfaktoren zerfällt,

$$p = (-1)^n (X - c_1)^{r_1} \cdots (X - c_k)^{r_k},$$

mit paarweise verschiedenen $c_1, \ldots, c_k \in \mathbb{K}$, ergibt sich somit ein effektives Verfahren zur Bestimmung des Minimalpolynoms m, das ja dann die Form

$$m = (X - c_1)^{s_1} \cdots (X - c_k)^{s_k}$$

hat mit $1 \leq s_i \leq r_i$ für $i = 1, \ldots, k$.

Im Spezialfall

$$p = (-1)^n (X - c)^n,$$

wenn also $H_c = V$ bzw. $H_c = \mathbb{K}^n$ gilt (und nur dann!), ist der Index s die kleinste natürliche Zahl, für die $(\Phi - c \operatorname{id}_V)^s$ die Nullabbildung bzw. $(A - c E_n)^s$ die Nullmatrix ist.

Wir wenden dieses Verfahren in den folgenden Beispielen an.

Beispiele. **(a)** Gegeben sei die Matrix

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

mit dem charakteristischen Polynom $p = (X - 1)(X + 3)^3$ und den Eigenwerten $c_1 = 1$ und $c_2 = -3$. Wir wollen die zugehörigen Haupträume angeben. Es ist

$$m = (X - 1)(X + 3)^s$$

mit $1 \leq s \leq 3$. Wir bestimmen s:

$$\operatorname{Rang}(A + 3 E_4) = \operatorname{Rang} \begin{bmatrix} -1 & 1 & 0 & 1 \\ -1 & 1 & 0 & 1 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = 2,$$

$$\operatorname{Rang}(A + 3 E_4)^2 = \operatorname{Rang} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 16 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = 1,$$

$$\text{Rang}\,(A + 3\,E_4)^3 = \text{Rang}\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 64 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = 1\,.$$

Also ist $s = 2$ und $m = (X-1)\,(X+3)^2$. Die zugehörigen Haupträume sind

$$H_{c_1} = E_{c_1} = [\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}]\,,\; H_{c_2} = [\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix},\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix},\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}] \supset E_{c_2} = [\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix},\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}]\,.$$

(b) Die Matrix

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

besitzt das charakteristische Polynom $p = (X - 2)^4$. Wir wollen das Minimalpolynom m und den Hauptraum H_2 bestimmen. Es ist $m = (X-2)^s$ mit $1 \leq s \leq 4$. Nun berechnen wir den Index s:

$$\text{Rang}\,(A - 2\,E_4) = \text{Rang}\begin{bmatrix} -1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 1 \\ -1 & 1 & 0 & 1 \end{bmatrix} = 1\,,$$

$$\text{Rang}\,(A - 2\,E_4)^2 = \text{Rang}\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = 0\,.$$

Damit ist auch Rang $(A - 2\,E_4)^3 = 0$, also $s = 2$ und $m = (X-2)^2$. Der Hauptraum ist hier trivialerweise der $\mathbb{R}^4$.

Wir kehren jetzt wieder zurück zur Zerlegung

$$p = (X-c)^r \cdot \overline{p}\,, \qquad \overline{p}\,(c) \neq 0\,,$$

mit der zugehörigen Darstellung $V = H_c \oplus \text{Kern}\,\overline{p}\,(\Phi)$ und der zu dieser Zerlegung gehörenden Abbildungsmatrix

$$A_\Phi = \begin{bmatrix} \boxed{A_c} & O \\ O & \boxed{\bar{A}} \end{bmatrix}$$

Unser nächstes Ziel ist, im Hauptraum H_c eine geeignete Basis zu finden, bezüglich der die Abbildungsmatrix A_c von $\Phi|_{H_c}$ eine besonders einfache Form hat. Hierzu betrachten wir die aufsteigende Kette

$$\{o\} \subsetneq E_c = \operatorname{Kern}(\Phi - c\,\mathrm{id}_V) \subsetneq \cdots \subsetneq \operatorname{Kern}(\Phi - c\,\mathrm{id}_V)^s = H_c .$$

Zur Abkürzung setzen wir $U_j := \operatorname{Kern}(\Phi - c\,\mathrm{id}_V)^j$ für $j = 1,\dots,s$ und $U_0 := \{o\}$.

Wir stellen H_c nun als direkte Summe von s Untervektorräumen dar:

$$\begin{array}{ll}
H_c = U_{s-1} \oplus W_1 & \dim W_1 =: q_1 \geq 1 \\
\quad U_{s-1} = U_{s-2} \oplus W_2 & \dim W_2 =: q_2 \geq 1 \\
\qquad \ddots & \quad \vdots \\
\qquad\quad U_2 = E_c \oplus W_{s-1} & \dim W_{s-1} =: q_{s-1} \geq 1 \\
 & \dim E_c =: q_s =: q
\end{array}$$

Dann gilt

$$H_c = W_1 \oplus W_2 \oplus \cdots \oplus W_{s-1} \oplus E_c .$$

Die Untervektorräume W_i werden dabei nach dem folgenden Prinzip bestimmt:

Der erste Ergänzungsraum W_1 kann noch beliebig gewählt werden. Es sei $B_1 = (x_1^{(1)},\dots,x_{q_1}^{(1)})$ eine geordnete Basis von W_1. Wir wenden auf diese q_1 Vektoren die Abbildung $(\Phi - c\,\mathrm{id}_V)$ an.

Behauptung. Für $x_i^{(2)} := (\Phi - c\,\mathrm{id}_V)(x_i^{(1)})$, $i = 1,\dots,q_1$, gilt dann $x_i^{(2)} \in U_{s-1}$, $[x_1^{(2)},\dots,x_{q_1}^{(2)}] \cap U_{s-2} = \{o\}$, und die Vektoren $x_1^{(2)},\dots,x_{q_1}^{(2)}$ sind linear unabhängig.

