Ein Skript für Lineare Algebra I und II

Chris Preston

2003/04

Dies ist ein Skript für die Vorlesungen Lineare Algebra I und II. Die Texte von Jänich [5] und Fischer [3] haben die Darstellung beeinflusst.

Chris Preston
Pankow, im 'Sommer' 2004

Inhaltsverzeichnis

1 Lineare Gleichungssyteme

Sei $n \geq 1$; mit $\mathbb{R}^n$ wird die Menge aller n-Tupel reeller Zahlen bezeichnet. Ein Element von $\mathbb{R}^n$ hat also die Form $(x_1, \ldots, x_n)$ mit $x_1, \ldots, x_n$ Elementen aus $\mathbb{R}$. Ist $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ und $1 \leq k \leq n$, so heißt x_k die k-te *Komponente* von x. Das Element $(0, \ldots, 0) \in \mathbb{R}^n$ wird mit 0 bezeichnet.

Eine Addition auf $\mathbb{R}^n$ wird erklärt durch

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$$

und eine Multiplikation eines Elements von $\mathbb{R}^n$ mit einer reellen Zahl durch

$$x\,(y_1, y_2, \ldots, y_n) = (xy_1, xy_2, \ldots, xy_n)\ .$$

Sei $v \in \mathbb{R}^n$; für jedes $x \in \mathbb{R}$ heißt dann xv ein *Vielfaches* von v.

Seien m, $n \in \mathbb{N}$; eine $m \times n$ *reelle Matrix* ist eine Anordnung von mn Elementen von $\mathbb{R}$ nach folgendem Schema

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Die Menge der $m \times n$ reellen Matrizen wird mit $\mathrm{M}(m \times n, \mathbb{R})$ bezeichnet. Sei

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \in \mathrm{M}(m \times n, \mathbb{R})$$

eine $m \times n$ reelle Matrix; dann schreibt man auch $A = (a_{ij})_{1 \leq i \leq m,\, 1 \leq j \leq n}$ oder nur $A = (a_{ij})$. Seien $1 \leq i \leq m$, $1 \leq j \leq n$; der waagerecht geschriebene n-Tupel $(a_{i1}, \ldots, a_{in})$ wird die i-te *Zeile* von A und der senkrecht geschriebene m-Tupel

$$\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

die j-te *Spalte* von A genannt. Die Zeilen von A werden als Elemente von $\mathbb{R}^n$ und die Spalten von A als Elemente von $\mathbb{R}^m$ betrachtet. Das Element

$$\begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \in \mathrm{M}(m \times n, \mathbb{R})$$

wird mit 0 bezeichnet.

Seien m, $n \geq 1$ und $A = (a_{ij}) \in \mathrm{M}(m \times n, \mathbb{R})$, $b = (b_1, \ldots, b_m) \in \mathbb{R}^m$; dann heißt

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

das *zu A und b gehörige lineare Gleichungssystem.*

Ein Element $(y_1, \ldots, y_n) \in \mathbb{R}^n$ heißt *Lösung* des zu A und b gehörigen linearen Gleichungssystems, wenn y_1, …, y_n die m Gleichungen erfüllen, d.h., wenn

$$\begin{aligned} a_{11}y_1 + a_{12}y_2 + \cdots + a_{1n}y_n &= b_1 \\ a_{21}y_1 + a_{22}y_2 + \cdots + a_{2n}y_n &= b_2 \\ &\vdots \\ a_{m1}y_1 + a_{m2}y_2 + \cdots + a_{mn}y_n &= b_m \end{aligned}$$

Die Menge aller Lösungen des Systems wird mit $\mathrm{Lös}(A, b)$ bezeichnet.

Das Gleichungssystem heißt *lösbar*, wenn $\mathrm{Lös}(A, b)$ mindestens ein Element von $\mathbb{R}^n$ enthält. Es heißt *eindeutig lösbar*, wenn $\mathrm{Lös}(A, b)$ aus genau einem Element von $\mathbb{R}^n$ besteht.

Das zu A und b gehörige lineare Gleichungssystem heißt *homogen*, wenn $b = 0$, d.h., wenn $b_j = 0$ für jedes $j = 1, \ldots, m$. Ein homogenes Gleichungssystem besitzt stets die triviale Lösung $0 = (0, \ldots, 0)$. (Insbesondere ist ein homogenes Gleichungssystem stets lösbar.)

Eine Teilmenge U von $\mathbb{R}^n$ heißt *Untervektorraum* von $\mathbb{R}^n$, wenn gilt:

(U0) $0 \in U$,

(U1) $u + v \in U$ für alle u, $v \in U$,

(U2) $xu \in U$ für alle $x \in \mathbb{R}$, $u \in U$.

Satz 1.1 *Sei $A = (a_{ij}) \in \mathrm{M}(m \times n, \mathbb{R})$; dann ist* $\mathrm{Lös}(A, 0)$ *ein Untervektorraum von $\mathbb{R}^n$. (Die Lösungsmenge eines homogenen linearen Gleichungssytems für n Unbekannte ist ein Untervektorraum von $\mathbb{R}^n$.)*

Beweis (U0): Es wurde schon erwähnt, dass $0 = (0, \ldots, 0) \in \mathrm{Lös}(A, 0)$.

(U1): Seien $u = (x_1, \ldots, x_n)$, $v = (y_1, \ldots, y_n) \in \mathrm{Lös}(A, 0)$; d.h.,

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = 0 \quad \text{und} \quad a_{i1}y_1 + a_{i2}y_2 + \cdots + a_{in}y_n = 0$$

für jedes $i = 1, \ldots, m$. Dann gilt

$$\begin{aligned} a_{i1}(x_1 + y_1) + a_{i2}(x_2 + y_2) + \cdots + a_{in}(x_n + y_n) \\ &= a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n + a_{i1}y_1 + a_{i2}y_2 + \cdots + a_{in}y_n \\ &= 0 + 0 = 0 \end{aligned}$$

für jedes $i = 1, \ldots, m$ und damit ist $u + v \in \text{Lös}(A, 0)$.

(U2): Seien $u = (x_1, \ldots, x_n) \in \text{Lös}(A, 0)$, $x \in \mathbb{R}$; dann gilt

$$a_{i1}(xx_1) + a_{i2}(xx_2) + \cdots + a_{in}(xx_n) = x(a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n) = x0 = 0$$

für jedes $i = 1, \ldots, m$; folglich ist auch $xu \in \text{Lös}(A, 0)$. □

Satz 1.2 *Seien $A \in \text{M}(m \times n, \mathbb{R})$, $b \in \mathbb{R}^m$, $v \in \text{Lös}(A, b)$ und $v' \in \mathbb{R}^n$. Dann gilt $v' \in \text{Lös}(A, b)$ genau, wenn $v' = v + u$ für ein $u \in \text{Lös}(A, 0)$. (Man erhält also alle Lösungen des zu A und b gehörigen Gleichungssystems, indem man zu einer speziellen Lösung dieses Systems alle Lösungen des zu A und 0 gehörigen Gleichungssystems addiert.)*

Beweis Seien $A = (a_{ij})$, $b = (b_1, \ldots, b_m)$ und $v = (y_1, \ldots, y_n) \in \text{Lös}(A, b)$. Sei $u = (x_1, \ldots, x_n) \in \text{Lös}(A, 0)$, d.h.,

$$a_{i1}y_1 + a_{i2}y_2 + \cdots + a_{in}y_n = b_i \quad \text{und} \quad a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = 0$$

für jedes $i = 1, \ldots, m$. Dann gilt

$$\begin{aligned} a_{i1}(y_1 + x_1) + a_{i2}(y_2 + x_2) + \cdots + a_{in}(y_n + x_n) \\ &= a_{i1}y_1 + a_{i2}y_2 + \cdots + a_{in}y_n + a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \\ &= b_i + 0 = b_i \end{aligned}$$

für jedes $i = 1, \ldots, m$ und damit ist $v + u \in \text{Lös}(A, b)$.

Seien nun $v = (y_1, \ldots, y_n)$, $v' = (y'_1, \ldots, y'_n) \in \text{Lös}(A, b)$, und sei $u = (x_1, \ldots, x_n)$, wobei $x_i = y'_i - y_i$ für $i = 1, \ldots, m$. Dann ist $v' = v + u$ und

$$\begin{aligned} a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_2 \\ &= a_{i1}(y'_1 - y_1) + a_{i2}(y'_2 - y_2) + \cdots + a_{in}(y'_n - y_n) \\ &= a_{i1}y'_1 + a_{i2}y'_2 + \cdots + a_{in}y'_n - (a_{i1}y_1 + a_{i2}y_2 + \cdots + a_{in}y_n) \\ &= b_i - b_i = 0 \end{aligned}$$

für jedes $i = 1, \ldots, m$; folglich ist $u \in \text{Lös}(A, 0)$. □

Unter *elementaren Zeilenumformungen* einer Matrix versteht man die folgenden Operationen:

I Addition eines Vielfachen einer Zeile zu einer anderen Zeile.

II Vertauschen zweier Zeilen.

Lemma 1.1 *Wird eine Matrix A durch eine elementare Zeilenumformung zu einer Matrix A' verändert, so gilt $A' = 0$ genau dann, wenn $A = 0$.*

Beweis Übung. □

Seien m, $n \geq 1$ und $A = (a_{ij}) \in M(m \times n, \mathbb{R})$, $b = (b_1, \ldots, b_m) \in \mathbb{R}^m$; dann wird die $m \times (n+1)$ reelle Matrix

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix}$$

mit (A, b) bezeichnet.

Satz 1.3 *Wird (A, b) durch eine elementare Zeilenumformung zu einer Matrix (A', b') verändert, so gilt* $\text{Lös}(A, b) = \text{Lös}(A', b')$.

Beweis Entsteht (A', b') aus (A, b) durch Vertauschen zweier Zeilen, so ist es klar, dass $\text{Lös}(A, b) = \text{Lös}(A', b')$. Nehme also an, dass (A', b') aus (A, b) durch eine Zeilenumformung vom Typ I entsteht. Genauer wird angenommen, dass (A', b') durch Addition des x-fachen der p-ten Zeile zu der q-ten Zeile von (A, b) entsteht, (wobei $x \in \mathbb{R}$ und $p \neq q$). Seien $A' = (a'_{ij})$, $b' = (b'_1, \ldots, b'_m)$; dann gilt

- $a'_{ij} = a_{ij}$ für alle $j = 1, \ldots, n$, $i \neq q$,
- $a'_{qj} = a_{qj} + x a_{pj}$ für alle $j = 1, \ldots, n$,
- $b'_i = b_i$ für $i \neq q$,
- $b'_q = b_q + x b_p$.

Sei $y = (y_1, \ldots, y_n) \in \text{Lös}(A, b)$; d.h.,

$$a_{i1} y_1 + a_{i2} y_2 + \cdots + a_{in} y_n = b_i$$

für jedes $i = 1, \ldots, m$. Dann gilt

$$a'_{i1} y_1 + a'_{i2} y_2 + \cdots + a'_{in} y_n = a_{i1} y_1 + a_{i2} y_2 + \cdots + a_{in} y_n = b_i = b'_i$$

für jedes $i \neq q$ und

$$\begin{aligned} a'_{q1}y_1 + a'_{q2}y_2 + \cdots + a'_{qn}y_n \\ &= (a_{q1} + xa_{p1})y_1 + (a_{q2} + xa_{q2})y_2 + \cdots + (a_{qn} + xa_{pn})y_n \\ &= a_{q1}y_1 + a_{q2}y_2 + \cdots + a_{qn}y_n + x(a_{p1}y_1 + a_{p2}y_2 + \cdots + a_{pn}y_n) \\ &= b_q + xb_p \;=\; b'_q \;, \end{aligned}$$

und damit ist $y \in \text{Lös}(A', b')$. Sei umgekehrt $y = (y_1, \ldots, y_n) \in \text{Lös}(A', b')$; d.h.,

$$a'_{i1}y_1 + a'_{i2}y_2 + \cdots + a'_{in}y_n = b'_i$$

für jedes $i = 1, \ldots, m$. Dann gilt

$$a_{i1}y_1 + a_{i2}y_2 + \cdots + a_{in}y_n = a'_{i1}y_1 + a'_{i2}y_2 + \cdots + a'_{in}y_n = b'_i = b_i$$

für jedes $i \neq q$ und

$$\begin{aligned} a_{q1}y_1 + a_{q2}y_2 + \cdots + a_{qn}y_n \\ &= (a'_{q1} - xa_{p1})y_1 + (a'_{q2} - xa_{q2})y_2 + \cdots + (a'_{qn} - xa_{pn})y_n \\ &= a'_{q1}y_1 + a'_{q2}y_2 + \cdots + a'_{qn}y_n - x(a_{p1}y_1 + a_{p2}y_2 + \cdots + a_{pn}y_n) \\ &= b'_q - xb_p = b_q \;; \end{aligned}$$

folglich ist $y \in \text{Lös}(A, b)$. Dies zeigt also, dass $\text{Lös}(A, b) = \text{Lös}(A', b')$. □

Eine Matrix $A \in \text{M}(m \times n, \mathbb{R})$ hat *Zeilen-Stufen-Form*, wenn für jede Zeile der Matrix folgende zwei Bedingungen erfüllt sind:

— Sind die ersten p Elemente der Zeile Null für ein p mit $p < n$, so sind für alle folgenden Zeilen mindestens die ersten $p + 1$ Elemente Null.

— Sind alle Elemente der Zeile Null, so ist jedes Element von jeder der folgenden Zeilen Null.

(Eine solche Matrix sieht etwa so aus:

$$\begin{pmatrix} 0 & 0 & * & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & * & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & * & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & 0 & * & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & \cdot \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Dabei sind die mit einem Stern angedeuteten Elemente verschieden von Null und die mit einem Punkt angedeuteten Elemente beliebig. Die mit einem Stern angedeuteten Elemente bezeichnen also, soweit vorhanden, in jeder Zeile das erste nicht-verschwindende Element.)

Insbesondere hat die Null-Matrix 0 Zeilen-Stufen-Form.

Sei $A = (a_{ij}) \in \mathrm{M}(m \times n, \mathbb{R})$ mit $A \neq 0$ und für $i = 1, \ldots, m$ sei

$$p_i = \begin{cases} \min\{1 \leq j \leq n : a_{ij} \neq 0\} & \text{falls } (a_{i1}, \ldots, a_{in}) \neq 0 \ , \\ 0 & \text{sonst} \ . \end{cases}$$

Die Matrix A hat also Zeilen-Stufen-Form genau dann, wenn

$$p_1 < p_2 < \cdots < p_r \ ,$$

wobei r der Index der letzten von Null verschiedenen Zeile von A ist. Hat A Zeilen-Stufen-Form, so nennt man $p_1, \ldots, p_r$ die *Treppen-Folge* von A.

Satz 1.4 *Seien $A = (a_{ij}) \in \mathrm{M}(m \times n, \mathbb{R})$, $b = (b_1, \ldots, b_m) \in \mathbb{R}^m$ mit $(A, b) \neq 0$. Nehme an, dass (A, b) Zeilen-Stufen-Form hat, und sei $p_1, \ldots, p_r$ die Treppen-Folge von A. Dann ist das zu A und b gehörige lineare Gleichungssystem lösbar genau, wenn $p_r \neq n + 1$. Ferner ist das Gleichungssystem eindeutig lösbar genau dann, wenn $r = n$ und $p_j = j$ für jedes $j = 1, \ldots, n$.*

Beweis Übung. □

$$\begin{pmatrix} 0 & 0 & * & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & * & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & * & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & 0 & * & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

nicht lösbar

$$\begin{pmatrix}
0 & 0 & * & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & 0 & 0 & * & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & 0 & 0 & 0 & * & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & 0 & 0 & 0 & 0 & * & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & \cdot & \cdot \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}$$

lösbar

$$\begin{pmatrix}
* & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & * & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & * & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & 0 & * & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & 0 & 0 & * & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & 0 & 0 & 0 & * & \cdot & \cdot & \cdot \\
0 & 0 & 0 & 0 & 0 & 0 & * & \cdot & \cdot \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & * & \cdot \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}$$

eindeutig lösbar

Satz 1.5 *Jede Matrix in* $\mathrm{M}(m \times n, \mathbb{R})$ *läßt sich durch eine Folge von elementaren Zeilenumformungen in eine Matrix mit Zeilen-Stufen-Form überführen.*

Beweis Der Beweis besteht in der Angabe des *Gaußschen Algorithmus*, der die gesuchte Folge von elementaren Zeilenumformungen bestimmt. Der Hauptschritt dieses Algorithmus wird nun beschrieben.

Sei $A = (a_{ij}) \in \mathrm{M}(m \times n, \mathbb{R})$; für jedes $j = 1, \ldots, n$ bezeichne mit $b_j(A)$ die $m \times j$ reelle Matrix, die aus den ersten j Spalten von A besteht. Sei

$$q' = \begin{cases} \max\{j : b_j(A) \text{ hat Zeilen-Stufen-Form}\} & \text{falls es ein solches } j \text{ gibt ,} \\ 0 & \text{sonst ,} \end{cases}$$

und definiere $p' \geq 0$ wie folgt:

- $p' = 0$, falls $q' = 0$ oder $q' > 0$ und $b_{q'}(A) = 0$,
- p' sei der Index der letzten von Null verschiedenen Zeile von $b_{q'}(A)$, falls $q' > 0$ und $b_{q'}(A) \neq 0$.

Setze $p = p' + 1$ und $q = q' + 1$. Nehme jetzt an, dass A nicht Zeilen-Stufen-Form hat. Dann ist $q \leq n$; ferner ist $p \leq m$ und es gibt mindestens ein i mit $p \leq i \leq m$ und $a_{iq} \neq 0$ (sonst hätte $b_q(A)$ Zeilen-Stufen-Form). Der Hauptschritt des Gaußschen Algorithmus besteht in der Ausführung der folgenden elementaren Zeilenumformungen:

— Gegebenenfalls die Vertauschung der p-ten und der i-ten Zeilen für ein i mit $p < i \leq m$, um ein von Null verschiedenes Element in die pq-te Stelle der Matrix zu bringen.

— Für $i = p+1, \ldots, m$ die Addition des geeigneten Vielfachen der p-ten Zeile zu der i-ten Zeile, um eine Null in die iq-te Stelle der Matrix zu bringen.

Diese Folge von elementaren Umformungen überführt A in eine Matrix A' mit folgenden Eigenschaften:

— Für $i = 1, \ldots, p'$ ist die i-te Zeile von A' gleich die i-te Zeile von A, d.h., $a'_{ij} = a_{ij}$ für alle $i = 1, \ldots, p'$, $j = 1, \ldots, m$.

— $b_{q'}(A') = b_{q'}(A)$.

— $a'_{pq} \neq 0$.

— $a'_{iq} = 0$ für $i = p + 1, \ldots, m$.