Beweis. Es gilt $(\Phi - c\,\mathrm{id}_V)^{s-1}\,(x_i^{(2)}) = (\Phi - c\,\mathrm{id}_V)^s\,(x_i^{(1)}) = o$, also $x_i^{(2)} \in U_{s-1}$. Ist $v = a_1\,x_1^{(2)} + \cdots + a_{q_1}\,x_{q_1}^{(2)} \in U_{s-2}$, so ist $(\Phi - c\,\mathrm{id}_V)^{s-1}\,(a_1\,x_1^{(1)} + \cdots + a_{q_1}\,x_{q_1}^{(1)}) = (\Phi - c\,\mathrm{id}_V)^{s-2}\,(v) = o$, also $a_1\,x_1^{(1)} + \cdots + a_{q_1}\,x_{q_1}^{(1)} \in U_{s-1} \cap W_1$. Daraus folgt $a_1\,x_1^{(1)} + \cdots + a_{q_1}\,x_{q_1}^{(1)} = o$ und somit auch $v = o$.

Sei nun $a_1\,x_1^{(2)} + \cdots + a_{q_1}\,x_{q_1}^{(2)} = o$. Dann ist $(\Phi - c\,\mathrm{id}_V)\,(a_1\,x_1^{(1)} + \cdots + a_{q_1}\,x_{q_1}^{(1)}) = o$ und $a_1\,x_1^{(1)} + \cdots + a_{q_1}\,x_{q_1}^{(1)} \in U_{s-1} \cap W_1$. Daraus folgt $a_1\,x_1^{(1)} + \cdots + a_{q_1}\,x_{q_1}^{(1)} = o$ und daher $a_1 = \cdots = a_{q_1} = 0$. Also sind die Vektoren $x_1^{(2)},\ldots,x_{q_1}^{(2)}$ linear unabhängig.■

Der Komplementärraum W_2 wird nun so gewählt, daß er die linear unabhängigen Vektoren $x_1^{(2)},\ldots,x_{q_1}^{(2)}$ enthält. Diese werden zu einer Basis

$$B_2 = (x_1^{(2)},\ldots,x_{q_1}^{(2)}\,,\,x_{q_1+1}^{(2)},\ldots,x_{q_2}^{(2)})$$

von W_2 ergänzt. Auch hier gilt analog, daß die Vektoren

$$x_i^{(3)} := (\Phi - c\,\mathrm{id}_V)\,(x_i^{(2)})\ , \qquad i = 1,\ldots,q_2\ ,$$

in U_{s-2} liegen, linear unabhängig sind, und daß $[x_1^{(3)},\ldots,x_{q_2}^{(3)}] \cap U_{s-3} = \{o\}$ gilt. Wir wählen W_3 derart, daß W_3 diese Vektoren enthält und ergänzen sie zu einer Basis B_3 von W_3

$$B_3 = (x_1^{(3)},\ldots,x_{q_1}^{(3)},\ldots,x_{q_2}^{(3)},\ldots,x_{q_3}^{(3)}).$$

Fahren wir so fort, so erhalten wir schließlich eine Basis von W_{s-1}:

$$B_{s-1} = (x_1^{(s-1)},\ldots,x_{q_1}^{(s-1)},\ldots,x_{q_2}^{(s-1)},\ldots,\ldots,x_{q_{s-1}}^{(s-1)}).$$

Ergänzen wir jetzt noch die Vektoren

$$x_i^{(s)} := (\Phi - c\,\mathrm{id}_V)\,(x_i^{(s-1)})\ , \qquad i = 1,\ldots,q_{s-1}\ ,$$

die ja nach den obigen Überlegungen in $U_1 = E_c$ liegen, zu einer Basis B_s von E_c,

$$B_s = (x_1^{(s)},\ldots,x_{q_1}^{(s)},\ldots,x_{q_2}^{(s)},\ldots,\ldots,x_q^{(s)})\ ,$$

so sind wir fertig.

Wir schreiben die Basis $B_1 \cup \ldots \cup B_s$ des Hauptraums H_c als Schema hin und ordnen dann um:

$$\begin{array}{ll}
B_1: & x_1^{(1)},\ldots,x_{q1}^{(1)} \\[1ex]
B_2: & x_1^{(2)},\ldots,x_{q1}^{(2)},\ x_{q1+1}^{(2)},\ldots,x_{q2}^{(2)} \\[1ex]
B_3: & x_1^{(3)},\ldots,x_{q1}^{(3)},\ x_{q1+1}^{(3)},\ldots,x_{q2}^{(3)},\ x_{q2+1}^{(3)},\ldots,x_{q3}^{(3)} \\
& \quad\vdots \\
B_s: & x_1^{(s)},\ldots,x_{q1}^{(s)},\ x_{q1+1}^{(s)},\ldots,x_{q2}^{(s)},\ x_{q2+1}^{(s)},\ldots,x_{q3}^{(s)},\ldots,\ldots,\ x_{q_{s-1}+1}^{(s)},\ldots,x_q^{(s)}
\end{array}$$

Die geordnete Basis sei nun

$$B = (x_1^{(1)},\ldots,x_1^{(s)},x_2^{(1)},\ldots,x_2^{(s)},\ldots,x_{q1}^{(1)},\ldots,x_{q1}^{(s)},x_{q1+1}^{(2)},\ldots,x_{q1+1}^{(s)},\ldots,x_{q2}^{(2)},\ldots,x_{q2}^{(s)},x_{q2+1}^{(3)},\ldots)\,.$$

Hierbei gilt für $j = 1,\ldots,s-1$ und entsprechende i

$$(\Phi - c\,\mathrm{id}_V)\,(x_i^{(j)}) = x_i^{(j+1)}\,,$$

also

$$\Phi(x_i^{(j)}) = c\,x_i^{(j)} + x_i^{(j+1)}\,,$$

und für $j = s$, $i = 1,\ldots,q$ gilt

$$\Phi(x_i^{(s)}) = c\,x_i^{(s)}\,.$$

Die ersten s Spalten der Abbildungsmatrix A_c bezüglich B haben demnach die Gestalt

$$\begin{array}{l}
\left.\begin{array}{cccccc}
c & 0 & \cdots & \cdots & \cdots & 0 \\
1 & c & \ddots & & & \vdots \\
0 & 1 & c & \ddots & & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\
\vdots & & \ddots & \ddots & c & 0 \\
0 & \cdots & \cdots & 0 & 1 & c
\end{array}\right\} s \\
\begin{array}{cccccc}
0 & \cdots & \cdots & \cdots & \cdots & 0 \\
\vdots & & & & & \vdots \\
0 & \cdots & \cdots & \cdots & \cdots & 0
\end{array}
\end{array}$$

Insgesamt hat A_c die Form

$$A_c = \begin{bmatrix} \boxed{A_1} & & & O \\ & \boxed{A_2} & & \\ & & \ddots & \\ O & & & \boxed{A_q} \end{bmatrix}$$

wobei die A_i wiederum von der Form

$$A_i = \begin{bmatrix} c & 0 & \cdots & \cdots & \cdots & 0 \\ 1 & c & \ddots & & & \vdots \\ 0 & 1 & c & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & c & 0 \\ 0 & \cdots & \cdots & 0 & 1 & c \end{bmatrix}$$

sind.

Wir nennen jedes solche A_i ein *Jordan–Kästchen* und die gesamte Matrix A_c den *Jordan–Block* zum Eigenwert c.

Ist dim $H_c = t$, so hat der Jordan–Block A_c genau t Zeilen. Weil das charakteristische Polynom von A_c offensichtlich $(c - X)^t$ ist, gilt $p = (X - c)^r \cdot \overline{p} = (c - X)^t \cdot \tilde{p}$, wobei $\tilde{p}$ das charakteristische Polynom von $\overline{A}$ ist. Es ist $t \leq r$. Wäre $t < r$, so wäre c Nullstelle von $\tilde{p}$ und es gäbe somit zum Eigenwert c einen Eigenvektor $v \in$ Kern $\overline{p}(\Phi)$, im Widerspruch zu $H_c \cap$ Kern $\overline{p}(\Phi) = \{o\}$. Also gilt $t = r$ und $(-1)^r \overline{p} = \tilde{p}$. Somit ist dim $H_c = r$ und $(-1)^r \overline{p}$ ist das charakteristische Polynom von $\Phi|_{\text{Kern } \overline{p}(\Phi)}$.

Die Jordan–Kästchen A_i haben eine Zeilenzahl (Länge) zwischen 1 und dem Index s. Dabei treten

q_1	Kästchen der Länge s	(mindestens eins !)
$q_2 - q_1$	Kästchen der Länge $s - 1$	
$q_3 - q_2$	Kästchen der Länge $s - 2$	
$\vdots$	$\vdots$	
$q - q_{s-1}$	Kästchen der Länge 1	

auf.

Insgesamt gibt es im Jordan-Block A_c genau $q = \dim E_c$ Jordan-Kästchen A_i. Die Anzahl der Jordan-Kästchen, welche die Länge l besitzen, ist

$$q_{s-l+1} - q_{s-l} = \dim W_{s-l+1} - \dim W_{s-l}$$

$$= (\dim U_{s-(s-l+1)+1} - \dim U_{s-(s-l+1)}) - (\dim U_{s-(s-l)+1} - \dim U_{s-(s-l)})$$

$$= \dim U_l - \dim U_{l-1} - \dim U_{l+1} + \dim U_l$$

$$= 2 \dim \operatorname{Kern} (\Phi - c_i \operatorname{id}_V)^l - \dim \operatorname{Kern} (\Phi - c_i \operatorname{id}_V)^{l+1} - \dim \operatorname{Kern} (\Phi - c_i \operatorname{id}_V)^{l-1}.$$

$(l = 1,\ldots,s)$

Bemerkung. Durch Änderung der Reihenfolge in der Basis B kann man erreichen, daß die Einsen, die ja nur bei Jordan-Kästchen der Länge $l > 1$ auftreten, oberhalb der Diagonalen stehen, statt unterhalb.

Wir haben jetzt für den Hauptraum H_c eine Basis gefunden, für die die zugehörige Matrix A_c eine einfache Form annimmt. Wenn das charakteristische Polynom p von Φ in Linearfaktoren zerfällt,

$$p = (-1)^n (X - c_1)^{r_1} \cdots (X - c_k)^{r_k},$$

mit paarweise verschiedenen $c_1,\ldots,c_k \in \mathbb{K}$, also $V = H_{c_1} \oplus \cdots \oplus H_{c_k}$ gilt, können wir die Jordan-Blöcke $A_{c_1},\ldots,A_{c_k}$ zur Jordanschen Normalform A_Φ von Φ zusammensetzen.

Satz 21 (Jordansche Normalform). *Es seien V ein n-dimensionaler $\mathbb{K}$-Vektorraum und $\Phi : V \longrightarrow V$ eine lineare Abbildung mit dem charakteristischen Polynom*

$$p = (-1)^n (X - c_1)^{r_1} \cdots (X - c_k)^{r_k},$$

wo $c_1,\ldots,c_k \in \mathbb{K}$ paarweise verschieden sind, und dem Minimalpolynom

$$m = (X - c_1)^{s_1} \cdots (X - c_k)^{s_k}.$$

Dann gibt es eine geordnete Basis B von V, bezüglich der die Abbildungsmatrix A_Φ die Form

$$A_\Phi = \begin{bmatrix} \boxed{A_{c_1}} & & \\ & \ddots & \\ & & \boxed{A_{c_k}} \end{bmatrix}$$

besitzt mit Jordan–Blöcken A_{c_i} zu den Eigenwerten c_i,

$$A_{c_i} = \begin{bmatrix} \boxed{\begin{matrix} c_i & & \\ 1 & \ddots & \\ & \ddots & \ddots \\ & & 1 \quad c_i \end{matrix}} & & \\ & \ddots & \\ & & \boxed{c_i} \end{bmatrix}$$

Der Jordan–Block A_{c_i} hat die Länge r_i, $i = 1,\ldots,k$. Innerhalb des Jordan–Blockes A_{c_i} zum Eigenwert c_i gibt es

$$2 \dim \operatorname{Kern} (\Phi - c_i \operatorname{id}_V)^l - \dim \operatorname{Kern} (\Phi - c_i \operatorname{id}_V)^{l+1} - \dim \operatorname{Kern} (\Phi - c_i \operatorname{id}_V)^{l-1}$$

Jordan–Kästchen der Länge l, $l = 1,\ldots,s_i$.