Insbesondere hat dann $b_p(A')$ Zeilen-Stufen-Form und damit ist die Matrix A' "näher" an der Zeilen-Stufen-Form als A.

Nach höchstens n-maliger Wiederholung dieses Hauptschrittes wird jede Matrix in $\mathrm{M}(m \times n, \mathbb{R})$ in eine Matrix mit Zeilen-Stufen-Form überführt. □

Seien m, $n \geq 1$ und $A = (a_{ij}) \in \mathrm{M}(m \times n, \mathbb{R})$, $b = (b_1, \ldots, b_m) \in \mathbb{R}^m$. Es ist klar, dass $\mathrm{Lös}(0, 0) = \mathbb{R}^n$; nehme also an, dass $(A, b) \neq 0$. Mit Hilfe des Gaußschen Algorithmus kann die Matrix (A, b) in eine Matrix (A', b') mit Zeilen-Stufen-Form überführt werden und nach Satz 1.3 gilt dann $\mathrm{Lös}(A, b) = \mathrm{Lös}(A', b')$. Ferner gilt nach Lemma 1.1, dass $(A', b') \neq 0$. Folglich kann Satz 1.4 verwendet werden, um festzustellen, ob das zu A und b gehörige lineare Gleichungssystem lösbar bzw. eindeutig lösbar ist.

2 Körper

Ein 5-Tupel $(K, +, \cdot, 0, 1)$ bestehend aus einer Menge K, einer Verknüpfung

$$\begin{aligned} + : K \times K &\to K \\ (\lambda, \mu) &\mapsto \lambda + \mu \end{aligned}$$

(genannt *Addition*), einer Verknüpfung

$$\begin{aligned} \cdot : K \times K &\to K \\ (\lambda, \mu) &\mapsto \lambda\mu \end{aligned}$$

(genannt *Multiplikation*) und Elementen $0, 1 \in K$ mit $0 \neq 1$ heißt *Körper*, wenn folgendes gilt:

(K1) $(\lambda + \mu) + \nu = \lambda + (\mu + \nu)$ für alle λ, μ, $\nu \in K$.

(K2) $\lambda + \mu = \mu + \lambda$ für alle λ, $\mu \in K$.

(K3) $0 + \lambda = \lambda$ für alle $\lambda \in K$.

(K4) Zu jedem $\lambda \in K$ gibt es ein Element $-\lambda \in K$ mit $(-\lambda) + \lambda = 0$.

(K5) $(\lambda\mu)\nu = \lambda(\mu\nu)$ für alle λ, μ, $\nu \in K$.

(K6) $\lambda\mu = \mu\lambda$ für alle λ, $\mu \in K$.

(K7) Für alle $\lambda \in K$ gilt $1\lambda = \lambda$.

(K8) Zu jedem $\lambda \in K$ mit $\lambda \neq 0$ gibt es ein Element $\lambda^{-1} \in K$ mit $\lambda^{-1}\lambda = 1$.

(K9) $\lambda(\mu + \nu) = \lambda\mu + \lambda\nu$ für alle λ, μ, $\nu \in K$.

Bemerkung: Nach der üblichen Konvention soll die Addition in K weniger stark binden als die Multiplikation. ($\lambda\mu + \lambda\nu$ bedeutet also $(\lambda\mu) + (\lambda\nu)$.)

Die Elemente 0 und 1 heißen das *Nullelement* oder die *Null* bzw. das *Einselement* oder die *Eins*.

Lemma 2.1 *Sei $(K, +, \cdot, 0, 1)$ ein Körper.*

(1) Das Nullelement 0 ist eindeutig: Ist $0' \in K$ ein Element mit $0' + \lambda = \lambda$ für alle $\lambda \in K$, so ist $0' = 0$.

(2) Zu jedem $\lambda \in K$ gibt es genau ein Element $-\lambda \in K$ mit $(-\lambda) + \lambda = 0$.

(3) Das Einselement 1 ist eindeutig: Ist $1' \in K$ ein Element mit $1'\lambda = \lambda$ für alle $\lambda \in K$, so ist $1' = 1$.

(4) Zu jedem $\lambda \in K$ mit $\lambda \neq 0$ gibt es genau ein Element $\lambda^{-1} \in K$ mit $\lambda^{-1}\lambda = 1$.

Beweis (1) Sei $0' \in K$ ein Element mit $0' + \lambda = \lambda$ für alle $\lambda \in K$; insbesondere ist dann $0' + 0 = 0$. Da aber $0 + \lambda = \lambda$ für alle $\lambda \in K$, ist auch $0 + 0' = 0'$, und nach (K2) ist $0 + 0' = 0' + 0$. Damit ist $0' = 0 + 0' = 0' + 0 = 0$.

(2) Sei $\lambda \in K$ und sei $\lambda' \in K$ mit $\lambda' + \lambda = 0$. Unter Anwendung von (K1), (K2), (K3) und (K4) folgt dann, dass

$$\begin{aligned}\lambda' &\overset{(K3)}{=} 0 + \lambda' \overset{(K4)}{=} ((-\lambda) + \lambda) + \lambda' \overset{(K1)}{=} (-\lambda) + (\lambda + \lambda') \\ &\overset{(K2)}{=} (-\lambda) + (\lambda' + \lambda) = (-\lambda) + 0 \overset{(K2)}{=} 0 + (-\lambda) \overset{(K3)}{=} -\lambda\,.\end{aligned}$$

(3) Sei $1' \in K$ ein Element mit $1'\lambda = \lambda$ für alle $\lambda \in K$; insbesondere ist dann $1' \cdot 1 = 1$. Da aber $1\lambda = \lambda$ für alle $\lambda \in K$, ist auch $1 \cdot 1' = 1'$, und nach (K6) ist $1 \cdot 1' = 1' \cdot 1$. Damit ist $1' = 1 \cdot 1' = 1' \cdot 1 = 1$.

(4) Sei $\lambda \in K \setminus \{0\}$ und sei $\lambda' \in K$ mit $\lambda'\lambda = 1$. Unter Anwendung von (K5), (K6), (K7) und (K8) folgt dann, dass

$$\begin{aligned}\lambda' &\overset{(K7)}{=} 1\lambda' \overset{(K8)}{=} (\lambda^{-1}\lambda)\lambda' \overset{(K5)}{=} \lambda^{-1}(\lambda\lambda') \\ &\overset{(K6)}{=} \lambda^{-1}(\lambda'\lambda) = \lambda^{-1}1 \overset{(K6)}{=} 1\lambda^{-1} \overset{(K7)}{=} \lambda^{-1}\,.\end{aligned}$$ □

Nach (K3) ist $0 + 0 = 0$ und damit ist $-0 = 0$, da nach Lemma 2.1 (2) -0 das eindeutige Element $\mu \in K$ mit $\mu + 0 = 0$ ist. Nach (K7) ist $1 \cdot 1 = 1$ und damit ist $1^{-1} = 1$, da nach Lemma 2.1 (4) 1^{-1} das eindeutige Element $\mu \in K$ mit $\mu 1 = 1$ ist.

Wenn aus dem Kontext klar ist, welche Verknüpfungen $+$ und $\cdot$ und Elemente 0 und 1 gemeint sind, dann wird lediglich K statt $(K, +, \cdot, 0, 1)$ geschrieben.

Ist K ein Körper, so wird eine Verknüpfung

$$\begin{aligned}- : K \times K &\to K \\ (\lambda, \mu) &\mapsto \lambda - \mu\end{aligned}$$

(genannt *Subtraktion*) durch $\lambda - \mu = \lambda + (-\mu)$ definiert.

Beispiele von Körpern

1. $\mathbb{R}$ mit der üblichen Addition und Multiplikation ist ein Körper. Hier ist das Nullelement 0 und das Einselement 1.

2. Die rationalen Zahlen $\mathbb{Q}$ (wieder mit der üblichen Addition und Multiplikation) bilden auch ein Körper. (Wieder ist das Nullelement 0 und das Einselement 1.)

3. *Die komplexen Zahlen* Sei $\mathbb{C} = \mathbb{R}^2 = \{(x, y) : x, y \in \mathbb{R}\}$; nun definiere eine Addition $+ : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ durch

$$(x, y) + (x', y') = (x + x', y + y')$$

und eine Multiplikation $\cdot : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ durch

$$(x, y)(x', y') = (xx' - yy', xy' + yx') \ .$$

Dann ist $(\mathbb{C}, +, \cdot)$ ein Körper. $(0, 0)$ ist die Null in $\mathbb{C}$ und $-(x, y) = (-x, -y)$ für alle $(x, y) \in \mathbb{C}$. Ferner ist das Element $(1, 0)$ die Eins in $\mathbb{C}$. Ist $(x, y) \in \mathbb{C}$ mit $(x, y) \neq (0, 0)$, so ist $\sigma = x^2 + y^2 > 0$ und

$$(x/\sigma, -y/\sigma)(x, y) = (x^2/\sigma + y^2/\sigma, xy/\sigma - yx/\sigma) = (1, 0) \ ,$$

d.h., $(x/\sigma, -y/\sigma) = (x, y)^{-1}$.

In der Praxis wird eine andere Schreibweise für die Elemente von $\mathbb{C}$ verwendet: Da $(x, 0) + (x', 0) = (x + x', 0)$ und $(x, 0)(x', 0) = (xx', 0)$ für alle x, $x' \in \mathbb{R}$, kann zunächst die Teilmenge $\mathbb{R} \times \{0\} = \{(x, 0) : x \in \mathbb{R}\}$ von $\mathbb{C}$ mit dem Körper $\mathbb{R}$ identifiziert werden, indem man x statt $(x, 0)$ schreibt. Auf diese Weise wird $\mathbb{R}$ als Teilmenge von $\mathbb{C}$ betrachtet. Das Element $(0, 1)$ in $\mathbb{C}$ wird *imaginäre Einheit* genannt und wird mit i bezeichnet. Für jedes $(x, y) \in \mathbb{C}$ gilt nun

$$(x, y) = (x, 0) + (0, y) = (x, 0) + (0, 1)(y, 0) = x + iy \ ,$$

also kann man $x + iy$ statt (x, y) schreiben. Mit dieser Schreibweise gilt

$$\begin{aligned} (x + iy) + (x' + iy') &= (x + x') + i(y + y') \\ (x + iy)(x' + iy') &= (xx' - yy') + i(xy + yx') \ . \end{aligned}$$

Außerdem ist 0 die Null und 1 die Eins in $\mathbb{C}$. Ferner ist

$$(0, 1)(0, 1) = (-1, 0) = -(1, 0) \ ,$$

d.h., $i^2 = -1$, und daher schreibt man manchmal $i = \sqrt{-1}$.

4. *Endliche Körper* Seien m, $n \in \mathbb{Z}$; dann schreibt man $m|n$ (m teilt n), wenn es $\ell \in \mathbb{Z}$ mit $n = \ell m$ gibt. Sei $n \geq 2$, und a, $b \in \mathbb{Z}$; man schreibt $a = b \bmod n$, wenn $n|(a - b)$. Zu jedem $m \in \mathbb{Z}$ gibt es ein eindeutiges Element $m \operatorname{Rest} n$ aus der Menge $\{0, \ldots, n - 1\}$ mit $m \operatorname{Rest} n = m \bmod n$.

Für $n \geq 2$ sei $\mathbb{Z}_n = \{0, 1, \ldots, n - 1\}$. Eine Addition $+ : \mathbb{Z}_n \times \mathbb{Z}_n \to \mathbb{Z}_n$ wird definiert durch

$$\ell + m = (\ell + m) \operatorname{Rest} n$$

und eine Multiplikation $\cdot : \mathbb{Z}_n \times \mathbb{Z}_n \to \mathbb{Z}_n$ durch

$$\ell m = (\ell m) \operatorname{Rest} n \ .$$

Für jedes $n \geq 2$ sind (K1), (K2), (K3), (K4), (K5), (K6), (K7) und (K9) erfüllt.

Satz 2.1 *Die Bedingung (K8) ist erfüllt genau dann, wenn n Primzahl ist. Damit ist $(\mathbb{Z}_n, +, \cdot, 0, 1)$ ein Körper genau dann, wenn n Primzahl ist.*

Beweis Nehme zuerst an, dass n keine Primzahl ist. Dann gibt es ℓ, $m > 1$ mit $\ell m = n$. Da $\ell < n$, $m < n$, kann ℓ und m als Elemente von $\mathbb{Z}_n$ betrachtet werden und in $\mathbb{Z}_n$ gilt $\ell m = 0$, da $n \operatorname{Rest} n = 0$. Daher kann $(\mathbb{Z}_n, +, \cdot)$ kein Körper sein. (*Übung:* Sei $(K, +, \cdot)$ ein Körper und λ, $\mu \in K$ mit $\lambda \neq 0$, $\mu \neq 0$; dann ist auch $\lambda\mu \neq 0$.)

Für die Umkehrung braucht man folgendes: Seien n, $m \geq 1$; dann gibt es $\ell, k \in \mathbb{Z}$, so dass $\ell n + km = (m, n)$, wobei (m, n) der größte gemeinsame Teiler von m und n ist. (Dieses Ergebnis kann man in jedem Buch über elementare Zahlentheorie finden.) Nehme jetzt an, dass p Primzahl ist und sei $m \in \mathbb{Z}_p$ mit $m \neq 0$. Dann gilt $(m, p) = 1$ und folglich gibt es ℓ, $k \in \mathbb{Z}$ mit $\ell p + km = 1$. Sei $k' = k \operatorname{Rest} p$; also gilt $k' - k = \ell' p$ für ein $\ell' \in \mathbb{Z}$, und daraus ergibt sich, dass

$$k'm - 1 = km - 1 + \ell' mp = (\ell + \ell' m)p \; .$$

Damit ist auch $k'm \operatorname{Rest} p = 1$; d.h., $k' = m^{-1}$ in $\mathbb{Z}_p$. Dies zeigt, dass in $\mathbb{Z}_p$ die Bedingung (K8) erfüllt ist. □

Wenn p Primzahl ist, dann bezeichnet man den Körper $\mathbb{Z}_p$ meistens mit $\mathbb{F}_p$.

Sei K ein Körper mit Einselement 1. Für $n \geq 1$ wird $n1$ als $1 + \cdots + 1$ (mit n Summanden) verstanden. Gilt $n1 \neq 0$ für alle $n \geq 1$, so nennt man K einen *Körper der Charakteristik Null*. Im anderen Falle ist die *Charakteristik* $\operatorname{char} K$ definiert als die kleinste positive natürliche Zahl p, für die $p1 = 0$ gilt.

Insbesondere ist $\operatorname{char} \mathbb{Q} = \operatorname{char} \mathbb{R} = \operatorname{char} \mathbb{C} = 0$ und $\operatorname{char} \mathbb{F}_p = p$.

Lemma 2.2 *Ist* $\operatorname{char} K \neq 0$*, dann ist* $\operatorname{char} K$ *eine Primzahl.*

Beweis Übung. □

Sei K ein endlicher Körper; dann gibt es eine Primzahl p und $n \geq 1$, so dass K aus genau p^n Elementen besteht (und $\operatorname{char} K = p$). Ist umgekehrt p eine Primzahl und $n \geq 1$, so gibt es einen (im wesentlichen eindeutigen) endlichen Körper, der aus genau p^n Elementen besteht.

Lineare Gleichungssysteme können über einem beliebigen Körper K betrachtet werden.

Sei $n \geq 1$; mit K^n wird die Menge aller n-Tupel von Elementen aus K bezeichnet. Ein Element von K^n hat also die Form $(\lambda_1, \ldots, \lambda_n)$ mit λ_1, …, λ_n Elementen aus K. Das Element $(0, \ldots, 0) \in K^n$ wird mit 0 bezeichnet.

Eine Addition auf K^n wird erklärt durch

$$(\lambda_1, \lambda_2, \ldots, \lambda_n) + (\mu_1, \mu_2, \ldots, \mu_n) = (\lambda_1 + \mu_1, \lambda_2 + \mu_2, \ldots, \lambda_n + \mu_n)$$

und eine Multiplikation eines Elements von K^n mit einem Element von K durch

$$\lambda(\mu_1, \mu_2, \ldots, \mu_n) = (\lambda\mu_1, \lambda\mu_2, \ldots, \lambda\mu_n) \; .$$

Sei $v \in K^n$; für jedes $\lambda \in K$ heißt dann λv ein *Vielfaches* von v.

Seien m, $n \geq 1$; eine $m \times n$ *Matrix* über K ist eine Anordnung von mn Elementen von K nach folgendem Schema

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Die Menge der $m \times n$ Matrizen über K wird mit $\mathrm{M}(m \times n, K)$ bezeichnet. Die Schreibweise $A = (a_{ij})_{1 \leq i \leq m,\, 1 \leq j \leq n}$ und $A = (a_{ij})$ wird weiter verwendet. Zeilen und Spalten werden wie im Fall $K = \mathbb{R}$ definiert; die Zeilen von A werden als Elemente von K^n und die Spalten von A als Elemente von K^m betrachtet. Das Element

$$\begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \in \mathrm{M}(m \times n, K)$$

wird mit 0 bezeichnet.

Seien m, $n \geq 1$ und $A = (a_{ij}) \in \mathrm{M}(m \times n, K)$, $b = (b_1, \ldots, b_m) \in K^m$; dann heißt

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

das *zu A und b gehörige lineare Gleichungssystem*.

Ein Element $(y_1, \ldots, y_n) \in K^n$ heißt *Lösung* des zu A und b gehörigen linearen Gleichungssystems, wenn y_1, …, y_n die m Gleichungen erfüllen, d.h., wenn

$$\begin{aligned} a_{11}y_1 + a_{12}y_2 + \cdots + a_{1n}y_n &= b_1 \\ a_{21}y_1 + a_{22}y_2 + \cdots + a_{2n}y_n &= b_2 \\ &\vdots \\ a_{m1}y_1 + a_{m2}y_2 + \cdots + a_{mn}y_n &= b_m \end{aligned}$$

Die Menge aller Lösungen des Systems wird mit $\mathrm{Lös}(A, b)$ bezeichnet.

Das Gleichungssystem heißt *lösbar*, wenn $\mathrm{Lös}(A, b)$ mindestens ein Element von K^n enthält. Das Gleichungssystem heißt *eindeutig lösbar*, wenn $\mathrm{Lös}(A, b)$ aus genau einem Element von K^n besteht.

Das zu A und b gehörige lineare Gleichungssystem heißt *homogen*, wenn $b = 0$. Ein homogenes Gleichungssystem besitzt stets die triviale Lösung $0 = (0, \ldots, 0)$. (Insbesondere ist ein homogenes Gleichungssystem stets lösbar.)

Eine Teilmenge U von K^n heißt *Untervektorraum* von K^n, wenn gilt:

(U0) $0 \in U$,

(U1) $u + v \in U$ für alle u, $v \in U$,

(U2) $\lambda u \in U$ für alle $\lambda \in K$, $u \in U$.

Satz 2.2 *Sei $A = (a_{ij}) \in \mathrm{M}(m \times n, K)$; dann ist $\mathrm{Lös}(A, 0)$ ein Untervektorraum von K^n. (Die Lösungsmenge eines homogenen linearen Gleichungssytems für n Unbekannte ist ein Untervektorraum von K^n.)*

Beweis Dieser ist fast identisch mit dem Beweis für Satz 1.1. □

Satz 2.3 *Seien $A \in \mathrm{M}(m \times n, K)$, $b \in K^m$, $v \in \mathrm{Lös}(A, b)$ und $v' \in K^n$. Dann gilt $v' \in \mathrm{Lös}(A, b)$ genau, wenn $v' = v + u$ für ein $u \in \mathrm{Lös}(A, 0)$. (Man erhält also alle Lösungen des zu A und b gehörigen Gleichungssystems, indem man zu einer speziellen Lösung dieses Systems alle Lösungen des zu A und 0 gehörigen Gleichungssystems addiert.)*

Beweis Dieser ist fast identisch mit dem Beweis für Satz 1.2. □

Für eine Matrix über einem beliebigen Körper K werden nun *elementare Zeilenumformungen* wie im reellen Fall definiert. Sie sind:

I Addition eines Vielfachen einer Zeile zu einer anderen Zeile.

II Vertauschen zweier Zeilen.

Lemma 2.3 *Wird eine Matrix A durch eine elementare Zeilenumformung zu einer Matrix A' verändert, so gilt $A' = 0$ genau dann, wenn $A = 0$.*

Beweis Dieser ist fast identisch mit dem Beweis für Lemma 1.1. □

Seien m, $n \geq 1$ und $A = (a_{ij}) \in M(m \times n, K)$, $b = (b_1, \ldots, b_m) \in K^m$; dann wird die $m \times (n + 1)$ reelle Matrix

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix}$$

mit (A, b) bezeichnet.

Satz 2.4 *Wird* (A, b) *durch eine elementare Zeilenumformung zu einer Matrix* (A', b') *verändert, so gilt* $\text{Lös}(A, b) = \text{Lös}(A', b')$.

Beweis Dieser ist fast identisch mit dem Beweis für Satz 1.3. □

Eine Matrix $A \in \text{M}(m \times n, K)$ hat *Zeilen-Stufen-Form*, wenn für jede Zeile der Matrix folgende zwei Bedingungen erfüllt sind:

— Sind die ersten p Elemente der Zeile Null für ein p mit $p < n$, so sind für alle folgenden Zeilen mindestens die ersten $p + 1$ Elemente Null.

— Sind alle Elemente der Zeile Null, so ist jedes Element von jeder der folgenden Zeilen Null.

Insbesondere hat die Null-Matrix 0 Zeilen-Stufen-Form.

Sei $A = (a_{ij}) \in \text{M}(m \times n, K)$ mit $A \neq 0$ und für $i = 1, \ldots, m$ sei

$$p_i = \begin{cases} \min\{1 \leq j \leq n : a_{ij} \neq 0\} & \text{falls } (a_{i1}, \ldots, a_{in}) \neq 0 \ , \\ 0 & \text{sonst} \ . \end{cases}$$

Die Matrix A hat also Zeilen-Stufen-Form genau dann, wenn

$$p_1 < p_2 < \cdots < p_r \ ,$$

wobei r der Index der letzten von Null verschiedenen Zeile von A ist. Hat A Zeilen-Stufen-Form, so nennt man $p_1, \ldots, p_r$ die *Treppen-Folge* von A.

Satz 2.5 *Seien* $A = (a_{ij}) \in \text{M}(m \times n, K)$, $b = (b_1, \ldots, b_m) \in K^m$ *mit* $(A, b) \neq 0$. *Nehme an, dass* (A, b) *Zeilen-Stufen-Form hat, und sei* $p_1, \ldots, p_r$ *die Treppen-Folge von* A. *Dann ist das zu* A *und* b *gehörige lineare Gleichungssystem lösbar genau, wenn* $p_r \neq n + 1$. *Ferner ist das Gleichungssystem eindeutig lösbar genau dann, wenn* $r = n$ *und* $p_j = j$ *für jedes* $j = 1, \ldots, n$.

Beweis Dieser ist fast identisch mit dem Beweis für Satz 1.4. □

Satz 2.6 *Jede Matrix in* $\text{M}(m \times n, K)$ *läßt sich durch eine Folge von elementaren Zeilenumformungen in eine Matrix mit Zeilen-Stufen-Form überführen.*

Beweis Der Gaußsche Algorithmus macht einen Sinn (und funktioniert) für einen beliebigen Körper K. □

Seien $m, n \geq 1$ und $A = (a_{ij}) \in \text{M}(m \times n, K)$, $b = (b_1, \ldots, b_m) \in K^m$. Es ist klar, dass $\text{Lös}(0, 0) = K^n$; nehme also an, dass $(A, b) \neq 0$. Mit Hilfe des Gaußschen Algorithmus kann die Matrix (A, b) in eine Matrix (A', b') mit Zeilen-Stufen-Form überführt werden und nach Satz 2.4 gilt dann $\text{Lös}(A, b) = \text{Lös}(A', b')$. Ferner gilt nach Lemma 2.3, dass $(A', b') \neq 0$. Folglich kann Satz 2.5 verwendet werden, um festzustellen, ob das zu A und b gehörige lineare Gleichungssystem lösbar bzw. eindeutig lösbar ist.

3 Vektorräume

Im Folgenden sei K ein Körper. Ein *Vektorraum* über K ist ein 4-Tupel $(V, +, \cdot, 0)$ bestehend aus einer Menge V, einer Verknüpfung (*Addition*)

$$\begin{aligned} + : V \times V &\to V \\ (\lambda, \mu) &\mapsto \lambda + \mu \end{aligned}$$

einer Verknüpfung (*Multiplikation mit Skalaren*)

$$\begin{aligned} \cdot : K \times V &\to V \\ (\lambda, v) &\mapsto \lambda v \end{aligned}$$

und einem Element $0 \in V$, für das folgendes gilt:

(1) $(u + v) + w = u + (v + w)$ für alle u, v, $w \in V$.

(2) $u + v = v + u$ für alle u, $v \in V$.

(3) $0 + v = v$ für alle $v \in V$.

(4) Zu jedem $v \in V$ gibt es ein Element $-v \in V$ mit $(-v) + v = 0$.

(5) $(\lambda\mu)v = \lambda(\mu v)$ für alle λ, $\mu \in K$, $v \in V$.

(6) $1v = v$ für alle $v \in V$.

(7) $\lambda(u + v) = \lambda u + \lambda v$ für alle $\lambda \in K$, u, $v \in V$.

(8) $(\lambda + \mu)v = \lambda v + \mu v$ für alle λ, $\mu \in K$, $v \in V$.

Bemerkung: Nach der üblichen Konvention soll die Addition in V weniger stark binden als die Multiplikation mit Skalaren. ($\lambda u + \lambda v$ bedeutet also $(\lambda u) + (\lambda v)$ und $\lambda v + \mu v$ bedeutet $(\lambda v) + (\mu v)$.)

Das Element 0 heißt das *Nullelement* oder die *Null*. Ein Vektorraum über K wird auch K*-Vektorraum* genannt. Einen Vektorraum über $K = \mathbb{R}$ bzw. $K = \mathbb{C}$ nennt man *reellen Vektorraum* bzw. *komplexen Vektorraum*.

Lemma 3.1 *Sei $(V, +, \cdot, 0)$ ein Vektorraum über K.*

(1) Das Nullelement 0 ist eindeutig: Ist $0' \in V$ ein Element mit $0' + v = v$ für alle $v \in V$, so ist $0' = 0$.

(2) Zu jedem $v \in V$ gibt es genau ein Element $-v \in V$ mit $(-v) + v = 0$.

Beweis Übung. □

Wenn aus dem Kontext klar ist, welche Verknüpfungen + und · und welches Element 0 gemeint sind, dann wird lediglich V statt $(V, +, \cdot, 0)$ geschrieben.

Lemma 3.2 *Sei V ein Vektorraum über K und seien $\lambda \in K$, $v \in V$. Dann ist $\lambda v \neq 0$ genau, wenn $\lambda \neq 0$ und $v \neq 0$.*

Beweis Übung. □

Beispiele von Vektorräumen:

1. Sei $n \geq 1$; definiere $+ : K^n \times K^n \to K^n$ und $\cdot : K \times K^n \to K^n$ durch

$$(\lambda_1, \ldots, \lambda_n) + (\mu_1, \ldots, \mu_n) = (\lambda_1 + \mu_1, \ldots, \lambda_n + \mu_n) ,$$
$$\lambda(\mu_1, \ldots, \mu_n) = (\lambda\mu_1, \ldots, \lambda\mu_n) .$$

Dann ist $(K^n, +, \cdot, 0)$ ein Vektorraum über K, wobei $0 = (0, \ldots, 0)$. Für jedes $(\lambda_1, \ldots, \lambda_n) \in K^n$ ist $-(\lambda_1, \ldots, \lambda_n) = (-\lambda_1, \ldots, -\lambda_n)$.

2. Sei X eine Menge. Für f, $g \in \mathrm{Abb}(X, K)$ und $\lambda \in K$ definiere Abbildungen $f + g$, $\lambda f \in \mathrm{Abb}(X, K)$ durch $(f + g)(x) = f(x) + g(x)$ und $(\lambda f)(x) = \lambda f(x)$. Mit diesen Verknüpfungen $+ : \mathrm{Abb}(X, K) \times \mathrm{Abb}(X, K) \to \mathrm{Abb}(X, K)$ und $\cdot : K \times \mathrm{Abb}(X, K) \to \mathrm{Abb}(X, K)$ ist $\mathrm{Abb}(X, K)$ ein Vektorraum über K. Die Nullabbildung $0 : X \to K$ (mit $0(x) = 0$ für alle $x \in X$) ist die Null und für jedes $f \in \mathrm{Abb}(X, K)$ ist $-f \in \mathrm{Abb}(X, K)$ durch $(-f)(x) = -f(x)$ für alle $x \in X$ gegeben.

3. Seien m, $n \geq 1$ und sei $\mathrm{M}(m \times n, K)$ die Menge aller $m \times n$ Matrizen über K. Für $A = (a_{ij})$, $B = (b_{ij}) \in \mathrm{M}(m \times n, K)$ und $\lambda \in K$ definiere Matrizen $A + B$, $\lambda A \in \mathrm{M}(m \times n, K)$ durch $A + B = (a_{ij} + b_{ij})$ und $\lambda A = (\lambda a_{ij})$. Mit diesen Verknüpfungen $+ : \mathrm{M}(m \times n, K) \times \mathrm{M}(m \times n, K) \to \mathrm{M}(m \times n, K)$ und $\cdot : K \times \mathrm{M}(m \times n, K) \to \mathrm{M}(m \times n, K)$ ist $\mathrm{M}(m \times n, K)$ ein K-Vektorraum. Die Nullmatrix 0 ist die Null in $\mathrm{M}(m \times n, K)$ und für jedes $A = (a_{ij}) \in \mathrm{M}(m \times n, K)$ ist $-A = (-a_{ij})$.

4. (Verallgemeinerung von 2.) Sei X eine Menge und sei V ein K-Vektorraum. Für f, $g \in \mathrm{Abb}(X, V)$ und $\lambda \in K$ definiere Abbildungen $f + g$, $\lambda f \in \mathrm{Abb}(X, V)$ durch $(f + g)(x) = f(x) + g(x)$ und $(\lambda f)(x) = \lambda f(x)$. Mit diesen Verknüpfungen $+ : \mathrm{Abb}(X, V) \times \mathrm{Abb}(X, V) \to \mathrm{Abb}(X, V)$ und $\cdot : K \times \mathrm{Abb}(X, V) \to \mathrm{Abb}(X, V)$ ist $\mathrm{Abb}(X, V)$ ein Vektorraum über K. Die Nullabbildung $0 : X \to V$ ist die Null und für $f \in \mathrm{Abb}(X, V)$ ist $-f \in \mathrm{Abb}(X, V)$ durch $(-f)(x) = -f(x)$ gegeben.

Sei V ein Vektorraum über K. Eine Teilmenge $U \subset V$ heißt *Untervektorraum* von V, wenn $0 \in U$ und $\lambda u + \mu v \in U$ für alle u, $v \in U$, λ, $\mu \in K$. Insbesondere

ist V selbst Untervektorraum von V. Ferner ist $\{0\}$ stets Untervektorraum von V, (da nach Lemma 3.2 $\lambda 0 = 0$ für jedes $\lambda \in K$).

Sei U ein Untervektorraum von V. Dann induzieren die Verknüpfungen $+$ und $\cdot$ Verknüpfungen $+ : U \times U \to U$ und $\cdot : K \times U \to U$. Mit diesen induzierten Verknüpfungen (und mit dem Nullelement 0 aus V) ist U ein K-Vektorraum.

Lemma 3.3 *Sei U ein Untervektorraum von V und seien $v_1, \ldots, v_m$ Elemente von U (mit $m \geq 1$). Für alle $\lambda_1, \ldots, \lambda_m \in K$ ist dann $\lambda_1 v_1 + \cdots + \lambda_m v_m$ wieder ein Element von U.*

Beweis Es wird durch Induktion nach n gezeigt, dass $\lambda_1 v_1 + \cdots + \lambda_n v_n$ ein Element von U ist für jedes $n = 1, \ldots, m$. Nach (U2) ist $\lambda_1 v_1$ ein Element von U. Sei nun n mit $1 \leq n < m$ und nehme an, dass $\lambda_1 v_1 + \cdots + \lambda_n v_n$ ein Element von U ist. Nach (U2) ist $\lambda_{n+1} v_{n+1}$ ein Element von U, und daraus folgt nach (U1), dass

$$\lambda_1 v_1 + \cdots + \lambda_{n+1} v_{n+1} = (\lambda_1 v_1 + \cdots + \lambda_n v_n) + \lambda_{n+1} v_{n+1}$$

auch ein Element von U ist. Damit ist $\lambda_1 v_1 + \cdots + \lambda_n v_n$ ein Element von U für jedes $n = 1, \ldots, m$. Insbesondere ist $\lambda_1 v_1 + \cdots + \lambda_m v_m$ ein Element von U. □

Im folgenden sei V ein Vektorraum über dem Körper K.

Seien $v_1, \ldots, v_m \in V$ (mit $m \geq 1$). Ein Element $v \in V$ heißt *Linearkombination* von $v_1, \ldots, v_m$, wenn es $\lambda_1, \ldots, \lambda_m \in K$ gibt, so dass

$$v = \lambda_1 v_1 + \cdots + \lambda_m v_m \ .$$

Die Menge aller Linearkombinationen von $v_1, \ldots, v_m$ nennt man die *lineare Hülle* von $v_1, \ldots, v_m$; sie wird mit $\mathrm{L}(v_1, \ldots, v_m)$ bezeichnet.

Wichtiges Beispiel für den Vektorraum K^m: Für $j = 1, \ldots, m$ sei

$$e_j = (0, \ldots, 0, 1, 0, \ldots, 0)$$

mit der Eins in der j-ten Komponente. Dann gilt

$$\lambda_1 e_1 + \cdots + \lambda_m e_m = (\lambda_1, \ldots, \lambda_m)$$

für alle $\lambda_1, \ldots, \lambda_m \in K$ und daraus ergibt sich, dass $\mathrm{L}(e_1, \ldots, e_m) = K^m$.

Seien nun $v_1, \ldots, v_m \in V$ (mit $m \geq 1$).

Satz 3.1 $\mathrm{L}(v_1, \ldots, v_m)$ *ist ein Untervektorraum von V mit $v_j \in \mathrm{L}(v_1, \ldots, v_m)$ für jedes $j = 1, \ldots, m$.*

Beweis Nach Lemma 3.2 ist $0 = 0v_1 + \cdots + 0v_m \in \mathrm{L}(v_1, \ldots, v_m)$. Seien nun u, $v \in \mathrm{L}(v_1, \ldots, v_m)$; dann gibt es Elemente $\lambda_1, \ldots, \lambda_m, \mu_1, \ldots, \mu_m \in K$, so dass $u = \lambda_1 v_1 + \cdots + \lambda_m v_m$ und $v = \mu_1 v_1 + \cdots + \mu_m v_m$. Folglich gilt

$$u+v = \lambda_1 v_1 + \cdots + \lambda_m v_m + \mu_1 v_1 + \cdots + \mu_m v_m = (\lambda_1 + \mu_1)v_1 + \cdots + (\lambda_m + \mu_m)v_m \ ,$$

und damit ist $u + v \in \mathrm{L}(v_1, \ldots, v_m)$. Für jedes $\lambda \in K$ gilt ferner

$$\lambda u = \lambda(\lambda_1 v_1 + \cdots + \lambda_m v_m) = (\lambda\lambda_1)v_1 + \cdots + (\lambda\lambda_m)v_m \ ,$$

und folglich ist auch $\lambda u \in \mathrm{L}(v_1, \ldots, v_m)$. Dies zeigt also, dass $\mathrm{L}(v_1, \ldots, v_m)$ ein Untervektorraum von V ist.

Sei nun j mit $1 \leq j \leq m$ und definiere $\lambda_1, \ldots, \lambda_m \in K$ durch

$$\lambda_i = \begin{cases} 1 & \text{falls } i = j \ , \\ 0 & \text{falls } i \neq j \ . \end{cases}$$

Dann ist $v_j = 0 + \cdots + 0 + 1v_j + 0 + \cdots + 0 = \lambda_1 v_1 + \cdots + \lambda_m v_m$, und damit ist $v_j \in \mathrm{L}(v_1, \ldots, v_m)$. □

Satz 3.2 *Sei U ein Untervektorraum von V mit $v_j \in U$ für jedes j. Dann gilt $\mathrm{L}(v_1, \ldots, v_m) \subset U$. Damit ist $\mathrm{L}(v_1, \ldots, v_m)$ der kleinste Untervektorraum von V, der die Elemente $v_1, \ldots, v_m$ enthält.*

Beweis Sei $v \in \mathrm{L}(v_1, \ldots, v_m)$; dann gibt es Elemente $\lambda_1, \ldots, \lambda_m \in K$, so dass $v = \lambda_1 v_1 + \cdots + \lambda_m v_m$ und nach Lemma 3.3 ist $\lambda_1 v_1 + \cdots + \lambda_m v_m$ ein Element von U, d.h., $v \in U$. Folglich ist $\mathrm{L}(v_1, \ldots, v_m) \subset U$. □

In den folgenden Lemmas sei $m \geq 1$ und seien $v_1, \ldots, v_m \in V$.