Im Jordan–Block A_{c_i} treten insgesamt $\dim E_{c_i}$ *Jordan–Kästchen auf. Es gibt mindestens ein Kästchen der Maximallänge s_i.*

Bemerkungen und Bezeichnungen. **(a)** A_Φ heißt die *Jordansche Normalform* von Φ, und B heißt eine *Jordan–Basis* von Φ. Die Jordansche Normalform ist, bis auf die Reihenfolge der Kästchen, eindeutig.

(b) Ist $A \in \mathbb{K}^{n\times n}$ und zerfällt das charakteristische Polynom p von A in Linearfaktoren, $p = (-1)^n (X - c_1)^{r_1} \cdots (X - c_k)^{r_k}$, $c_1,\ldots,c_k \in \mathbb{K}$ paarweise verschieden, so gilt für die zugehörige lineare Abbildung $\Phi : \mathbb{K}^n \longrightarrow \mathbb{K}^n$, $\Phi(x) = A\,x$, der eben bewiesene Satz 21.

A ist dann ähnlich zu der Jordanschen Normalform von Φ. Diese Matrix $\tilde{A}$ nennen wir die *Jordansche Normalform* der Matrix A.

Existieren die Jordanschen Normalformen der Matrizen $A,B \in \mathbb{K}^{n\times n}$, so sind diese Normalformen genau dann gleich, wenn A und B ähnlich sind.

(c) Die Forderung, daß das charakteristische Polynom bzw. das Minimalpolynom in Linearfaktoren zerfällt, kann entfallen, wenn $\mathbb{K} = \mathbb{C}$ ist.

Beispiele. **(a)** Der Endomorphismus $\Phi : V \longrightarrow V$ sei bezüglich der Basis (v_1,v_2,v_3) durch

$$\begin{aligned} \Phi(v_1) &= 3\,v_1 - v_2 + \,v_3 \\ \Phi(v_2) &= 4\,v_1 \phantom{{}-v_2} + 2\,v_3 \\ \Phi(v_3) &= 3\,v_1 - v_2 + 3\,v_3 \end{aligned}$$

gegeben. Die zugehörige Abbildungsmatrix A hat dann die Form

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

Damit folgt $p = -(X-2)^3$ sowie $m = (X-2)^s$ mit $s \leq 3$. Nun gilt:

$$\text{Rang}\,(A - 2\,E_3) = \text{Rang} \begin{bmatrix} 1 & 4 & 3 \\ -1 & -2 & -1 \\ 1 & 2 & 1 \end{bmatrix} = 2\,.$$

Also ist $\dim E_c = 1$ und es gibt nur ein Jordan–Kästchen. Dieses muß dann zwangsläufig die Länge 3 besitzen, woraus für den Index $s = 3$ folgt. Die Jordansche Normalform ist also

$$A_\Phi = \begin{bmatrix} \boxed{\begin{matrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{matrix}} \end{bmatrix}$$

Wir wollen nun auch eine Basis B bestimmen, bezüglich der diese Form angenommen wird. Es gilt

$$\begin{aligned} V = H_c &= \operatorname{Kern}\,(\Phi - 2\,\mathrm{id}_V)^3\,, \\ U_2 &= \operatorname{Kern}\,(\Phi - 2\,\mathrm{id}_V)^2\,, \\ U_1 &= \operatorname{Kern}\,(\Phi - 2\,\mathrm{id}_V) = E_c\,. \end{aligned}$$

Zur Bestimmung von U_2 lösen wir das LGS $(A - 2\,E_3)^2\,\hat{v} = o$:

$$\begin{bmatrix} 0 & 2 & 2 \\ 0 & -2 & -2 \\ 0 & 2 & 2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Es ergibt sich $U_2 = \{v \in V \mid \hat{v} = (a_1, a_2, -a_2)\,,\ a_1, a_2 \in \mathbb{K}\} = [v_1, v_2 - v_3]$.

Zur Bestimmung von U_1 lösen wir das LGS $(A - 2\,E_3)\,\hat{v} = o$:

$$\begin{bmatrix} 1 & 4 & 3 \\ -1 & -2 & -1 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \iff \begin{bmatrix} 1 & 4 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Es ergibt sich $U_1 = \{v \in V \mid \hat{v} = (a_3, -a_3, a_3)\,,\ a_3 \in \mathbb{K}\} = [v_1 - v_2 + v_3]$.

Nun wählen wir $x_1 \in H_c \setminus U_2$ beliebig, etwa $x_1 = v_2$; dann ist

$$x_2 := (\Phi - 2\,\mathrm{id}_V)\,x_1 \quad \text{und } x_2 \in U_2 \setminus U_1\,,$$

$$x_3 := (\Phi - 2\,\mathrm{id}_V)\,x_2 = (\Phi - 2\,\mathrm{id}_V)^2\,x_1 \quad \text{und} \quad x_3 \in U_1 = E_c\,.$$

Die zugehörigen Koordinatendarstellungen sind dann

$$\hat{x}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \hat{x}_2 = \begin{bmatrix} 1 & 4 & 3 \\ -1 & -2 & -1 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 4 \\ -2 \\ 2 \end{bmatrix}, \quad \hat{x}_3 = \begin{bmatrix} 0 & 2 & 2 \\ 0 & -2 & -2 \\ 0 & 2 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ 2 \end{bmatrix}.$$