Lemma 3.4 *Es gilt $\mathrm{L}(v_1, \ldots, v_m) = \{0\}$ genau dann, wenn $v_j = 0$ für jedes $j = 1, \ldots, m$.*

Beweis Nach Satz 3.1 enthält $\mathrm{L}(v_1, \ldots, v_m)$ die Elemente $v_1, \ldots, v_m$ und folglich ist $\mathrm{L}(v_1, \ldots, v_m) \neq \{0\}$, wenn $v_j \neq 0$ für ein j. Ist andererseits $v_j = 0$ für jedes $j = 1, \ldots, m$, dann gilt $\lambda_1 v_1 + \cdots + \lambda_m v_m = 0$ für alle $\lambda_1, \ldots, \lambda_m \in K$ und damit ist $\mathrm{L}(v_1, \ldots, v_m) = \{0\}$. □

Lemma 3.5 *Seien $\lambda_1, \ldots, \lambda_m \in K$ mit $\lambda_j \neq 0$ für jedes $j = 1, \ldots, m$. Dann gilt $\mathrm{L}(\lambda_1 v_1, \ldots, \lambda_m v_m) = \mathrm{L}(v_1, \ldots, v_m)$.*

Beweis Nach Satz 3.1 ist $\lambda_j v_j \in \mathrm{L}(v_1, \ldots, v_m)$ für jedes $j = 1, \ldots, m$ und daraus folgt nach Satz 3.2, dass $\mathrm{L}(\lambda_1 v_1, \ldots, \lambda_m v_m) \subset \mathrm{L}(v_1, \ldots, v_m)$. Umgekehrt ist

$$v_j = 1v_j = (\lambda_j^{-1}\lambda_j)v_j = \lambda_j^{-1}(\lambda_j v_j) \in \mathrm{L}(\lambda_1 v_1, \ldots, \lambda_m v_m)$$

für jedes j und damit ist auch $\mathrm{L}(v_1, \ldots, v_m) \subset \mathrm{L}(\lambda_1 v_1, \ldots, \lambda_m v_m)$. □

Lemma 3.6 *Für jede Permutation* $\{i_1, \ldots, i_m\}$ *von* $\{1, \ldots, m\}$ *gilt*

$$\mathrm{L}(v_{i_1}, \ldots, v_{i_m}) = \mathrm{L}(v_1, \ldots, v_m) \,.$$

Beweis Nach Satz 3.1 ist $v_{i_j} \in \mathrm{L}(v_1, \ldots, v_m)$ für jedes $j = 1, \ldots, m$ und daraus folgt nach Satz 3.2, dass $\mathrm{L}(v_{i_1}, \ldots, v_{i_m}) \subset L(v_1, \ldots, v_m)$. Genauso gilt dann auch $\mathrm{L}(v_1, \ldots, v_m) \subset \mathrm{L}(v_{i_1}, \ldots, v_{i_m})$. □

Lemma 3.7 *Sei* $u \in V$. *Dann gilt* $\mathrm{L}(v_1, \ldots, v_m) \subset \mathrm{L}(v_1, \ldots, v_m, u)$. *Ferner gilt* $\mathrm{L}(v_1, \ldots, v_m) = \mathrm{L}(v_1, \ldots, v_m, u)$ *genau dann, wenn* $u \in \mathrm{L}(v_1, \ldots, v_m)$.

Beweis Nach Satz 3.1 ist $v_j \in \mathrm{L}(v_1, \ldots, v_m, u)$ für jedes j und daraus ergibt sich nach Satz 3.2, dass $\mathrm{L}(v_1, \ldots, v_m) \subset \mathrm{L}(v_1, \ldots, v_m, u)$.

Gilt $\mathrm{L}(v_1, \ldots, v_m) = \mathrm{L}(v_1, \ldots, v_m, u)$, so ist $u \in \mathrm{L}(v_1, \ldots, v_m)$, da nach Satz 3.1 u in $\mathrm{L}(v_1, \ldots, v_m, u)$ liegt. Ist umgekehrt $u \in \mathrm{L}(v_1, \ldots, v_m)$, so ist nach Satz 3.2 $\mathrm{L}(v_1, \ldots, v_m, u) \subset \mathrm{L}(v_1, \ldots, v_m)$ und damit $\mathrm{L}(v_1, \ldots, v_m) = \mathrm{L}(v_1, \ldots, v_m, u)$, da nach Satz 3.1 $v_j \in \mathrm{L}(v_1, \ldots, v_m)$ für jedes j. □

Lemma 3.8 *Seien* $u, w \in V$. *Es gilt* $u \in \mathrm{L}(v_1, \ldots, v_m, w) \setminus \mathrm{L}(v_1, \ldots, v_m)$ *genau dann, wenn* $w \in \mathrm{L}(v_1, \ldots, v_m, u) \setminus \mathrm{L}(v_1, \ldots, v_m)$.

Beweis Nehme zunächst an, dass $w \in \mathrm{L}(v_1, \ldots, v_m, u) \setminus \mathrm{L}(v_1, \ldots, v_m)$. Dann ist insbesondere $\mathrm{L}(v_1, \ldots, v_m, u) \neq \mathrm{L}(v_1, \ldots, v_m)$ und daraus folgt nach Lemma 3.7, dass $u \notin \mathrm{L}(v_1, \ldots, v_m)$. Da $w \in \mathrm{L}(v_1, \ldots, v_m, u)$, gibt es $\lambda_1, \ldots, \lambda_{m+1} \in K$, so dass $w = \lambda_1 v_1 + \cdots + \lambda_m v_m + \lambda_{m+1} u$. Ferner ist $\lambda_{m+1} \neq 0$, sonst würde w in $\mathrm{L}(v_1, \ldots, v_m)$ liegen. Damit ist

$$u = \mu_1 v_1 + \cdots + \mu_m v_m + \mu_{m+1} w \,,$$

wobei $\mu_j = (-1)\lambda_{m+1}^{-1}\lambda_j$ für $j = 1, \ldots, m$ und $\mu_{m+1} = \lambda_{m+1}^{-1}$. Dies zeigt also, dass $u \in \mathrm{L}(v_1, \ldots, v_m, w) \setminus \mathrm{L}(v_1, \ldots, v_m)$.

Der Beweis für die Umkehrung ist identisch, die Rollen von u und w müssen lediglich getauscht werden. □

Für den Vektorraum K^m gibt es einen engen Zusammenhang zwischen der Frage, ob ein Vektor $v \in K^m$ in der linearen Hülle von Vektoren $v_1, \ldots, v_n$ liegt, und der Frage, ob ein entsprechendes Gleichungssystem lösbar ist.

Seien $v_1, \ldots, v_n \in K^m$ (mit $n \geq 1$) mit $v_j = (\alpha_{1j}, \ldots, \alpha_{mj})$ für jedes $j = 1, \ldots, n$. Dann gibt es die Matrix $A = (\alpha_{ij}) \in \mathrm{M}(m \times n, K)$, in der die Vektoren $v_1, \ldots, v_n$ als Spalten vorkommen.

Satz 3.3 *Für jedes $v \in K^m$ gilt $v \in \mathrm{L}(v_1, \ldots, v_n)$ genau dann, wenn das zu A und v gehörige lineare Gleichungssystem lösbar ist.*

Beweis Seien $\lambda_1, \ldots, \lambda_n \in K$; dann bedeutet $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$ genau, dass $\lambda_1 \alpha_{i1} + \cdots + \lambda_n \alpha_{in} = \beta_i$ für jedes $i = 1, \ldots, m$, wobei $v = (\beta_1, \ldots, \beta_m)$. Also gilt $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$ genau dann, wenn $(\lambda_1, \ldots, \lambda_n) \in \mathrm{Lös}(A, v)$, und damit liegt v in $\mathrm{L}(v_1, \ldots, v_n)$ genau dann, wenn $\mathrm{Lös}(A, v) \neq \varnothing$. □

Seien $v_1, \ldots, v_m \in V$ (mit $m \geq 1$). Man sagt, dass $v_1, \ldots, v_m$ *linear abhängig* sind, wenn es $\lambda_1, \ldots, \lambda_m \in K$ mit $\lambda_j \neq 0$ für mindestens ein j gibt, so dass

$$\lambda_1 v_1 + \cdots + \lambda_m v_m = 0 \, .$$

Die Vektoren $v_1, \ldots, v_m$ sind *linear unabhängig*, wenn sie nicht linear abhängig sind. Mit anderen Worten sind die Vektoren $v_1, \ldots, v_m$ linear unabhängig genau dann, wenn aus $\lambda_1 v_1 + \cdots + \lambda_m v_m = 0$ stets folgt, dass $\lambda_1 = \cdots = \lambda_m = 0$, d.h., wenn eine Linearkombination von $v_1, \ldots, v_m$ nur dann Null sein kann, wenn alle "Koeffizienten" verschwinden.

Wichtiges Beispiel für den Vektorraum K^m: Für $j = 1, \ldots, m$ sei wieder

$$e_j = (0, \ldots, 0, 1, 0, \ldots, 0)$$

mit der Eins in der j-ten Komponenten. Dann sind $e_1, \ldots, e_m$ linear unabhängig: Sind $\lambda_1, \ldots, \lambda_m \in K$ mit $\lambda_1 e_1 + \cdots + \lambda_m e_m = 0$, so ist $(\lambda_1, \ldots, \lambda_m) = 0$, d.h., $\lambda_j = 0$ für jedes $j = 1, \ldots, m$, da $\lambda_1 e_1 + \cdots + \lambda_m e_m = (\lambda_1, \ldots, \lambda_m)$.

In den folgenden Lemmas sei $m \geq 1$ und seien $v_1, \ldots, v_m \in V$.

Lemma 3.9 *Sind $v_1, \ldots, v_m$ linear unabhängig, so ist $v_j \neq 0$ für jedes j und ferner ist $v_j \neq v_k$, falls $j \neq k$.*

Beweis Nehme an, dass $v_j = 0$ für ein j, und definiere $\lambda_1, \ldots, \lambda_m \in K$ durch

$$\lambda_i = \begin{cases} 1 & \text{falls } i = j \, , \\ 0 & \text{falls } i \neq j \, . \end{cases}$$

Dann ist $\lambda_i \neq 0$ für mindestens ein i und nach Lemma 3.2 ist

$$\begin{aligned}\lambda_1 v_1 + \cdots + \lambda_m v_m &= 0v_1 + \cdots + 0v_{j-1} + 1v_j + 0v_{j+1} + \cdots + 0v_m \\ &= 0 + \cdots + 0 + 1v_j + 0 + \cdots + 0 = 1v_j = v_j = 0 \ .\end{aligned}$$

Damit wären $v_1, \ldots, v_m$ linear abhängig. Folglich ist $v_j \neq 0$ für jedes j.

Nehme nun an, dass es j, k mit $j \neq k$ und $v_j = v_k$ gibt, und definiere diesmal

$$\lambda_i = \begin{cases} -1 & \text{falls } i = j \ , \\ 1 & \text{falls } i = k \ , \\ 0 & \text{sonst .} \end{cases}$$

Dann ist $\lambda_i \neq 0$ für mindestens ein i und nach Lemma 3.1 und Lemma 3.2 ist

$$\lambda_1 v_1 + \cdots + \lambda_m v_m = (-1)v_j + 1v_k = (-v_j) + v_j = 0 \ ,$$

und wieder wären $v_1, \ldots, v_m$ linear abhängig. Folglich ist $v_j \neq v_k$, falls $j \neq k$. □

Lemma 3.10 *Seien $\lambda_1, \ldots, \lambda_m \in K$ mit $\lambda_j \neq 0$ für jedes $j = 1, \ldots, m$. Die Vektoren $\lambda_1 v_1, \ldots, \lambda_m v_m$ sind genau dann linear unabhängig, wenn $v_1, \ldots, v_m$ linear unabhängig sind.*

Beweis Es wird gezeigt, dass $v_1, \ldots, v_m$ genau dann linear abhängig sind, wenn $\lambda_1 v_1, \ldots, \lambda_m v_m$ linear abhängig sind. Nehme zunächst an, dass $\lambda_1 v_1, \ldots, \lambda_m v_m$ linear abhängig sind. Dann gibt es $\mu_1, \ldots, \mu_m \in K$ mit $\mu_j \neq 0$ für mindestens ein j, so dass $\mu_1(\lambda_1 v_1) + \cdots + \mu_m(\lambda_m v_m) = 0$. Für $j = 1, \ldots, m$ sei $\mu'_j = \mu_j \lambda_j$; dann ist $\mu'_j \neq 0$ für mindestens ein j, da $\mu'_j = 0$ genau dann, wenn $\mu_j = 0$, und

$$\mu'_1 v_1 + \cdots + \mu'_m v_m = (\mu_1\lambda_1)v_1 + \cdots + (\mu_m\lambda_m)v_m = \mu_1(\lambda_1 v_1) + \cdots + \mu_m(\lambda_m v_m) = 0 \ .$$

Damit sind $v_1, \ldots, v_m$ linear abhängig.

Nehme nun umgekehrt an, dass $v_1, \ldots, v_m$ linear abhängig sind. Dann gibt es $\mu_1, \ldots, \mu_m \in K$ mit $\mu_j \neq 0$ für mindestens ein j, so dass $\mu_1 v_1 + \cdots + \mu_m v_m = 0$. Für $j = 1, \ldots, m$ sei $\mu'_j = \mu_j \lambda_j^{-1}$; dann ist $\mu'_j \neq 0$ für mindestens ein j, da $\mu'_j = 0$ genau dann, wenn $\mu_j = 0$. Ferner gilt $\mu'_j \lambda_j = \mu_j \lambda_j^{-1} \lambda_j = \mu_j 1 = \mu_j$ für jedes j und daraus ergibt sich, dass

$$\begin{aligned}\mu'_1(\lambda_1 v_1) + \cdots + \mu'_m(\lambda_m v_m) &= (\mu'_1\lambda_1)v_1 + \cdots + (\mu'_m\lambda_m)v_m \\ &= \mu_1(\lambda_1 v_1) + \cdots + \mu_m(\lambda_m v_m) = 0 \ .\end{aligned}$$

Folglich sind $\lambda_1 v_1, \ldots, \lambda_m v_m$ linear abhängig. □

Lemma 3.11 *Seien $i_1, \ldots, i_n$ (mit $n \geq 1$) paarweise verschiedene Elemente aus der Menge $\{1, 2, \ldots, m\}$. (Es gilt also $n \leq m$, $1 \leq i_j \leq m$ für $j = 1, \ldots, n$ und $i_j \neq i_k$, falls $j \neq k$.) Sind $v_1, \ldots, v_m$ linear unabhängig, so sind auch $v_{i_1}, \ldots, v_{i_n}$ linear unabhängig.*

Beweis Nehme an, dass $v_{i_1}, \ldots, v_{i_n}$ linear abhängig sind. Es gibt also Elemente $\lambda_{i_1}, \ldots, \lambda_{i_n} \in K$ mit $\lambda_{i_j} \neq 0$ für mindestens ein j, so dass $\lambda_{i_1} v_{i_1} + \cdots + \lambda_{i_n} v_{i_n} = 0$. Definiere $\mu_1, \ldots, \mu_m \in K$ durch

$$\mu_k = \begin{cases} \lambda_{i_j} & \text{falls } k = i_j \text{ für ein } j = 1, \ldots, n \text{ , ,} \\ 0 & \text{sonst .} \end{cases}$$

Dann ist $\mu_k \neq 0$ für mindestens ein k und nach Lemma 3.2 ist

$$\mu_1 v_1 + \cdots + \mu_m v_m = \lambda_{i_1} v_{i_1} + \cdots + \lambda_{i_n} v_{i_n} = 0 \, .$$

Aber dies ist nicht möglich, da dann $v_1, \ldots, v_m$ linear abhängig wären. Also sind $v_{i_1}, \ldots, v_{i_n}$ linear unabhängig. □

Seien $v_1, \ldots, v_m$ linear unabhängig. Dann gibt es folgende spezielle Fälle von Lemma 3.11:

— Für jedes $n = 1, \ldots, m$ sind $v_1, \ldots, v_n$ linear unabhängig.

— Ist $m \geq 2$, so sind $v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_m$ linear unabhängig für jedes $j = 1, \ldots, m$.

— Für jede Permutation $\{i_1, \ldots, i_m\}$ von $\{1, \ldots, m\}$ sind $v_{i_1}, \ldots, v_{i_m}$ linear unabhängig.

Satz 3.4 *(1) Sei $v \in V$; dann ist v linear unabhängig genau, wenn $v \neq 0$.*

(2) Seien $v_1, \ldots, v_m \in V$ mit $m \geq 2$. Dann sind $v_1, \ldots, v_m$ linear unabhängig genau, wenn $v_k \notin L(v_1, \ldots, v_{k-1}, v_{k+1}, \ldots, v_m)$ für jedes $k = 1, \ldots, m$.

Beweis (1) Sei $\lambda \in K$ mit $\lambda \neq 0$; nach Lemma 3.2 gilt, dass $\lambda v = 0$ genau dann, wenn $v = 0$. Folglich ist v linear abhängig genau dann, wenn $v = 0$, d.h., v ist linear unabhängig genau dann, wenn $v \neq 0$,

(2) Es wird gezeigt, dass $v_k \in \mathrm{L}(v_1, \ldots, v_{k-1}, v_{k+1}, \ldots, v_m)$ für ein k genau dann gilt, wenn $v_1, \ldots, v_m$ linear abhängig sind.

Nehme zunächst an, dass $v_k \in \mathrm{L}(v_1, \ldots, v_{k-1}, v_{k+1}, \ldots, v_m)$ für ein k. Dann gibt es $\lambda_1, \ldots, \lambda_{k-1}, \lambda_{k+1}, \ldots, \lambda_m \in K$, so dass

$$v_k = \lambda_1 v_1 + \cdots + \lambda_{k-1} v_{k-1} + \lambda_{k+1} v_{k+1} + \cdots + \lambda_m v_m \, .$$

Setze $\lambda_k = -1$; dann ist $\lambda_k \neq 0$ und

$$\begin{aligned}\lambda_1 v_1 + \cdots + \lambda_m v_m &= \lambda_1 v_1 + \cdots + \lambda_{k-1} v_{k-1} + (-1) v_k + \lambda_{k+1} v_{k+1} + \cdots + \lambda_m v_m \\ &= (-1) v_k + \lambda_1 v_1 + \cdots + \lambda_{k-1} v_{k-1} + \lambda_{k+1} v_{k+1} + \cdots + \lambda_m v_m \\ &= (-1) v_k + v_k = (-v_k) + v_k = 0 \,,\end{aligned}$$

(da $-v = (-1)v$ für jedes $v \in V$). Damit sind $v_1, \ldots, v_m$ linear abhängig.

Nehme nun umgekehrt an, dass $v_1, \ldots, v_m$ linear abhängig sind. Dann gibt es $\lambda_1, \ldots, \lambda_m \in K$ mit $\lambda_j \neq 0$ für mindestens ein j, so dass $\lambda_1 v_1 + \cdots + \lambda_m v_m = 0$. Wähle k mit $\lambda_k \neq 0$ und für jedes j mit $j \neq k$ sei $\mu_j = \nu \lambda_j$, wobei $\nu = (-1)\lambda_k^{-1}$. Dann ist $\nu \lambda_k = (-1)(\lambda_k^{-1} \lambda_k) = (-1)1 = -1$ und

$$\begin{aligned}v_k &= 0 + v_k = \nu 0 + v_k = \nu(\lambda_1 v_1 + \cdots + \lambda_m v_m) + v_k \\ &= (\nu\lambda_1) v_1 + \cdots + (\nu\lambda_m) v_m + v_k \\ &= \mu_1 v_1 + \cdots + \mu_{k-1} v_{k-1} + \mu_{k+1} v_{k+1} + \cdots + \mu_m v_m + (\nu\lambda_k) v_k + v_k \\ &= \mu_1 v_1 + \cdots + \mu_{k-1} v_{k-1} + \mu_{k+1} v_{k+1} + \cdots + \mu_m v_m + (-1) v_k + v_k \\ &= \mu_1 v_1 + \cdots + \mu_{k-1} v_{k-1} + \mu_{k+1} v_{k+1} + \cdots + \mu_m v_m\end{aligned}$$

und damit ist $v_k \in \mathrm{L}(v_1, \ldots, v_{k-1}, v_{k+1}, \ldots, v_m)$. □

Lemma 3.12 *Seien $v_1, \ldots, v_m \in V$ linear unabhängig und sei $u \in V$. Dann sind $v_1, \ldots, v_m, u$ linear unabhängig genau, wenn $u \notin \mathrm{L}(v_1, \ldots, v_m)$.*

Beweis Sind $v_1, \ldots, v_m, u$ linear unabhängig, dann gilt nach Satz 3.4 (2), dass insbesondere $u \notin \mathrm{L}(v_1, \ldots, v_m)$. Es bleibt also zu zeigen, dass $u \in \mathrm{L}(v_1, \ldots, v_m)$, wenn $v_1, \ldots, v_m, u$ linear abhängig sind.

Sind die Vektoren $v_1, \ldots, v_m, u$ linear abhängig, dann gibt es $\lambda_1, \ldots, \lambda_{m+1} \in K$ mit $\lambda_j \neq 0$ für mindestens ein j, so dass $\lambda_1 v_1 + \cdots + \lambda_m v_m + \lambda_{m+1} u = 0$. Nehme an, dass $\lambda_{m+1} = 0$; dann ist $\lambda_j \neq 0$ für mindestens ein j mit $1 \leq j \leq m$ und $\lambda_1 v_1 + \cdots + \lambda_m v_m = 0$. Dies steht aber im Widerspruch zu der Annahme, dass $v_1, \ldots, v_m$ linear unabhängig sind, und daraus ergibt sich, dass $\lambda_{m+1} \neq 0$. Für $j = 1, \ldots, m$ setze $\mu_j = \nu \lambda_j$, wobei $\nu = (-1)\lambda_{m+1}^{-1}$. Dann ist $\nu \lambda_{m+1} = -1$ und

$$\begin{aligned}u = 0 + u &= \nu 0 + u = \nu(\lambda_1 v_1 + \cdots + \lambda_m v_m + \lambda_{m+1} u) + u \\ &= (\nu\lambda_1) v_1 + \cdots + (\nu\lambda_m) v_m + (\nu\lambda_{m+1}) u + u \\ &= \mu_1 v_1 + \cdots + \mu_m v_m + (-1) u + u \\ &= \mu_1 v_1 + \cdots + \mu_m v_m + 0 = \mu_1 v_1 + \cdots + \mu_m v_m \,;\end{aligned}$$

d.h., $u \in \mathrm{L}(v_1, \ldots, v_m)$. □

Lemma 3.13 *Die Vektoren $v_1, \ldots, v_m$ sind linear unabhängig genau dann, wenn $v_1 \neq 0$ und $v_n \notin \mathrm{L}(v_1, \ldots, v_{n-1})$ für jedes $n = 2, \ldots, m$.*

Beweis Sind $v_1, \ldots, v_m$ linear unabhängig, dann ist nach Lemma 3.9 $v_1 \neq 0$ und nach Satz 3.4 (2) gilt $v_n \notin \mathrm{L}(v_1, \ldots, v_{n-1})$ für jedes $n = 2, \ldots, m$. Nehme nun umgekehrt an, dass die Vektoren $v_1, \ldots, v_m$ linear abhängig sind. Dann gibt es $\lambda_1, \ldots, \lambda_m \in K$ mit $\lambda_j \neq 0$ für mindestens ein j, so dass $\lambda_1 v_1 + \cdots + \lambda_m v_m = 0$. Sei $n = \max\{1 \leq j \leq m : \lambda_j \neq 0\}$; also ist $\lambda_n \neq 0$ und $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$. Falls $n = 1$, so ist $\lambda_1 \neq 0$ und $\lambda_1 v_1 = 0$, und hier ist nach Lemma 3.2, $v_1 = 0$. Wenn aber $n \geq 2$, so ist $v_n = \mu_1 v_1 + \cdots + \mu_{n-1} v_{n-1}$, wobei $\mu_j = (-1)\lambda_n^{-1}\lambda_j$ für $j = 1, \ldots, n-1$, und hier ist $v_n \in \mathrm{L}(v_1, \ldots, v_{n-1})$. Daraus ergibt sich, dass $v_1, \ldots, v_m$ linear unabhängig sein müssen, wenn $v_1 \neq 0$ und $v_n \notin \mathrm{L}(v_1, \ldots, v_{n-1})$ für jedes $n = 2, \ldots, m$. □

Für den Vektorraum K^m gibt es einen engen Zusammenhang zwischen der Frage, ob Vektoren $v_1, \ldots, v_n \in K^m$ linear unabhängig sind, und der Frage, ob ein entsprechendes homogenes Gleichungssystem eindeutig lösbar ist.

Seien $v_1, \ldots, v_n \in K^m$ (mit $n \geq 1$) mit $v_j = (\alpha_{1j}, \ldots, \alpha_{mj})$ für jedes $j = 1, \ldots, n$. Es gibt dann die Matrix $A = (\alpha_{ij}) \in \mathrm{M}(m \times n, K)$, in der die Vektoren $v_1, \ldots, v_n$ als Spalten vorkommen.

Satz 3.5 *Die Vektoren $v_1, \ldots, v_n$ sind linear unabhängig genau dann, wenn das zu A und 0 gehörige lineare Gleichungssystem eindeutig lösbar ist (d.h., genau dann, wenn $\mathrm{Lös}(A, 0) = \{0\}$).*

Beweis Genau wie im Beweis für Satz 3.3 gilt $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$ genau dann, wenn $(\lambda_1, \ldots, \lambda_n) \in \mathrm{Lös}(A, 0)$. Damit sind $v_1, \ldots, v_n$ linear unabhängig genau dann, wenn $\mathrm{Lös}(A, v) = \{0\}$. □

Für Untervektorräume $U_1, \ldots, U_m$ ($m \geq 2$) von V bezeichne mit $U_1 + \cdots + U_m$ die Menge aller Elemente von V, die eine Darstellung der Form $u_1 + \cdots + u_m$ haben mit $u_j \in U_j$ für $j = 1, \ldots, m$, und nenne $U_1 + \cdots + U_m$ die *Summe* der Untervektorräume $U_1, \ldots, U_m$.

Satz 3.6 *(1) Die Summe $U_1 + \cdots + U_m$ der Untervektorräume $U_1, \ldots, U_m$ ist ein Untervektorraum von V und $U_j \subset U_1 + \cdots + U_m$ für jedes $j = 1, \ldots, m$.*

(2) Ist U ein Untervektorraum von V mit $U_j \subset U$ für jedes $j = 1, \ldots, m$, so ist $U_1 + \cdots + U_m \subset U$.

Beweis Übung. □

Nach Satz 3.6 ist die Summe $U_1 + \cdots + U_m$ der kleinste Untervektorraum von V, der die Untervektorräume $U_1, \ldots, U_m$ enthält.

4 Endlichdimensionale Vektorräume

Im folgenden sei wieder K ein Körper und V ein Vektorraum über K.

Sei U ein Untervektorraum von V und seien $v_1, \ldots, v_m \in V$ (mit $m \geq 1$). Dann heißt das m-Tupel $(v_1, \ldots, v_m)$ *Basis* von U, wenn gilt:

(B1) Die Vektoren $v_1, \ldots, v_m$ sind linear unabhängig,

(B2) $\mathrm{L}(v_1, \ldots, v_m) = U$.

Sei $(v_1, \ldots, v_m)$ eine Basis eines Untervektorraumes U von V. Nach Lemma 3.9 ist dann $v_j \neq 0$ für jedes j, und daraus folgt nach Lemma 3.4, dass $U \neq \{0\}$. Der triviale Untervektorraum $\{0\}$ von V kann also im obigen Sinn keine Basis besitzen. Es erweist sich aber als nützlich, die leere Menge $\varnothing$ als Basis von $\{0\}$ anzusehen.

Wichtiges Beispiel für den Vektorraum K^m: Für $j = 1, \ldots, m$ sei wieder

$$e_j = (0, \ldots, 0, 1, 0, \ldots, 0)$$

mit der Eins in der j-ten Komponenten. Dann sind $e_1, \ldots, e_m$ linear unabhängig und es gilt auch $\mathrm{L}(e_1, \ldots, e_m) = K^m$. Damit ist $(e_1, \ldots, e_m)$ eine Basis von K^m. Sie wird *die kanonische Basis* von K^m genannt.

Lemma 4.1 *Sei $U \neq \{0\}$ ein Untervektorraum von V und sei $(v_1, \ldots, v_m)$ eine Basis von U.*

(1) Seien $\lambda_1, \ldots, \lambda_m \in K$ mit $\lambda_j \neq 0$ für jedes j; dann ist $(\lambda_1 v_1, \ldots, \lambda_m v_m)$ auch eine Basis von U.

(2) Sei $\{i_1, \ldots, i_m\}$ eine Permutation von $\{1, \ldots, m\}$; dann ist $(v_{i_1}, \ldots, v_{i_m})$ auch eine Basis von U.

Beweis (1) Nach Lemma 3.5 ist $\mathrm{L}(\lambda_1 v_1, \ldots, \lambda_m v_m) = U$ und nach Lemma 3.10 sind $\lambda_1 v_1, \ldots, \lambda_m v_m$ linear unabhängig. Damit ist $(\lambda_1 v_1, \ldots, \lambda_m v_m)$ eine Basis von U.

(2) Nach Lemma 3.11 sind $v_{i_1}, \ldots, v_{i_m}$ linear unabhängig und nach Lemma 3.6 ist $\mathrm{L}(v_{i_1}, \ldots, v_{i_m}) = U$. Damit ist $(v_{i_1}, \ldots, v_{i_m})$ eine Basis von U. □

Satz 4.1 *Sei $U \neq \{0\}$ ein Untervektorraum von V und sei $(v_1, \ldots, v_m)$ eine Basis von U. Dann gibt es zu jedem $u \in U$ genau ein $(\lambda_1, \ldots, \lambda_m) \in K^m$, so dass*

$$u = \lambda_1 v_1 + \cdots + \lambda_m v_m \, .$$

Beweis Sei $u \in U$; da $U = \mathrm{L}(v_1, \ldots, v_m)$, gibt es mindestens ein m-Tupel $(\lambda_1, \ldots, \lambda_m) \in K^m$, so dass $u = \lambda_1 v_1 + \cdots + \lambda_m v_m$. Nehme an, dass auch $u = \mu_1 v_1 + \cdots + \mu_m v_m$ für ein $(\mu_1, \ldots, \mu_m) \in K^m$. Dann ist

$$\begin{aligned}((-1)\lambda_1 + \mu_1)v_1 + \cdots + ((-1)\lambda_m + \mu_m)v_m \\ &= (-1)(\lambda_1 v_1 + \cdots + \lambda_m v_m) + \mu_1 v_1 + \cdots + \mu_m v_m \\ &= (-1)u + u = (-u) + u = 0 \,,\end{aligned}$$

und daraus folgt, dass $(-1)\lambda_j + \mu_j = 0$ für jedes $j = 1, \ldots, m$, da $v_1, \ldots, v_m$ linear unabhängig sind. Damit ist $\lambda_j = \mu_j$ für jedes $j = 1, \ldots, m$. (*Übung:* Sind $\lambda, \mu \in K$ mit $(-1)\lambda + \mu = 0$, so ist $\lambda = \mu$.) Dies zeigt also, dass es genau ein $(\lambda_1, \ldots, \lambda_m) \in K^m$ mit $u = \lambda_1 v_1 + \cdots + \lambda_m v_m$ gibt. □

Satz 4.2 *Seien $(v_1, \ldots, v_m)$, $(w_1, \ldots, w_n)$ zwei Basen eines Untervektorraumes U von V mit $U \neq \{0\}$. Dann ist $m = n$.*

Beweis Für den Beweis wird folgendes *Austauschlemma* gebraucht:

Lemma 4.2 *Sei $U \neq \{0\}$ ein Untervektorraum von V und seien $(v_1, \ldots, v_m)$, $(w_1, \ldots, w_n)$ zwei Basen von U mit $m \geq 2$ und $n \geq 1$. Dann gibt es zu jedem $j = 1, \ldots, m$ ein k mit $1 \leq k \leq n$, so dass $(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_m, w_k)$ wieder eine Basis von U ist.*

Beweis Sei j mit $1 \leq j \leq m$ fest. Nehme an, $w_k \in \mathrm{L}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_m)$ für jedes $k = 1, \ldots, n$. Dann gilt nach (B2) und Satz 3.2, dass

$$U = \mathrm{L}(w_1, \ldots, w_n) \subset \mathrm{L}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_m) \,,$$

und nach Satz 3.1 ist $v_j \in U$; d.h., $v_j \in \mathrm{L}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_m)$. Dies steht aber im Widerspruch zu Satz 3.4 (2), da $v_1, \ldots, v_m$ linear unabhängig sind. Damit gibt es ein k mit $1 \leq k \leq n$, so dass $w_k \notin \mathrm{L}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_m)$. Es wird nun gezeigt, dass $\{v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_m, w_k\}$ eine Basis von U ist.

Nach Lemma 3.11 sind $v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_m$ linear unabhängig, und folglich sind nach Lemma 3.12 $v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_m, w_k$ auch linear unabhängig.

Es bleibt zu zeigen, dass $\mathrm{L}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_m, w_k) = U$. Nach (B2), Satz 3.1 und Lemma 3.6 ist

$$w_k \in U = \mathrm{L}(v_1, \ldots, v_m) = \mathrm{L}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_m, v_j) \,.$$

Aber $w_k \notin \mathrm{L}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_m)$, d.h.,

$$w_k \in \mathrm{L}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_m, v_j) \setminus \mathrm{L}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_m) \,,$$

und daraus folgt nach Lemma 3.8, dass $v_j \in \mathrm{L}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_m, w_k)$. Nach Satz 3.2 ist also $U = \mathrm{L}(v_1, \ldots, v_m) \subset \mathrm{L}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_m, w_k)$; und daher ist $U = \mathrm{L}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_m, w_k)$. □

Nun zum Beweis für Satz 4.2, und ohne Beschränkung der Allgemeinheit kann man annehmen, dass $m \geq n \geq 1$. Ist $m = 1$, dann ist $m = n$ trivial richtig; es kann also weiter angenommen werden, dass $m \geq 2$. Nun wird Lemma 4.2 wiederholt angewendet.

Zuerst gibt es nach Lemma 4.2 ein k_1 mit $1 \leq k_1 \leq n$, so dass $(v_2, \ldots, v_m, w_{k_1})$ eine Basis von U ist. Nun sind $(v_2, \ldots, v_m, w_{k_1})$ und $(w_1, \ldots, w_n)$ Basen von U; also gibt es nach Lemma 4.2 ein k_2 mit $1 \leq k_2 \leq n$, so dass $(v_3, \ldots, v_m, w_{k_1}, w_{k_2})$ eine Basis von U ist. Nach m-maliger Anwendung dieses Verfahrens erhält man eine Basis $(w_{k_1}, \ldots, w_{k_m})$ von U. Insbesondere sind dann $w_{k_1}, \ldots, w_{k_m}$ linear unabhängig und daraus folgt nach Lemma 3.9, dass $w_{k_i} \neq w_{k_j}$ und damit $k_i \neq k_j$, falls $i \neq j$. Da aber $1 \leq k_j \leq n$ für jedes $j = 1, \ldots, m$, ist dies nur möglich, wenn $m = n$. □

Ein Untervektorraum U von V heißt nun *endlichdimensional*, wenn U eine Basis besitzt. Sei $U \neq \{0\}$ endlichdimensional; nach Satz 4.2 gibt es dann $n \geq 1$, so dass jede Basis von U aus genau n Elementen besteht und diese Zahl n wird die *Dimension* von U gennant und wird mit $\dim U$ abgekürzt. Per Definition wird der Untervektorraum $\{0\}$ die Dimension 0 zugeordnet ($\dim\{0\} = 0$).

Der Vektorraum V selber heißt *endlichdimensional*, wenn V eine Basis besitzt (d.h., wenn V als Untervektorraum von V endlichdimensional ist). Dann ist die *Dimension* von V die Anzahl der Elemente in einer Basis von V.

Für jedes $m \geq 1$ ist der Vektorraum K^m endlichdimensional und $\dim K^m = m$, da $(e_1, \ldots, e_m)$ eine Basis von K^m ist.

Nicht jeder Vektorraum ist endlichdimensional: Zum Beispiel ist der Vektorraum $\mathrm{Abb}(X, K)$ genau dann endlichdimensional, wenn die Menge X endlich ist. (Der Beweis dafür ist eine Übung.) In der Vorlesung *Linearer Algebra* werden aber hauptsächlich nur endlichdimensionale Vektorräume betrachtet.

Lemma 4.3 *Seien $m \geq 0$, $n \geq 1$, seien $u_1, \ldots, u_m \in V$ linear unabhängig und $w_1, \ldots, w_n \in V$. Dann gibt es $i_1, i_2, \ldots, i_k$ mit $k \geq 0$ und $1 \leq i_1 < \cdots < i_k \leq n$, so dass $u_1, \ldots, u_m, w_{i_1}, \ldots, w_{i_k}$ linear unabhängig sind und*

$$\mathrm{L}(u_1, \ldots, u_m, w_{i_1}, \ldots, w_{i_k}) = \mathrm{L}(u_1, \ldots, u_m, w_1, \ldots, w_n) \,.$$

(In dieser Formulierung werden $v_1, \ldots, v_p$ als linear unabhängig angesehen und $\mathrm{L}(v_1, \ldots, v_p)$ als $\{0\}$ interpretiert, wenn $p = 0$.)