Die gesuchte Jordan-Basis ist also $B = (x_1, x_2, x_3)$ mit

$$\begin{aligned} x_1 &= v_2 \\ x_2 &= 4\,v_1 - 2\,v_2 + 2\,v_3 \\ x_3 &= 2\,v_1 - 2\,v_2 + 2\,v_3 \end{aligned}$$

(b) (vgl. S.199) Wir betrachten die reelle Matrix

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

Es ist

$$p = (X-1)(X+3)^3 \,, \text{ also } c_1 = 1, r_1 = 1 \text{ und } c_2 = -3, r_2 = 3 \,,$$

$$m = (X-1)\,(X+3)^2 \,, \text{ also } s_1 = 1 \text{ und } s_2 = 2 \,.$$

Es gibt daher einen Jordan–Block der Länge $r_1 = 1$ zum Eigenwert $c_1 = 1$ und einen solchen der Länge $r_2 = 3$ zum Eigenwert $c_2 = -3$. In diesem Block existiert mindestens ein Jordan–Kästchen der maximalen Länge $s_2 = 2$. Also lautet die Jordansche Normalform von A

$$\tilde{A} = \begin{bmatrix} -3 & 0 & 0 & 0 \\ 1 & -3 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Wir wollen eine Transformationsmatrix S bestimmen, für die $S^{-1} A\, S = \tilde{A}$ gilt. Dazu bestimmen wir eine Jordan–Basis der zugehörigen linearen Abbildung Φ :

Für den Hauptraum zum Eigenwert $c_1 = 1$ erhalten wir

$$H_{c_1} = \{v \in \mathbb{R}^4 \mid (A - E_4)\,v = o\} = [\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}] \,.$$

Für den Hauptraum zum Eigenwert $c_2 = -3$ erhalten wir

$$H_{c_2} = \{v \in \mathbb{R}^4 \mid (A + 3\,E_4)^2\, v = o\} = U_2 = [\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}]$$

und für den Eigenraum E_{c_2}

$$E_{c_2} = \{v \in \mathbb{R}^4 \mid (A + 3\,E_4)\, v = o\} = U_1 = [\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}] .$$

Dabei sind

$$(A + 3\,E_4) = \begin{bmatrix} -1 & 1 & 0 & 1 \\ -1 & 1 & 0 & 1 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{und} \quad (A + 3\,E_4)^2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 16 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Wir wählen $x_1 \in U_2 \setminus U_1$ beliebig; der Basisvektor x_2 ist dann festgelegt:

$$x_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad , \quad x_2 = (A + 3\,E_4)\, x_1 = \begin{bmatrix} -1 \\ -1 \\ 0 \\ 0 \end{bmatrix} \in U_1 .$$

Nun ergänzen wir x_2 mit dem Vektor

$$x_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \in U_1$$

zu einer Basis von U_1. Bezüglich x_1, x_2, x_3 ergibt sich dann der Jordan-Block zum Eigenwert $c_2 = -3$. Insgesamt erhalten wir als Transformationsmatrix

$$S = \begin{bmatrix} 1 & -1 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Kontrolle: $S^{-1}A\,S = \tilde{A} \iff A\,S = S\,\tilde{A}$,

$$A\,S = \begin{bmatrix} -4 & 1 & 0 & 1 \\ -1 & -2 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} -4 & 3 & -3 & 0 \\ -1 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -3 & 0 \end{bmatrix}$$

$$S\,\tilde{A} = \begin{bmatrix} 1 & -1 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} -3 & 0 & 0 & 0 \\ 1 & -3 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -4 & 3 & -3 & 0 \\ -1 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -3 & 0 \end{bmatrix}$$

(c) (vgl. S.183) Sei

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

Hier ist

$$p = (X+1)^2\,(X-1)\,(X-2)\;,\quad c_1 = -1,\; c_2 = 1,\; c_3 = 2\;,$$

und $1 \le s_1 \le r_1 = 2$, $r_2 = s_2 = 1$, $r_3 = s_3 = 1$. Weil A nicht diagonalisierbar ist, muß $s_1 > 1$ gelten. Also ist $m = p$.

Die Haupträume H_{c_2} und H_{c_3} sind also eindimensional, der Hauptraum H_{c_1} hat die Dimension 2. Im zugehörigen Jordan-Block A_{c_1} gibt es ein Kästchen der maximalen Länge 2.

Damit ergibt sich als Jordansche Normalform für A die Matrix

$$\tilde{A} = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

Wir erhalten wieder eine Transformationsmatrix S mit $S^{-1}A\,S = \tilde{A}$, indem wir in den einzelnen Haupträumen jeweils Jordan-Basen bestimmen:

Für den Hauptraum H_{c_1} ergibt sich

$$H_{c_1} = \{v \in \mathbb{R}^4 \mid (A + E_4)^2 v = o\} = [\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}],$$

denn es gilt

$$\begin{bmatrix} 5 & -5 & 5 & -1 \\ -4 & 4 & -4 & 8 \\ 5 & -5 & 5 & -1 \\ 5 & -5 & 5 & -1 \end{bmatrix} v = o \iff v = \begin{bmatrix} a_1 \\ a_1 + a_3 \\ a_3 \\ 0 \end{bmatrix}, \quad a_1, a_3 \in \mathbb{K}.$$

Für den Untervektorraum $U_1 \subset H_{c_1}$ erhalten wir

$$U_1 = \{v \in \mathbb{R}^4 \mid (A + E_4) v = o\} = [\begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}],$$

denn es gilt

$$\begin{bmatrix} 1 & -1 & 1 & 1 \\ -1 & 2 & -2 & 3 \\ 2 & -1 & 1 & 0 \\ 1 & -1 & 1 & 1 \end{bmatrix} v = o \iff v = \begin{bmatrix} 0 \\ a \\ a \\ 0 \end{bmatrix}, \quad a \in \mathbb{K}.$$

Wir wählen $x_1 \in H_{c_1} \setminus U_1$ beliebig; dann ist x_2 festgelegt:

$$x_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad x_2 = (A + E_4)\, x_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}.$$