Beweis Sei $k \geq 0$ die größte Zahl, für die es $1 \leq i_1 < \cdots < i_k \leq n$ gibt, so dass $u_1, \ldots, u_m, w_{i_1}, \ldots, w_{i_k}$ linear unabhängig sind. Seien $i_1, \ldots, i_k$ wie in der Definition von k. Es wird nun gezeigt, dass

$$\mathrm{L}(u_1, \ldots, u_m, w_{i_1}, \ldots, w_{i_k}) = \mathrm{L}(u_1, \ldots, u_m, w_1, \ldots, w_n) \; .$$

Sei j mit $j \neq i_p$ für jedes $p = 1, \ldots, n$. Nach Lemma 3.11 sind dann die Vektoren $u_1, \ldots, u_m, w_{i_1}, \ldots, w_{i_k}, w_j$ linear abhängig und daraus folgt nach Lemma 3.12, dass $w_j \in \mathrm{L}(u_1, \ldots, u_m, w_{i_1}, \ldots, w_{i_k})$. Also ist $v \in \mathrm{L}(u_1, \ldots, u_m, w_{i_1}, \ldots, w_{i_k})$ für jedes $v \in \{u_1, \ldots, u_m, w_1, \ldots, w_n\}$ und folglich gilt nach Satz 3.2, dass

$$\mathrm{L}(u_1, \ldots, u_m, w_1, \ldots, w_n) \subset \mathrm{L}(u_1, \ldots, u_m, w_{i_1}, \ldots, w_{i_k}) \; ,$$

d.h., $\mathrm{L}(u_1, \ldots, u_m, w_1, \ldots, w_n) = \mathrm{L}(u_1, \ldots, u_m, w_{i_1}, \ldots, w_{i_k})$. □

Satz 4.3 *Ein Untervektorraum U von V ist endlichdimensional genau dann, wenn es ein $n \geq 1$ und $v_1, \ldots, v_n \in U$ gibt, so dass $U = \mathrm{L}(v_1, \ldots, v_n)$. Ist ferner $U \neq \{0\}$ und $U = \mathrm{L}(v_1, \ldots, v_n)$, dann gibt es $i_1, \ldots, i_k$ mit $k \geq 1$ und $1 \leq i_1 < \cdots < i_k \leq n$, so dass $(v_{i_1}, v_{i_2}, \ldots, v_{i_k})$ eine Basis von U ist.*

Beweis Da $\{0\}$ endlichdimensional ist und $\{0\} = L(0)$, kann man annehmen, dass $U \neq \{0\}$. Ist U endlichdimensional, dann gilt $U = \mathrm{L}(v_1, \ldots, v_n)$ für jede Basis $(v_1, \ldots, v_n)$ von U. Nehme also umgekehrt an, dass es $n \geq 1$ und $v_1, \ldots, v_n \in U$ mit $U = \mathrm{L}(v_1, \ldots, v_n)$ gibt. Nach Lemma 4.3 (mit $m = 0$) gibt es dann $k \geq 0$ und $1 \leq i_1 < \cdots < i_k \leq n$, so dass $v_{i_1}, v_{i_2}, \ldots, v_{i_k}$ linear unabhängig sind und $\mathrm{L}(v_{i_1}, v_{i_2}, \ldots, v_{i_k}) = \mathrm{L}(v_1, \ldots, v_n) = U$ (und also ist $k \geq 1$, da $U \neq \{0\}$). Damit ist $(v_{i_1}, v_{i_2}, \ldots, v_{i_k})$ eine Basis von U; insbesondere ist U endlichdimensional. □

Lemma 4.4 *Sei W ein endlichdimensionaler Untervektorraum von V und seien $u_1, \ldots, u_m \in W$ linear unabhängig. Dann gilt $m \leq \dim W$.*

Beweis Sei $(w_1, \ldots, w_n)$ eine Basis von W. Nach Lemma 4.3 gibt es $i_1, \ldots, i_k$ mit $k \geq 0$ und $1 \leq i_1 < \cdots < i_k \leq n$, so dass $u_1, \ldots, u_m, w_{i_1}, \ldots, w_{i_k}$ linear unabhängig sind und $\mathrm{L}(u_1, \ldots, u_m, w_{i_1}, \ldots, w_{i_k}) = \mathrm{L}(u_1, \ldots, u_m, w_1, \ldots, w_n)$. Da aber $u_j \in W$ für jedes $j = 1, \ldots, m$ und $W = \mathrm{L}(w_1, \ldots, w_n)$, ist nach Satz 3.1 und Satz 3.2 $\mathrm{L}(u_1, \ldots, u_m, w_1, \ldots, w_n) = W$. Damit ist $(u_1, \ldots, u_m, w_{i_1}, \ldots, w_{i_k})$ eine Basis von W. Insbesondere ist $m \leq n = \dim W$. □

Satz 4.4 *Seien U, W Untervektorräume von V mit W endlichdimensional und $U \subset W$. Dann ist U endlichdimensional und $\dim U \leq \dim W$.*

Beweis Man kann annehmen, dass $U \neq \{0\}$ (und damit ist auch $W \neq \{0\}$), da die Behauptung trivial richtig ist, wenn $U = \{0\}$.

Da $U \neq \{0\}$, gibt es $u_1 \in U$ mit $u_1 \neq 0$ und nach Satz 3.4 (1) ist u_1 linear unabhängig. Gilt $U = \mathrm{L}(u_1)$, dann ist (u_1) schon eine Basis von U.

Ist dagegen $U \neq \mathrm{L}(u_1)$, so gibt es $u_2 \in U \setminus \mathrm{L}(u_1)$ und nach Lemma 3.12 sind dann u_1, u_2 linear unabhängig. Gilt $U = \mathrm{L}(u_1, u_2)$, dann ist (u_1, u_2) eine Basis von U.

Ist dagegen $U \neq \mathrm{L}(u_1, u_2)$, so gibt es $u_3 \in U \setminus \mathrm{L}(u_1, u_2)$ und nach Lemma 3.12 sind dann u_1, u_2, u_3 linear unabhängig. Gilt $U = \mathrm{L}(u_1, u_2, u_3)$, dann ist (u_1, u_2, u_3) eine Basis von U.

Ist dagegen $U \neq \mathrm{L}(u_1, u_2, u_3)$, so gibt es $u_4 \in U \setminus \mathrm{L}(u_1, u_2, u_3)$ und ... .

Entweder hört dieses Verfahren auf, indem $U = \mathrm{L}(v_1, \ldots, v_m)$ für ein $m \geq 1$, und damit ist $(v_1, \ldots, v_m)$ eine Basis von U, oder das Verfahren hört nie auf, und in diesem Fall ist $\{v_k\}_{k\geq 1}$ eine Folge von Elementen aus U mit der Eigenschaft, dass für jedes $m \geq 1$ die Vektoren $v_1, \ldots, v_m$ linear unabhängig sind. Aber nach Lemma 4.4 ist das Letztere nicht möglich und daraus ergibt sich, dass U endlichdimensional ist. Ist ferner $(u_1, \ldots, u_m)$ eine Basis von U, dann gilt nach Lemma 4.4, dass $m \leq \dim W$, d.h., $\dim U \leq \dim W$. □

Satz 4.5 (Basisergänzungssatz) *Seien U, W Untervektorräume von V mit W endlichdimensional und $\{0\} \neq U \subset W$; sei $(u_1, \ldots, u_m)$ eine Basis von U. Dann gibt es $k \geq 0$ und $v_1, \ldots, v_k \in W$, so dass $(u_1, \ldots, u_m, v_1, \ldots, v_k)$ eine Basis von W ist.*

Beweis Sei $(w_1, \ldots, w_n)$ eine Basis von W. Nach Lemma 4.3 gibt es $i_1, \ldots, i_k$ mit $k \geq 0$ und $1 \leq i_1 < \cdots < i_k \leq n$, so dass $u_1, \ldots, u_m, w_{i_1}, \ldots, w_{i_k}$ linear unabhängig sind und $\mathrm{L}(u_1, \ldots, u_m, w_{i_1}, \ldots, w_{i_k}) = \mathrm{L}(u_1, \ldots, u_m, w_1, \ldots, w_n)$. Wie im Beweis für Lemma 4.4 ist nun $(u_1, \ldots, u_m, w_{i_1}, \ldots, w_{i_k})$ eine Basis von W. □

Satz 4.6 *Sei $U \neq \{0\}$ ein endlichdimensionaler Untervektorraum von V und seien $u_1, \ldots, u_m \in U$ (mit $m \geq 1$).*

(1) Sind $u_1, \ldots, u_m$ linear unabhängig, so ist $m \leq \dim U$.

(2) Sind $u_1, \ldots, u_m$ linear unabhängig und $m = \dim U$, so ist $(u_1, \ldots, u_m)$ schon eine Basis von U.

(3) Gilt $\mathrm{L}(u_1, \ldots, u_m) = U$, so ist $m \geq \dim U$.

(4) Gilt $\mathrm{L}(u_1, \ldots, u_m) = U$ und $m = \dim U$, so ist $(u_1, \ldots, u_m)$ schon eine Basis von U.

Beweis (1) und (2): Setze $U' = \mathrm{L}(u_1, \ldots, u_m)$, und also ist $(u_1, \ldots, u_m)$ eine Basis von U'. Nach Satz 4.5 gibt es dann $k \geq 0$ und $v_1, \ldots, v_k \in U$, so dass $(u_1, \ldots, u_m, v_1, \ldots, v_k)$ eine Basis von U ist. Insbesondere ist $m \leq m+k = \dim U$. Ist ferner $m = \dim U$, d.h., $k = 0$, so ist $(u_1, \ldots, u_m)$ eine Basis von U.

(3) und (4): Wie im Beweis für Satz 4.3 gibt es $k \geq 1$ und $1 \leq i_1 < \cdots < i_k \leq m$, so dass $(v_{i_1}, v_{i_2}, \ldots, v_{i_k})$ eine Basis von U ist. Insbesondere ist $\dim U = k \leq m$. Ist ferner $m = \dim U$, d.h., $k = m$, so ist $i_j = j$ für jedes $j = 1, \ldots, m$ und damit ist $(v_1, v_2, \ldots, v_m) = (v_{i_1}, v_{i_2}, \ldots, v_{i_k})$ eine Basis von U. □

Satz 4.7 *Seien U, W Untervektorräume von V mit W endlichdimensional und $U \subset W$. Dann ist $\dim U = \dim W$ gleichbedeutend mit $U = W$.*

Beweis Dies ist klar, wenn $U = \{0\}$ und also kann man annehmen, dass $U \neq \{0\}$. Sei $(u_1, \ldots, u_m)$ eine Basis von U; dann sind $u_1, \ldots, u_m \in W$ linear unabhängig. Ist $m = \dim U = \dim W$, so ergibt sich nach Satz 4.6 (2), dass $(u_1, \ldots, u_m)$ auch eine Basis von W ist, und insbesondere ist $U = \mathrm{L}(u_1, \ldots, u_m) = W$. □

Erinnerung: Für Untervektorräume $U_1, \ldots, U_m$ $(m \geq 2)$ von V ist die Summe $U_1 + \cdots + U_m$ der Untervektorraum bestehend aus allen Elementen von V, die eine Darstellung der Form $u_1 + \cdots + u_m$ haben mit $u_j \in U_j$ für $j = 1, \ldots, m$.

Sind U und W Untervektorräume von V, dann ist es klar, dass $U \cap W$ auch ein Untervektorraum von V ist.

Satz 4.8 *Seien U und W endlichdimensionale Untervektorräume von V. Dann sind die Untervektorräume $U + W$ und $U \cap W$ auch endlichdimensional und*

$$\dim(U + W) + \dim(U \cap W) = \dim U + \dim W \, .$$

Beweis Nach Satz 4.4 ist $U \cap W$ endlichdimensional, da U endlichdimensional ist und $U \cap W \subset U$. Sei $(v_1, \ldots, v_m)$ eine Basis von $U \cap W$, (wobei $m = 0$, falls $U \cap W = \{0\}$). Da $U \cap W \subset U$ und $U \cap W \subset W$, gibt es nach Satz 4.6 Vektoren $u_1, \ldots, u_p \in U$ (mit $p \geq 0$) und Vektoren $w_1, \ldots, w_q \in W$ (mit $q \geq 0$), so dass $(v_1, \ldots, v_m, u_1, \ldots, u_p)$ eine Basis von U und $(v_1, \ldots, v_m, w_1, \ldots, w_q)$ eine Basis von W ist. Dann ist $\dim(U \cap W) = m$, $\dim U = m + p$ und $\dim W = m + q$, und also genügt es zu zeigen, dass $(v_1, \ldots, v_m, u_1, \ldots, u_p, w_1, \ldots, w_q)$ eine Basis von $U + W$ ist.

Sei $v \in U + W$; dann gibt es $u \in U$ und $w \in W$, so dass $v = u + w$. Ferner gibt es $\lambda_1, \ldots, \lambda_m, \mu_1, \ldots, \mu_p \in K$ mit $u = \lambda_1 v_1 + \cdots + \lambda_m v_m + \mu_1 u_1 + \cdots + \mu_p u_p$, da $(v_1, \ldots, v_m, u_1, \ldots, u_p)$ eine Basis von U ist, und $\lambda'_1, \ldots, \lambda'_m, \nu_1, \ldots, \nu_q \in K$

mit $w = \lambda'_1 v_1 + \cdots + \lambda'_m v_m + \nu_1 w_1 + \cdots + \nu_q w_q$, da $(v_1, \ldots, v_m, w_1, \ldots, w_q)$ eine Basis von W ist. Daraus folgt, dass

$$v = u + w = \lambda''_1 v_1 + \cdots + \lambda''_m v_m + \mu_1 u_1 + \cdots + \mu_p u_p + \nu_1 w_1 + \cdots + \nu_q w_q \ ,$$

wobei $\lambda''_j = \lambda_j + \lambda'_j$ für jedes j, d.h., $v \in \mathrm{L}(v_1, \ldots, v_m, u_1, \ldots, u_p, w_1, \ldots, w_q)$. Damit ist $U + W \subset \mathrm{L}(v_1, \ldots, v_m, u_1, \ldots, u_p, w_1, \ldots, w_q)$. Andererseits folgt

$$\mathrm{L}(v_1, \ldots, v_m, u_1, \ldots, u_p, w_1, \ldots, w_q) \subset U + W$$

unmittelbar aus Satz 3.2, da $v_1, \ldots, v_m, u_1, \ldots, u_p, w_1, \ldots, w_q \in U + W$, d.h., es gilt $\mathrm{L}(v_1, \ldots, v_m, u_1, \ldots, u_p, w_1, \ldots, w_q) = U + W$.

Es wird nun gezeigt, dass $v_1, \ldots, v_m, u_1, \ldots, u_p, w_1, \ldots, w_q$ linear unabhängig sind. Seien $\lambda_1, \ldots, \lambda_m, \mu_1, \ldots, \mu_p, \nu_1, \ldots, \nu_q \in K$ mit

$$\lambda_1 v_1 + \cdots + \lambda_m v_m + \mu_1 u_1 + \cdots + \mu_p u_p + \nu_1 w_1 + \cdots + \nu_q w_q = 0 \ .$$

Setze $u = \lambda_1 v_1 + \cdots + \lambda_m v_m + \mu_1 u_1 + \cdots + \mu_p u_p$ und $w = \nu_1 w_1 + \cdots + \nu_q w_q$. Dann ist $u \in \mathrm{L}(v_1, \ldots, v_m, u_1, \ldots, u_p) = U$ und $w \in \mathrm{L}(w_1, \ldots, w_q) = W$. Aber $u + w = 0$, d.h., $w = -u = (-1)u$ und damit ist auch $w \in U$, d.h., $w \in U \cap W$. Da $(v_1, \ldots, v_m)$ eine Basis von $U \cap W$ ist, gibt es $\lambda'_1, \ldots, \lambda'_m \in K$, so dass $w = \lambda'_1 v_1 + \cdots + \lambda'_m v_m$. Es gilt also

$$\begin{aligned} 0 &= u + w \\ &= \lambda_1 v_1 + \cdots + \lambda_m v_m + \mu_1 u_1 + \cdots + \mu_p u_p + \lambda'_1 v_1 + \cdots + \lambda'_m v_m \\ &= (\lambda_1 + \lambda'_1) v_1 + \cdots + (\lambda_m + \lambda'_m) v_m + \mu_1 u_1 + \cdots + \mu_p u_p \end{aligned}$$

und $v_1, \ldots, v_m, u_1, \ldots, u_p$ sind linear unabhängig. Insbesondere ist dann $\mu_i = 0$ für jedes $i = 1, \ldots, p$ und folglich ist $\lambda_1 v_1 + \cdots + \lambda_m v_m + \nu_1 w_1 + \cdots + \nu_q w_q = 0$. Aber $v_1, \ldots, v_m, w_1, \ldots, w_q$ sind linear unabhängig und daraus ergibt sich, dass auch $\lambda_j = 0$ für $j = 1, \ldots, m$ und $\nu_k = 0$ für $k = 1, \ldots, q$. Dies zeigt, dass $v_1, \ldots, v_m, u_1, \ldots, u_p, w_1, \ldots, w_q$ linear unabhängig sind. □

Sei $A = (a_{ij}) \in \mathrm{M}(m \times n, K)$ eine $m \times n$ Matrix über K. Seien $v_1, \ldots, v_n \in K^m$ die Spalten von A; es gilt also $v_j = (a_{1j}, \ldots, a_{mj})$ für jedes $j = 1, \ldots, n$.

Es wird nun gezeigt, wie man eine Basis von $\mathrm{Lös}(A, 0)$ konstruieren kann. Wenn $A = 0$, dann ist $\mathrm{Lös}(A, 0) = K^n$, und in diesem Fall ist jede Basis von K^n auch eine Basis von $\mathrm{Lös}(A, 0)$. Es kann also angenommen werden, dass $A \neq 0$. Nach Satz 4.3 gibt es dann $k_1, \ldots, k_q$ mit $q \geq 1$ und $1 \leq k_1 < \cdots < k_q \leq n$, so dass $(v_{k_1}, \ldots, v_{k_q})$ eine Basis von $\mathrm{L}(v_1, \ldots, v_n)$ ist.

Seien $\ell_1, \ldots, \ell_p$ mit $1 \leq \ell_1 < \cdots < \ell_p \leq n$ die restlichen Indizes (d.h., die Indizes, die nicht in der Menge $\{k_1, \ldots, k_q\}$ vorkommen); insbesondere ist dann $p = n - q = n - \dim \mathrm{L}(v_1, \ldots, v_n)$.