Für die Haupträume H_{c_2} und H_{c_3} ergibt sich

$$H_{c_2} = E_{c_2} = [\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}], \quad H_{c_3} = E_{c_3} = [\begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}].$$

Damit erhalten wir als Transformationsmatrix

$$S = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

Kontrolle:

$$A\,S = \begin{bmatrix} 0 & -1 & 1 & 1 \\ -1 & 1 & -2 & 3 \\ 2 & -1 & 0 & 0 \\ 1 & -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 1 & 2 \\ 0 & -1 & 1 & 0 \\ 1 & -1 & 1 & 2 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

$$S\,\tilde{A} = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 1 & 2 \\ 0 & -1 & 1 & 0 \\ 1 & -1 & 1 & 2 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

(d) (abstrakt) Es sei $A \in \mathbb{R}^{12\times 12}$ mit $p = (2 - X)^{12}$, d.h. $c = 2$ ist 12-facher Eigenwert und der Hauptraum H_c ist der Gesamtraum $\mathbb{R}^{12}$. Weiterhin gelte

$$\text{Rang}\,(A - 2\,E_{12}) = 7 \;,\; \text{Rang}\,(A - 2\,E_{12})^2 = 4 \;,$$

$$\text{Rang}\,(A - 2\,E_{12})^3 = 2 \;,\; \text{Rang}\,(A - 2\,E_{12})^4 = 0.$$

Damit ist $(A - 2\,E_{12})^4 = O$, also $s = 4$ und $m = (X - 2)^4$. Für die Untervektorräume U_i gilt:

$$E_c = U_1 \subsetneq U_2 \subsetneq U_3 \subsetneq U_4 = H_c \,.$$

Die zugehörigen Dimensionen sind:

$$\dim U_1 = 12 - 7 = 5 \;,\; \dim U_2 = 12 - 4 = 8 \;,\; \dim U_3 = 12 - 2 = 10 \;,\; \dim U_4 = 12.$$

Wegen $\dim E_c = 5$ gibt es 5 Jordan-Kästchen und zwar

$2 \dim U_4 - \dim U_5 - \dim U_3 = 24 - 12 - 10 = 2$ Kästchen der Länge 4 ,

$2 \dim U_3 - \dim U_4 - \dim U_2 = 20 - 12 - 8 = 0$ Kästchen der Länge 3 ,

$2 \dim U_2 - \dim U_3 - \dim U_1 = 16 - 10 - 5 = 1$ Kästchen der Länge 2 ,

$2 \dim U_1 - \dim U_2 - \dim U_0 = 10 - 8 - 0 = 2$ Kästchen der Länge 1 .

Damit ergibt sich folgende Jordansche Normalform von A (die freien Stellen sind mit Nullen aufzufüllen),

$$\tilde{A} = \begin{bmatrix} 2&0&0&0&&&&&&&& \\ 1&2&0&0&&&&&&&& \\ 0&1&2&0&&&&&&&& \\ 0&0&1&2&&&&&&&& \\ &&&&2&0&0&0&&&& \\ &&&&1&2&0&0&&&& \\ &&&&0&1&2&0&&&& \\ &&&&0&0&1&2&&&& \\ &&&&&&&&2&0&& \\ &&&&&&&&1&2&& \\ &&&&&&&&&&2& \\ &&&&&&&&&&&2 \end{bmatrix}$$

Zum Schluß dieses Paragraphen wollen wir noch den Fall ansprechen, daß das charakteristische Polynom p **nicht** in Linearfaktoren zerfällt, also die Form

$$p = (X - c_1)^{r_1} \cdots (X - c_k)^{r_k} \cdot \overline{p}$$

hat, wobei $\overline{p} \in \mathbb{K}[X]$ keine Nullstelle besitzt und Grad $\overline{p} > 1$ ist. Hier ergibt sich die Normalform

$$A_\Phi = \begin{bmatrix} \boxed{A_{c_1}} &&& \\ & \ddots && \\ && \boxed{A_{c_k}} & \\ &&& \boxed{\overline{A}} \end{bmatrix}$$

die der Zerlegung $V = H_{c_1} \oplus \cdots \oplus H_{c_k} \oplus \overline{V}$ entspricht. Neben den Jordan–Blöcken A_{c_i}, $i = 1,\ldots,k$, tritt also noch der Block $\overline{A}$ auf, der zu dem invarianten Unterraum $\overline{V} = \operatorname{Kern} \overline{p}\,(\Phi)$ gehört. Es gilt

$$\overline{V} = \bigcap_{i=1}^{k} \operatorname{Bild}\, (\Phi - c_i\, \mathrm{id}_V)^{r_i}$$

(Übungsaufgabe).

Im Fall $\mathbb{K} = \mathbb{R}$ können wir auch für $\overline{A}$ eine einfachere Form angeben. Diese *reelle Normalform* wird im nächsten Abschnitt hergeleitet werden.

§ 5 Reelle Jordansche Normalform

Zur Herleitung der reellen Jordanschen Normalform benutzen wir folgende Bemerkung über die Zerlegung reeller Polynome.

Bemerkung. Jedes normierte Polynom $p \in \mathbb{R}[X]$ besitzt eine Zerlegung

$$p = (X - c_1)^{r_1} \cdots (X - c_k)^{r_k} (X^2 + a_1 X + b_1)^{t_1} \cdots (X^2 + a_m X + b_m)^{t_m}$$

mit $c_i, a_j, b_j \in \mathbb{R}$, wobei die Polynome $(X^2 + a_j X + b_j)^{t_j}$, $j = 1,\ldots,m$, keine reelle Nullstelle haben.