Lemma 4.5 *Seien $\mu_1, \ldots, \mu_p \in K$; dann gibt es genau eine Lösung $(\lambda_1, \ldots, \lambda_n)$ in* $\mathrm{Lös}(A, 0)$ *mit $\lambda_{\ell_i} = \mu_i$ für jedes $i = 1, \ldots, p$.*

Beweis Sei $v = \mu_1 v_{\ell_1} + \cdots + \mu_p v_{\ell_p}$; dann ist $v \in \mathrm{L}(v_1, \ldots, v_n)$ und folglich gibt es $\nu_1, \ldots, \nu_q \in K$, so dass $v = \nu_1 v_{k_1} + \cdots + \nu_q v_{k_q}$, da $(v_{k_1}, \ldots, v_{k_q})$ eine Basis von $\mathrm{L}(v_1, \ldots, v_n)$ ist. Definiere ein Element $(\lambda_1, \ldots, \lambda_n)$ von K^n durch

$$\lambda_m = \begin{cases} \mu_i & \text{falls } m = \ell_i \text{ für ein } i\,, \\ -\nu_j & \text{falls } m = k_j \text{ für ein } j\,. \end{cases}$$

Dann gilt $\lambda_{\ell_i} = \mu_i$ für jedes $i = 1, \ldots, p$ und

$$\begin{aligned} \lambda_1 v_1 + \cdots + \lambda_n v_n &= \lambda_{\ell_1} v_{\ell_1} + \cdots + \lambda_{\ell_p} v_{\ell_p} + \lambda_{k_1} v_{k_1} + \cdots + \lambda_{k_q} v_{k_q} \\ &= \mu_1 v_{\ell_1} + \cdots + \mu_p v_{\ell_p} + (-\nu_1) v_{k_1} + \cdots + (-\nu_q) v_{k_q} \\ &= v - (\nu_1 v_{k_1} + \cdots + \nu_q v_{k_q}) = v - v = 0\,, \end{aligned}$$

und damit ist $(\lambda_1, \ldots, \lambda_n) \in \mathrm{Lös}(A, 0)$.

Sei nun $(\lambda'_1, \ldots, \lambda'_n)$ ein beliebiges Element von $\mathrm{Lös}(A, 0)$ mit $\lambda'_{\ell_i} = \mu_i$ für jedes $i = 1, \ldots, p$. Dann gilt

$$\begin{aligned} 0 = 0 - 0 &= (\lambda_1 v_1 + \cdots + \lambda_n v_n) - (\lambda'_1 v_1 + \cdots + \lambda'_n v_n) \\ &= (\lambda_1 - \lambda'_1) v_1 + \cdots + (\lambda_n - \lambda'_n) v_n \\ &= (\lambda_{\ell_1} - \lambda'_{\ell_1}) v_{\ell_1} + \cdots + (\lambda_{\ell_p} - \lambda'_{\ell_p}) v_{\ell_p} + (\lambda_{k_1} - \lambda'_{k_1}) v_{k_1} + \cdots + (\lambda_{k_q} - \lambda'_{k_q}) v_{k_q} \\ &= (\lambda_{k_1} - \lambda'_{k_1}) v_{k_1} + \cdots + (\lambda_{k_q} - \lambda'_{k_q}) v_{k_q}\,, \end{aligned}$$

da $\lambda_{\ell_i} = \lambda'_{\ell_i}$ für $i = 1, \ldots, p$. Daraus ergibt sich, dass $\lambda_{k_j} = \lambda'_{k_j}$ für $j = 1, \ldots, q$, da $v_{k_1}, \ldots, v_{k_q}$ linear unabhängig sind. Damit ist $(\lambda_1, \ldots, \lambda_n) = (\lambda'_1, \ldots, \lambda'_n)$; d.h., es gibt genau ein $(\lambda_1, \ldots, \lambda_n) \in \mathrm{Lös}(A, 0)$ mit $\lambda_{\ell_i} = \mu_i$ für jedes $i = 1, \ldots, p$. □

Nach Lemma 4.5 gibt es für jedes $i = 1, \ldots, p$ eine eindeutige Lösung

$$u_i = (\lambda^i_1, \ldots, \lambda^i_n) \in \mathrm{Lös}(A, 0)$$

mit $\lambda^i_{\ell_i} = 1$ und $\lambda^i_{\ell_j} = 0$ für jedes $j = 1, \ldots, p$ mit $j \neq i$.

Satz 4.9 $(u_1, \ldots, u_p)$ *ist eine Basis von* $\mathrm{Lös}(A, 0)$.

Beweis Sei $u = (\lambda_1, \ldots, \lambda_n) \in \mathrm{Lös}(A, 0)$ und setze $u' = u - (\lambda_{\ell_1} u_1 + \cdots + \lambda_{\ell_p} u_p)$. Dann ist $u' = (\lambda'_1, \ldots, \lambda'_n) \in \mathrm{Lös}(A, 0)$, da $\mathrm{Lös}(A, 0)$ ein Untervektorraum von K^n ist, und für jedes $i = 1, \ldots, p$ gilt

$$\begin{aligned} \lambda'_{\ell_i} &= \lambda_{\ell_i} - (\lambda_{\ell_1} \lambda^1_{\ell_i} + \cdots + \lambda_{\ell_p} \lambda^p_{\ell_i}) \\ &= \lambda_{\ell_i} - (\lambda_{\ell_1} 0 + \cdots + \lambda_{\ell_{i-1}} 0 + \lambda_{\ell_i} 1 + \lambda_{\ell_{i+1}} 0 + \cdots + \lambda_{\ell_p} 0) = \lambda_{\ell_i} - \lambda_{\ell_i} = 0\,. \end{aligned}$$

Also ist $(\lambda'_1, \ldots, \lambda'_n) \in \text{Lös}(A, 0)$ mit $\lambda'_{\ell_i} = 0$ für $i = 1, \ldots, p$ und daraus folgt, dass $(\lambda'_1, \ldots, \lambda'_n) = (0, \ldots, 0)$, da nach Lemma 4.5 $(0, \ldots, 0)$ das einzige Element von $\text{Lös}(A, 0)$ mit dieser Eigenschaft ist. Dies zeigt, dass $u' = 0$, d.h.,

$$u = \lambda_{\ell_1} u_1 + \cdots \lambda_{\ell_p} u_p$$

und insbesondere ist $\text{L}(u_1, \ldots, u_p) = \text{Lös}(A, 0)$.

Seien nun $\lambda_1, \ldots, \lambda_p \in K$ mit $\lambda_1 u_1 + \cdots + \lambda_p u_p = 0$. Für $i = 1, \ldots, p$ gilt dann

$$0 = \lambda_1 \lambda^1_{\ell_i} + \cdots + \lambda_p \lambda^p_{\ell_i} = \lambda_1 0 + \cdots + \lambda_{i-1} 0 + \lambda_i 1 + \lambda_{i+1} 0 + \cdots + \lambda_p 0 \;=\; \lambda_i \;,$$

d.h., $\lambda_i = 0$ für jedes $i = 1, \ldots, p$ und daher sind $u_1, \ldots, u_p$ linear unabhängig. Dies zeigt, dass $(u_1, \ldots, u_p)$ eine Basis von $\text{Lös}(A, 0)$ ist. □

Nach Satz 2.4 und Satz 2.6 braucht man eigentlich nur eine Basis von $\text{Lös}(A, 0)$ zu finden, wenn die Matrix $A \neq 0$ Zeilen-Stufen-Form hat. Nehme nun also an, dass dies der Fall ist und sei q der Index der letzten von Null verschiedenen Zeile von A. Für $i = 1, \ldots, m$ sei

$$k_i = \begin{cases} \min\{1 \leq j \leq n : a_{ij} \neq 0\} & \text{falls } (a_{i1}, \ldots, a_{in}) \neq 0 \;, \\ 0 & \text{sonst .} \end{cases}$$

Da die Matrix A Zeilen-Stufen-Form hat, gilt $k_1 < k_2 < \cdots < k_q$.

Lemma 4.6 *$(v_{k_1}, \ldots, v_{k_q})$ ist eine Basis von $\text{L}(v_1, \ldots, v_n)$.*

Beweis Übung. (Den Beweis findet man später im Beweis für Satz 7.2.) □

Es ist gerade diese Basis, die benutzt werden soll, um die Basis $(u_1, \ldots, u_n)$ von $\text{Lös}(A, 0)$ zu konstruieren.

Beispiel: Sei $A \in \mathrm{M}(4 \times 6, \mathbb{R})$ folgende Matrix:

$$\begin{pmatrix} 1 & 2 & 3 & 0 & 7 & -1 \\ 1 & 2 & 2 & 4 & 7 & 0 \\ 2 & 4 & 6 & 1 & 12 & -1 \\ 1 & 2 & 2 & 5 & 5 & 1 \end{pmatrix}$$

Durch eine geeignete Folge von elementaren Zeilenumformungen läßt sich A in folgende Matrix $A' \in \mathrm{M}(4 \times 6, \mathbb{R})$ überführen:

$$\begin{pmatrix} 1 & 2 & 3 & 0 & 7 & -1 \\ 0 & 0 & -1 & 4 & 0 & 1 \\ 0 & 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

und nach Satz 2.4 ist $\mathrm{Lös}(A', 0) = \mathrm{Lös}(A, 0)$. Seien v_1, v_2, v_3, v_4, v_5, v_6 die Spalten von A'. Dann ist (v_1, v_3, v_4) eine Basis von $\mathrm{L}(v_1, v_2, v_3, v_4, v_5, v_6)$. Nach Lemma 4.5 gibt es also eindeutige Lösungen

$$u_1 = (\lambda_1^1, \lambda_2^1, \lambda_3^1, \lambda_4^1, \lambda_5^1, \lambda_6^1) \in \mathrm{Lös}(A', 0) \text{ mit } \lambda_2^1 = 1,\ \lambda_5^1 = 0 \text{ und } \lambda_6^1 = 0,$$
$$u_2 = (\lambda_1^2, \lambda_2^2, \lambda_3^2, \lambda_4^2, \lambda_5^2, \lambda_6^2) \in \mathrm{Lös}(A', 0) \text{ mit } \lambda_2^2 = 0,\ \lambda_5^2 = 1 \text{ und } \lambda_6^2 = 0,$$
$$u_3 = (\lambda_1^3, \lambda_2^3, \lambda_3^3, \lambda_4^3, \lambda_5^3, \lambda_6^3) \in \mathrm{Lös}(A', 0) \text{ mit } \lambda_2^3 = 0,\ \lambda_5^3 = 0 \text{ und } \lambda_6^3 = 1,$$

und nach Satz 4.9 ist dann (u_1, u_2, u_3) eine Basis von $\mathrm{Lös}(A', 0)$ und damit auch von $\mathrm{Lös}(A, 0)$. In der Tat ist hier

$$\begin{aligned} u_1 &= (-2,\ 1,\ 0,\ 0,\ 0,\ 0)\ , \\ u_2 &= (-31,\ 0,\ 8,\ 2,\ 1,\ 0)\ , \\ u_3 &= (10,\ 0,\ -3,\ -1,\ 0,\ 1)\ . \end{aligned}$$

5 Lineare Abbildungen

Seien X, Y, Z Mengen. Sind $f : X \to Y$ und $g : Y \to Z$ Abbildungen, so gibt es eine Abbildung $g \circ f : X \to Z$, die definiert ist durch

$$(g \circ f)(x) = g(f(x))$$

für jedes $x \in X$. Die Operation $\circ$ is assoziativ: Sind $f : X \to Y$, $g : Y \to Z$ und $h : Z \to W$ Abbildungen, so gilt $h \circ (g \circ f) = (h \circ g) \circ f$, und folglich kann man einfach $h \circ g \circ f$ schreiben.

Die *Identitätsabbildung* $\mathrm{id}_X : X \to X$ wird definiert durch $\mathrm{id}_X(x) = x$ für jedes $x \in X$. Es gilt $f \circ \mathrm{id}_X = f$ für jede Abbildung $f : X \to Y$ und $\mathrm{id}_X \circ g = g$ für jede Abbildung $g : Z \to X$.

Ist $f : X \to Y$ eine Abbildung und $X' \subset X$ eine Teilmenge von X, so wird die Teilmenge $\{y \in Y : y = f(x) \text{ für ein } x \in X'\}$ von Y mit $f(X')$ bezeichnet.

Eine Abbildung $f : X \to Y$ heißt *injektiv*, wenn $f(x_1) \neq f(x_2)$ für alle $x_1 \neq x_2$. Die Abbildung f heißt *surjektiv*, wenn $f(X) = Y$ (d.h., wenn es zu jedem $y \in Y$ ein $x \in X$ gibt, so dass $f(x) = y$). Ist f injektiv und surjektiv, so heißt f *bijektiv*.

Lemma 5.1 *(1) Sei $f : X \to Y$ bijektiv; dann gibt es eine eindeutige Abbildung $f^{-1} : Y \to X$, so dass $f^{-1} \circ f = \mathrm{id}_X$ und $f \circ f^{-1} = \mathrm{id}_Y$. Die Abbildung f^{-1} heißt Umkehrabbildung von f und sie is ebenfalls bijektiv.*

(2) Sei $f : X \to Y$ eine Abbildung; gibt es dann eine Abbildung $g : Y \to X$ mit $g \circ f = \mathrm{id}_X$ und $f \circ g = \mathrm{id}_Y$, so ist f bijektiv und es gilt $g = f^{-1}$.

(3) Seien $f : X \to Y$ und $g : Y \to X$ Abbildungen mit $g \circ f = \mathrm{id}_X$. Dann ist f injektiv und g surjektiv.

Beweis Übung. □

Im folgenden sei K ein Körper und seien V, W Vektorräume über K.

Eine Abbildung $f : V \to W$ heißt *linear* oder *Homomorphismus*, wenn

$$f(\lambda u + \mu v) = \lambda f(u) + \mu f(v)$$

für alle u, $v \in V$ und alle λ, $\mu \in K$ gilt. Die Menge aller Homomorphismen von V nach W wird mit $\mathrm{Hom}(V, W)$ bezeichnet. Sei $f \in \mathrm{Hom}(V, W)$; dann gilt

(1) $f(u + v) = f(u) + f(v)$ für alle u, $v \in V$,

(2) $f(\lambda v) = \lambda f(v)$ für alle $v \in V$, $\lambda \in K$,

da $f(\lambda v) = f(\lambda v + 00) = \lambda f(v) + 0f(0) = \lambda f(v)$ und

$$f(u+v) = f(1u+1v) = 1f(u) + 1f(v) = f(u) + f(v) .$$

Sind umgekehrt (1) und (2) erfüllt, so ist f linear, da

$$f(\lambda u + \mu v) = f(\lambda u) + f(\mu v) = \lambda f(u) + \mu f(v) .$$

Ist $f : V \to W$ eine lineare Abbildung, so gilt insbesondere $f(0) = 0$, da

$$f(0) = f(00 + 00) = 0f(0) + 0f(0) = 0 + 0 = 0 .$$

Ferner ist $f(-v) = -f(v)$, da $f(-v) = f((-1)v) = (-1)f(v) = -f(v)$ für alle $v \in V$. Sind $v_1, \ldots, v_m \in V$ (mit $m \geq 2$), dann gilt

$$f(\lambda_1 v_1 + \cdots + \lambda_m v_m) = \lambda_1 f(v_1) + \cdots + \lambda_m f(v_m)$$

für alle $\lambda_1, \ldots, \lambda_m \in K$.

Satz 5.1 *Sei $A = (a_{ij}) \in \mathrm{M}(m \times n, K)$ und seien $v_1, \ldots, v_n \in K^m$ die Spalten von A (es gilt also $v_j = (a_{1j}, \ldots, a_{mj})$ für jedes $j = 1, \ldots, n$). Sei $\varphi_A : K^n \to K^m$ die Abbildung, die definiert ist durch*

$$\varphi_A((\lambda_1, \ldots, \lambda_n)) = \lambda_1 v_1 + \cdots + \lambda_n v_n$$

für jedes $(\lambda_1, \ldots, \lambda_n) \in K^n$. Dann ist φ_A linear (d.h., $\varphi_A \in \mathrm{Hom}(K^n, K^m)$). Sei umgekehrt $f \in \mathrm{Hom}(K^n, K^m)$; dann gibt es ein eindeutiges $A \in \mathrm{M}(m \times n, K)$, so dass $f = \varphi_A$.

Beweis Seien $u, v \in K^n$ mit $u = (\lambda_1, \ldots, \lambda_n)$ und $v = (\mu_1, \ldots, \mu_n)$, und seien $\lambda, \mu \in K$. Dann ist

$$\begin{aligned}
\varphi_A(\lambda u + \mu v) &= \varphi_A((\lambda\lambda_1 + \mu\mu_1, \ldots, \lambda\lambda_n + \mu\mu_n)) \\
&= (\lambda\lambda_1 + \mu\mu_1)v_1 + \cdots + (\lambda\lambda_n + \mu\mu_n)v_n \\
&= (\lambda\lambda_1 v_1 + \cdots + \lambda\lambda_n v_n) + (\mu\mu_1 v_1 + \cdots + \mu\mu_n v_n) \\
&= \lambda(\lambda_1 v_1 + \cdots + \lambda_n v_n) + \mu(\mu_1 v_1 + \cdots + \mu_n v_n) \\
&= \lambda\varphi_A((\lambda_1, \ldots, \lambda_n)) + \mu\varphi_A((\mu_1, \ldots, \mu_n)) = \lambda\varphi_A(u) + \mu\varphi_A(v) .
\end{aligned}$$

Damit ist die Abbildung φ_A linear. Sei umgekehrt $f \in \mathrm{Hom}(K^n, K^m)$ und sei $(e_1, \ldots, e_n)$ die kanonische Basis von K^n. Dann sind $f(e_1), \ldots, f(e_n)$ Elemente aus K^m, sei also $A = (a_{ij}) \in \mathrm{M}(m \times n, K)$ die Matrix, die $f(e_1), \ldots, f(e_n)$ als Spalten hat (und folglich ist $(a_{1j}, a_{2j}, \ldots, a_{mj}) = f(e_j)$ für $j = 1, \ldots, n$). Sei $(\lambda_1, \ldots, \lambda_n) \in K^n$; dann ist $(\lambda_1, \ldots, \lambda_n) = \lambda_1 e_1 + \cdots + \lambda_n e_n$ und damit ist

$$\begin{aligned}
f((\lambda_1, \ldots, \lambda_n)) &= f(\lambda_1 e_1 + \cdots + \lambda_n e_n) \\
&= \lambda_1 f(e_1) + \cdots + \lambda_n f(e_n) = \varphi_A((\lambda_1, \ldots, \lambda_n)) ,
\end{aligned}$$

d.h., $f = \varphi_A$. Ferner ist es klar, dass A eindeutig durch f bestimmt ist, da $f(e_j) = \varphi_A(e_j)$ die j te Spalte von A sein muss. □

Sei $A = (a_{ij}) \in \mathrm{M}(m \times n, K)$. Dann gilt

$$\varphi_A((\lambda_1, \ldots, \lambda_n)) = \Big(\sum_{j=1}^{n} a_{1j}\lambda_j, \ldots, \sum_{j=1}^{n} a_{mj}\lambda_j\Big)$$

für alle $(\lambda_1, \ldots, \lambda_n) \in K^n$, da

$$\begin{aligned}\varphi_A((\lambda_1, \ldots, \lambda_n)) &= \lambda_1 v_1 + \cdots + \lambda_n v_n \\ &= \lambda_1(a_{11}, \ldots, a_{m1}) + \cdots + \lambda_n(a_{1n}, \ldots, a_{mn}) = \Big(\sum_{j=1}^{n} a_{1j}\lambda_j, \ldots, \sum_{j=1}^{n} a_{mj}\lambda_j\Big)\,,\end{aligned}$$

wobei $v_1, \ldots, v_n$ die Spalten von A sind.