Beweis. Das reelle Polynom $p = d_0 + d_1 X + \cdots + d_n X^n$, $d_i \in \mathbb{R}$, $d_n = 1$, läßt sich über $\mathbb{C}$ in Linearfaktoren zerlegen, $p = (X - c_1) \cdot \ldots \cdot (X - c_n)$, $c_i \in \mathbb{C}$. Für das Polynom $\overline{p} := (X - \overline{c}_1) \cdot \ldots \cdot (X - \overline{c}_n)$ gilt dann $\overline{p} = \overline{d}_0 + \overline{d}_1 X + \cdots + \overline{d}_n X^n = p$. Also ist mit jeder komplexen Nullstelle c von p der Vielfachheit t auch die konjugiert komplexe Zahl $\overline{c}$ eine Nullstelle der Vielfachheit t.

Seien nun $c_1,\ldots,c_k$ die reellen Nullstellen von p und $c_1', \overline{c_1'},\ldots,c_m', \overline{c_m'}$ die verschiedenen komplexen Nullstellen. Dann folgt

$$\begin{aligned} p &= (X - c_1)^{r_1} \cdots (X - c_k)^{r_k} (X - c_1')^{t_1} (X - \overline{c_1'})^{t_1} \cdots (X - c_m')^{t_m} (X - \overline{c_m'})^{t_m} \\ &= (X - c_1)^{r_1} \cdots (X - c_k)^{r_k} (X^2 - \underbrace{(c_1' + \overline{c_1'})}_{-a_1} X + \underbrace{c_1' \overline{c_1'}}_{b_1})^{t_1} \\ &\qquad \cdots (X^2 - \underbrace{(c_m' + \overline{c_m'})}_{-a_m} X + \underbrace{c_m' \overline{c_m'}}_{b_m})^{t_m} \end{aligned}$$

■

Wir betrachten jetzt die Matrix $A \in \mathbb{R}^{n \times n}$. Seien p das charakteristische Polynom von A und $c, \overline{c}$ ein Paar konjugiert komplexer Nullstellen von p. Bei der Zerlegung $c = a + i\, b$ in Real- und Imaginärteil können wir o.B.d.A. $b > 0$ voraussetzen. Wegen

$$(A - c\,E_n)^j\, z = o \iff (A - \bar{c}\,E_n)^j\, \bar{z} = o$$

für $z \in \mathbb{C}^n$, ist

$$\text{Rang}\,(A - c\,E_n)^j = \text{Rang}\,(A - \bar{c}\,E_n)^j .$$

Dies bedeutet, daß in der komplexen Jordanschen Normalform von $A = A_\Phi$ die Jordan–Kästchen immer in Paaren gleicher Länge auftreten,

$$\boxed{\begin{matrix} c & 0 & \cdots & 0 \\ 1 & c & \ddots & \vdots \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0\;1 & c \end{matrix}} \qquad \boxed{\begin{matrix} \bar{c} & 0 & \cdots & 0 \\ 1 & \bar{c} & \ddots & \vdots \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0\;1 & \bar{c} \end{matrix}}$$

wobei zum 1. Kästchen die Basis $z_1,\dots,z_q$, und zum 2. Kästchen die Basis $\bar{z}_1,\dots,\bar{z}_q$ gehört, d.h. auch die Basen können konjugiert komplex gewählt werden. Setzen wir für $j = 1,\dots,q$

$$x_j = \tfrac{1}{2}(z_j + \bar{z}_j)\ ,\quad y_j = -\tfrac{i}{2}(z_j - \bar{z}_j)\,,$$

so gilt im Vektorraum $\mathbb{C}^n$

$$[z_1,\dots,z_q,\bar{z}_1,\dots,\bar{z}_q] = [x_1,y_1,\dots,x_q,y_q]\,,$$

also sind x_1, y_1, ..., x_q, y_q linear unabhängig. Nach Konstruktion ist x_j, $y_j \in \mathbb{R}^n$ und für alle $j = 1,\dots,q-1$ gilt:

$$A\,x_j = \tfrac{1}{2}(A\,z_j + A\,\bar{z}_j) = \tfrac{1}{2}(c\,z_j + z_{j+1} + \bar{c}\,\bar{z}_j + \bar{z}_{j+1})$$

$$= a\,\frac{z_j + \bar{z}_j}{2} - y\,\frac{(-i)(z_j - \bar{z}_j)}{2} + \frac{z_{j+1} + \bar{z}_{j+1}}{2} = a\,x_j - b\,y_j + x_{j+1}$$

und

$$A\,y_j = -\tfrac{i}{2}(A\,z_j - A\,\bar{z}_j) = a\,y_j + b\,x_j + y_{j+1}\,.$$

Für $j = q$ erhalten wir

$$A\,x_q = \tfrac{1}{2}(c\,z_q + \bar{c}\,\bar{z}_q) = a\,x_q - b\,y_q$$

und

$$A\,y_q = a\,y_q + b\,x_q\,.$$

Also ergibt sich für diesen Teil der Abbildungsmatrix bezüglich der neuen Basis $x_1, y_1, \ldots, x_q, y_q$

$$\begin{bmatrix} \boxed{\begin{matrix} a & b \\ -b & a \end{matrix}} & & & & \\ \boxed{\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}} & \boxed{\begin{matrix} a & b \\ -b & a \end{matrix}} & & & \\ & \boxed{\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}} & \ddots & & \\ & & & \boxed{\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}} & \boxed{\begin{matrix} a & b \\ -b & a \end{matrix}} \end{bmatrix}$$

Wenden wir das Verfahren auf jedes Paar von Jordan–Kästchen der konjugiert komplexen Eigenwerte von A an, so erhalten wir das gewünschte Ergebnis.

Satz 22 (Reelle Normalform). *Es sei $A \in \mathbb{R}^{n\times n}$ mit charakteristischem Polynom*

$$p = (-1)^n (X - c_1)^{r_1} \cdots (X - c_k)^{r_k} (X^2 + a_1' X + b_1')^{t_1} \cdots (X^2 + a_m' X + b_m')^{t_m}\,,$$

wo $c_1, \ldots, c_k$ die paarweise verschiedenen Eigenwerte von A und $X^2 + a_1' X + b_1', \ldots, X^2 + a_m' X + b_m'$ paarweise verschiedene reelle Polynome ohne Nullstellen sind. Dann gibt es eine zu A ähnliche Matrix $\tilde{A}$ der Form

$$\tilde{A} = \begin{bmatrix} A_{c_1} & & & & & O \\ & \ddots & & & & \\ & & A_{c_k} & & & \\ & & & B_1 & & \\ & O & & & \ddots & \\ & & & & & B_m \end{bmatrix}$$

Dabei sind $A_{c_1},\dots,A_{c_k}$ *die Jordan–Blöcke zu den Eigenwerten* $c_1,\dots,c_k$ *und die Blöcke* B_l *haben die Form*

$$B_l = \begin{bmatrix} B_l^{(1)} & & O \\ & \ddots & \\ O & & B_l^{(n_l)} \end{bmatrix}, \quad l = 1,\dots,m\,.$$

Hierbei hat das Kästchen $B_l^{(j)}$ *die Gestalt*

$$B_l^{(j)} = \begin{bmatrix} \begin{matrix} a_l & b_l \\ -b_l & a_l \end{matrix} & & & \\ \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} & \begin{matrix} a_l & b_l \\ -b_l & a_l \end{matrix} & & \\ & \ddots & \ddots & \\ & & \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} & \begin{matrix} a_l & b_l \\ -b_l & a_l \end{matrix} \end{bmatrix}, \quad j = 1,\dots,n_l\,,$$

wo $a_l \pm i\, b_l$, $b_l > 0$, *die komplexen Nullstellen von* $X^2 + a_l'\,X + b_l'$ *sind.*

Bemerkungen. **(a)** Die Zahlen n_l, $l = 1,\dots,m$, und die Größe der Kästchen $B_l^{(j)}$ ergeben sich aus der komplexen Jordanschen Normalform der Matrix A, wenn man diese als Element aus $\mathbb{C}^{n\times n}$ auffaßt.

n_l ist gerade die Zahl der Jordan–Kästchen zum Eigenwert $a_l + i\,b_l$ und zu jedem Jordan–Kästchen der Länge q zum Eigenwert $a_l + i\,b_l$ gibt es ein Kästchen $B_l^{(j)}$ der Länge $2\,q$.

(b) Die Basis, bezüglich der man die Normalform $\tilde{A}$ von A erhält, läßt sich mit dem oben beschriebenen Verfahren aus den konjugiert komplexen Basisvektoren in der Jordan–Basis der komplexen Matrix A bestimmen.

Beipiel. Gegeben sei die Matrix

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

mit dem charakteristischen Polynom

$$p = (X^2 - 2\,X + 2)^2 = (X - (1+i))^2\,(X - (1-i))^2 .$$

Wir bestimmen zunächst die komplexe Jordansche Normalform A' von A. Die Nullstellen von $X^2 - 2\,X + 2$ sind $c_1 = 1 + i$ und $c_2 = 1 - i$. Der Eigenraum zu c_1 ergibt sich als Lösungsmenge des LGS $(A - (1+i)\,E_4)\,z = o$. Es ist

$$E_{1+i} = [\begin{bmatrix} 1-i \\ 1 \\ -i \\ 1-i \end{bmatrix}] .$$

Der Hauptraum zu c_1 ist Lösungsmenge des LGS $(A - (1+i)\,E_4)^2\,z = o$. Es ist

$$H_{1+i} = [\begin{bmatrix} i \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ i \\ 1 \\ 0 \end{bmatrix}] .$$

Im Hauptraum H_{1+i} bestimmen wir eine Jordan–Basis:

$$z_1 = \begin{bmatrix} i \\ 0 \\ 0 \\ 1 \end{bmatrix} \in H_{1+i} \setminus E_{1+i} \quad \text{und} \quad z_2 = (A - (1+i)\,E_4)\,z_1 = \begin{bmatrix} 1-i \\ 1 \\ -i \\ 1-i \end{bmatrix} .$$

Damit ergibt sich für den Eigenraum und den Hauptraum zur konjugiert komplexen Nullstelle $c_2 = 1 - i$ sowie für die Basis in diesem Hauptraum:

$$E_{1-i} = [\begin{bmatrix} 1+i \\ 1 \\ i \\ 1+i \end{bmatrix}] \quad \text{und} \quad H_{1-i} = [\begin{bmatrix} -i \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -i \\ 1 \\ 0 \end{bmatrix}] ,$$

$$\overline{z}_1 = \begin{bmatrix} -i \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad \text{und} \quad \overline{z}_2 = \begin{bmatrix} 1+i \\ 1 \\ i \\ 1+i \end{bmatrix}.$$

Die komplexe Jordansche Normalform und eine zugehörige Transformationsmatrix S lauten:

$$A' = \begin{bmatrix} 1+i & 0 & 0 & 0 \\ 1 & 1+i & 0 & 0 \\ 0 & 0 & 1-i & 0 \\ 0 & 0 & 1 & 1-i \end{bmatrix}, \quad S = \begin{bmatrix} i & 1-i & -i & 1+i \\ 0 & 1 & 0 & 1 \\ 0 & -i & 0 & i \\ 1 & 1-i & 1 & 1+i \end{bmatrix}$$

Damit ergeben sich die reelle Jordansche Normalform

$$\tilde{A} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & 1 \end{bmatrix}$$

und die reelle Transformationsmatrix

$$S = \begin{bmatrix} 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 1 & -1 \end{bmatrix}$$

Kontrolle: $\tilde{A} = S^{-1} A S \iff S \tilde{A} = A S$:

$$S\tilde{A} = \begin{bmatrix} 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 2 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & -1 & 1 & -1 \\ 2 & 0 & 2 & 0 \end{bmatrix}$$

$$A S = \begin{bmatrix} 0 & 2 & 0 & 0 \\ 0 & 0 & -2 & 1 \\ -1 & 2 & 2 & 0 \\ 0 & 0 & -2 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 2 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & -1 & 1 & -1 \\ 2 & 0 & 2 & 0 \end{bmatrix}$$