Es ist klar, dass für jeden Vektorraum V die Identitätsabbildung $\mathrm{id}_V : V \to V$ linear ist, d.h., $\mathrm{id}_V \in \mathrm{Hom}(V, V)$.

Lemma 5.2 *Seien $f \in \mathrm{Hom}(U, V)$, $g \in \mathrm{Hom}(V, W)$, (wobei U ein weiterer Vektorraum über K ist). Dann ist die Abbildung $g \circ f : U \to W$ linear, d.h., $g \circ f \in \mathrm{Hom}(U, W)$.*

Beweis Seien u, $v \in U$ und λ, $\mu \in K$. Dann ist

$$\begin{aligned}(g \circ f)(\lambda u + \mu v) &= g(f(\lambda u + \mu v)) = g(\lambda f(u) + \mu f(v)) \\ &= \lambda g(f(u)) + \mu g(f(v)) = \lambda(g \circ f)(u) + \mu(g \circ f)(v)\,.\end{aligned}$$

Damit ist die Abbildung $g \circ f : U \to W$ linear. □

Sei $f \in \mathrm{Hom}(V, W)$; dann heißt f *Isomorphismus*, wenn es ein $g \in \mathrm{Hom}(W, V)$ mit $g \circ f = \mathrm{id}_V$ und $f \circ g = \mathrm{id}_W$ gibt. In diesem Fall ist nach Lemma 5.1 (2) die Abbildung f bijektiv und g ist die Umkehrabbildung $f^{-1} : W \to V$. Aber die Umkehrung ist auch richtig:

Lemma 5.3 *Ist $f \in \mathrm{Hom}(V, W)$ bijektiv, so ist die Umkehrabbildung f^{-1} linear, d.h., $f^{-1} \in \mathrm{Hom}(W, V)$, und damit ist f ein Isomorphismus.*

Beweis Seien w, $w' \in W$ und λ, $\mu \in K$. Dann gilt

$$\begin{aligned}f(f^{-1}(\lambda w + \mu w')) &= \lambda w + \mu w' \\ &= \lambda f(f^{-1}(w)) + \mu f(f^{-1}(w')) = f(\lambda f^{-1}(w) + \mu f^{-1}(w'))\end{aligned}$$

und daraus folgt, dass $f^{-1}(\lambda w + \mu w') = \lambda f^{-1}(w) + \mu f^{-1}(w')$. □

Ist $f \in \mathrm{Hom}(V, W)$ ein Isomorphismus, dann ist $f^{-1} \in \mathrm{Hom}(W, V)$ ebenfalls ein Isomorphismus.

Lemma 5.4 *Sei $f : V \to W$ eine lineare Abbildung und U ein Untervektorraum von V. Dann ist $f(U)$ ein Untervektorraum von W.*

Beweis Da $0 = f(0)$ und $0 \in U$, ist $0 \in f(U)$. Seien w, $w' \in f(U)$. Es gibt also u, $u' \in U$ mit $f(u) = w$ und $f(u') = w'$ und folglich ist

$$f(u + u') = f(u) + f(u') = w + w' .$$

Aber $u+u' \in U$, da U ein Untervektorraum von V ist, und daher ist $w+w' \in f(U)$. Sei nun $\lambda \in K$; dann ist $f(\lambda u) = \lambda f(u) = \lambda w$ und $\lambda u \in U$. Damit ist $\lambda w \in f(U)$ für jedes $\lambda \in K$. Folglich ist $f(U)$ ein Untervektorraum von W. □

Für eine lineare Abbildung $f : V \to W$ setze

$$\operatorname{Kern} f = \{v \in V : f(v) = 0\} \quad \text{und} \quad \operatorname{Bild} f = f(V) .$$

Lemma 5.5 *Sei $f : V \to W$ eine lineare Abbildung. Dann ist* Kern f *ein Untervektorraum von V und* Bild f *ein Untervektorraum von W.*

Beweis Da $f(0) = 0$, ist $0 \in \operatorname{Kern} f$. Seien v, $v' \in \operatorname{Kern} f$ und $\lambda \in K$. Dann gilt $f(v + v') = f(v) + f(v') = 0 + 0 = 0$ und $f(\lambda v) = \lambda f(v) = \lambda\, 0 = 0$, und damit ist $v + v' \in \operatorname{Kern} f$ und $\lambda v \in \operatorname{Kern} f$. Folglich ist $\operatorname{Kern} f$ ein Untervektorraum von V. Nach Lemma 5.4 ist $\operatorname{Bild} f = f(V)$ ein Untervektorraum von W. □

Lemma 5.6 *Eine lineare Abbildung $f : V \to W$ ist injektiv genau dann, wenn* $\operatorname{Kern} f = \{0\}$.

Beweis Ist f injektiv, so gilt insbesondere, dass $f(v) \neq f(0) = 0$ für alle $v \neq 0$ und damit ist $\operatorname{Kern} f = \{0\}$. Nehme umgekehrt an, dass $\operatorname{Kern} f = \{0\}$ und seien u, $v \in V$ mit $u \neq v$. Dann ist $u + (-v) \neq 0$ und daher ist auch $f(u + (-v)) \neq 0$. Aber $f(u + (-v)) = f(u) + (-f(v))$ und daraus ergibt sich, dass $f(u) \neq f(v)$. □

Lemma 5.7 *Sei $b \in K^m$ und $x \in K^n$; dann gilt $x \in \operatorname{Lös}(A, b)$ genau, wenn $\varphi_A(x) = b$. Insbesondere ist* $\operatorname{Lös}(A, 0) = \operatorname{Kern} \varphi_A$.

Beweis Dies ist klar. □

Lemma 5.8 *Sei $f : V \to W$ eine lineare Abbildung. Dann gilt*

$$f(\mathrm{L}(v_1, \ldots, v_m)) = \mathrm{L}(f(v_1), \ldots, f(v_m))$$

für alle $v_1, \ldots, v_m \in V$.

Beweis Für alle $\lambda_1, \ldots, \lambda_m \in K$ ist

$$f(\lambda_1 v_1 + \cdots + \lambda_m v_m) = \lambda_1 f(v_1) + \cdots + \lambda_m f(v_m)$$

und daraus folgt unmittelbar, dass $f(\mathrm{L}(v_1, \ldots, v_m)) = \mathrm{L}(f(v_1), \ldots, f(v_m))$. □

Lemma 5.9 *Sei U ein endlichdimensionaler Untervektorraum von V und sei $f : V \to W$ eine lineare Abbildung. Dann ist der Untervektorraum $f(U)$ von W auch endlichdimensional und es gilt* $\dim f(U) \leq \dim U$.

Beweis Sei $(u_1, \ldots, u_m)$ eine Basis von U. Nach Lemma 5.8 ist dann

$$f(U) = f(\mathrm{L}(u_1, \ldots, u_m)) = \mathrm{L}(f(u_1), \ldots, f(u_m)) \ ,$$

und daraus folgt nach Satz 4.3 und Satz 4.6 (3), dass $f(U)$ endlichdimensional ist und $\dim f(U) \leq m = \dim U$. □

Satz 5.2 (Dimensionsformel) *Sei $f : V \to W$ eine lineare Abbildung. Ist V endlichdimensional, so sind* Kern f *und* Bild f *auch endlichdimensional und*

$$\dim V = \dim \mathrm{Kern}\, f + \dim \mathrm{Bild}\, f \ .$$

Beweis Nach Satz 4.4 und Lemma 5.9 sind Kern f und Bild f endlichdimensional. Sei $(u_1, \ldots, u_m)$ eine Basis von Kern f und $(w_1, \ldots, w_p)$ eine Basis von Bild f (mit $m = 0$ bzw. $p = 0$, falls Kern $f = \{0\}$ bzw. Bild $f = \{0\}$). Da $w_i \in \mathrm{Bild}\, f = f(V)$, gibt es für jedes $i = 1, \ldots, p$ einen Vektor $v_i \in V$, so dass $f(v_i) = w_i$. Es wird gezeigt, dass $(u_1, \ldots, u_m, v_1, \ldots, v_p)$ eine Basis von V ist.

$V = \mathrm{L}(u_1, \ldots, u_m, v_1, \ldots, v_p)$: Sei $v \in V$; da $f(v) \in \mathrm{Bild}\, f$ und $(w_1, \ldots, w_p)$ eine Basis von Bild f ist, gibt es $\mu_1, \ldots, \mu_p \in K$, so dass $f(v) = \mu_1 w_1 + \cdots + \mu_p w_p$. Setze $v' = \mu_1 v_1 + \cdots + \mu_p v_p$; dann ist

$$f(v') = f(\mu_1 v_1 + \cdots + \mu_p v_p) = \mu_1 f(v_1) + \cdots + \mu_p f(v_p) = \mu_1 w_1 + \cdots + \mu_p w_p = f(v)$$

und damit ist $f(v + (-v')) = f(v) + (-f(v')) = 0$, d.h., $v + (-v') \in \mathrm{Kern}\, f$. Also gibt es $\lambda_1, \ldots, \lambda_m \in K$, so dass $v + (-v') = \lambda_1 u_1 + \cdots + \lambda_m u_m$ und folglich ist

$$v = (v + (-v')) + v' = \lambda_1 u_1 + \cdots + \lambda_m u_m + \mu_1 v_1 + \cdots + \mu_p v_p \ ,$$

d.h., $v \in \mathrm{L}(u_1, \ldots, u_m, v_1, \ldots, v_p)$.

Die Vektoren $u_1, \ldots, u_m, v_1, \ldots, v_p$ sind linear unabhängig: Nehme also an, dass $\lambda_1 u_1 + \cdots + \lambda_m u_m + \mu_1 v_1 + \cdots + \mu_p v_p = 0$. Dann ist

$$\begin{aligned} 0 = f(0) &= f(\lambda_1 u_1 + \cdots + \lambda_m u_m + \mu_1 v_1 + \cdots + \mu_p v_p) \\ &= \lambda_1 f(u_1) + \cdots + \lambda_m f(u_m) + \mu_1 f(v_1) + \cdots + \mu_p f(v_p) = \mu_1 w_1 + \cdots + \mu_p w_p \ , \end{aligned}$$

da $u_1, \ldots, u_m \in \operatorname{Kern} f$ und $f(v_i) = w_i$ für $i = 1, \ldots, p$. Aber $w_1, \ldots, w_p$ sind linear unabhängig und daraus folgt, dass $\mu_i = 0$ für $i = 1, \ldots, p$. Nun ist $\lambda_1 u_1 + \cdots + \lambda_m u_m = 0$ und die Vektoren $u_1, \ldots, u_m$ sind linear unabhängig. Damit gilt auch $\lambda_j = 0$ für $j = 1, \ldots, m$, d.h., die Vektoren $u_1, \ldots, u_m, v_1, \ldots, v_p$ sind linear unabhängig.

Dies zeigt also, dass $(u_1, \ldots, u_m, v_1, \ldots, v_p)$ eine Basis von V ist, und insbesondere ist $\dim V = m + p = \dim \operatorname{Kern} f + \dim \operatorname{Bild} f$. □

Satz 5.3 *Seien V, W endlichdimensional mit $\dim V = \dim W$, sei $f : V \to W$ eine lineare Abbildung. Dann sind äquivalent:*

(1) f ist ein Isomorphismus.

(2) f ist injektiv.

(3) f ist surjektiv.

Beweis Nach Lemma 5.6 ist f injektiv genau dann, wenn $\operatorname{Kern} f = \{0\}$, und $\operatorname{Kern} f = \{0\}$ genau dann, wenn $\dim \operatorname{Kern} f = 0$. Andererseits ist nach Satz 4.7 f surjektiv genau dann, wenn $\dim \operatorname{Bild} f = \dim W$. Aber nach Satz 5.2 ist $\dim \operatorname{Kern} f = 0$ genau dann, wenn $\dim \operatorname{Bild} f = \dim W$, da $\dim V = \dim W$. Folglich ist f injektiv genau dann, wenn f surjektiv ist (und nach Lemma 5.3 ist f ein Isomorphismus genau dann, wenn f bijektiv ist). □

Eine lineare Abbildung $f : V \to V$ nennt man einen *Endomorphismus* von V. Ein Endomorphismus, der auch ein Isomorphismus ist, heißt *Automorphismus* von V. Mit f ist auch f^{-1} ein Automorphismus von V.

Sei V endlichdimensional und f ein Endomorphismus von V. Nach Satz 5.3 sind dann äquivalent:

(1) f ist ein Automorphismus.

(2) f ist injektiv.

(3) f ist surjektiv.

(Die folgenden Beispiele zeigen, dass diese Behauptung im Allgemeinen nicht richtig ist, wenn V nicht endlichdimensional ist. Sei $V = \operatorname{Abb}(\mathbb{N}, K)$ und definiere $f, g : V \to V$ durch $f(\{\lambda_n\}_{n\geq 0}) = \{\lambda'_n\}_{n\geq 0}$ und $g(\{\lambda_n\}_{n\geq 0}) = \{\lambda''_n\}_{n\geq 0}$, wobei $\lambda'_n = \lambda_{n+1}$ für alle $n \geq 0$ und $\lambda''_0 = 0$, $\lambda''_n = \lambda_{n-1}$ für alle $n \geq 1$. Dann sieht man leicht, dass f und g Endomorphismen von V sind. Die Abbildung f ist surjektiv aber nicht injektiv und g ist injektiv aber nicht surjektiv.)

Lemma 5.10 *Sei $f : V \to W$ eine lineare Abbildung und seien $v_1, \ldots, v_m \in V$.*

(1) Ist f injektiv und sind $v_1, \ldots, v_m$ linear unabhängig, so sind $f(v_1), \ldots, f(v_m)$ linear unabhängig.

(2) Ist f surjektiv und ist $\mathrm{L}(v_1, \ldots, v_m) = V$, so ist $\mathrm{L}(f(v_1), \ldots, f(v_m)) = W$.

(3) Ist f ein Isomorphismus und ist $(v_1, \ldots, v_m)$ eine Basis von V, dann ist auch $(f(v_1), \ldots, f(v_m))$ eine Basis von W.

(4) Ist $(v_1, \ldots, v_m)$ eine Basis von V und $(f(v_1), \ldots, f(v_m))$ eine Basis von W, dann ist f ein Isomorphismus.

Beweis (1) Seien $\lambda_1, \ldots, \lambda_m \in K$ mit $\lambda_1 f(v_1) + \cdots + \lambda_m f(v_m) = 0$. Dann gilt

$$f(\lambda_1 v_1 + \cdots + \lambda_m v_m) = \lambda_1 f(v_1) + \cdots + \lambda_m f(v_m) = 0$$

und damit $\lambda_1 v_1 + \cdots + \lambda_m v_m = 0$, da f injektiv ist. Folglich ist $\lambda_j = 0$ für jedes j, da $v_1, \ldots, v_m$ linear unabhängig sind. Also sind $f(v_1), \ldots, f(v_m)$ linear unabhängig.

(2) Dies folgt unmittelbar aus Lemma 5.8.

(3) Dies folgt unmittelbar aus (1) und (2).

(4) Nach Lemma 5.8 gilt

$$f(V) = f(\mathrm{L}(v_1, \ldots, v_m)) = \mathrm{L}(f(v_1), \ldots, f(v_m)) = W \ ,$$

d.h., f ist surjektiv. Aber $\dim V = n = \dim W$ und daraus folgt nach Satz 5.2, dass f ein Isomorphismus ist. (Man kann auch zeigen, dass f injektiv ist, ohne Satz 5.2 anzuwenden: Sei $v \in \mathrm{Kern}\, f$; da $(v_1, \ldots, v_m)$ eine Basis von V ist, gibt es $\lambda_1, \ldots, \lambda_m \in K$ mit $\lambda_1 v_1 + \cdots + \lambda_m v_m = 0$ und dann gilt

$$0 = f(v) = f(\lambda_1 v_1 + \cdots + \lambda_m v_m) = \lambda_1 f(v_1) + \cdots + \lambda_m f(v_m) \ .$$

Daraus folgt, dass $\lambda_j = 0$ für jedes $j = 1, \ldots, m$, da $f(v_1), \ldots, f(v_m)$ linear unabhängig sind; d.h., $v = 0$ und damit ist $\mathrm{Kern}\, f = \{0\}$.) □

Lemma 5.11 *Sei V endlichdimensional, sei $(v_1, \ldots, v_m)$ eine Basis von V und $w_1, \ldots, w_m \in W$. Dann gibt es eine eindeutige lineare Abbildung $f : V \to W$, so dass $f(v_j) = w_j$ für jedes $j = 1, \ldots, m$.*

Beweis Nach Satz 4.1 gibt es zu jedem $v \in V$ genau ein $(\lambda_1, \ldots, \lambda_m) \in K^m$, so dass $v = \lambda_1 v_1 + \cdots + \lambda_m v_m$. Folglich kann eine Abbildung $f : V \to W$ definiert werden durch

$$f(v) = \lambda_1 w_1 + \cdots + \lambda_m w_m \ , \text{ falls } v = \lambda_1 v_1 + \cdots + \lambda_m v_m \ .$$

Dann ist f linear: Seien u, $v \in V$, λ, $\mu \in K$ mit $u = \lambda_1 v_1 + \cdots + \lambda_m v_m$ und $v = \mu_1 v_1 + \cdots + \mu_m v_m$. Folglich ist

$$\begin{aligned}\lambda u + \mu v &= \lambda(\lambda_1 v_1 + \cdots + \lambda_m v_m) + \mu(\mu_1 v_1 + \cdots + \mu_m v_m) \\ &= (\lambda\lambda_1 + \mu\mu_1)v_1 + \cdots + (\lambda\lambda_m + \mu\mu_m)v_m\end{aligned}$$

und damit ist per Definition

$$\begin{aligned}f(\lambda u + \mu v) &= (\lambda\lambda_1 + \mu\mu_1)w_1 + \cdots + (\lambda\lambda_m + \mu\mu_m)w_m \\ &= \lambda(\lambda_1 w_1 + \cdots + \lambda_m w_m) + \mu(\mu_1 w_1 + \cdots + \mu_m w_m) = \lambda f(u) + \mu f(v)\ .\end{aligned}$$

Ferner gilt offensichtlich $f(v_j) = w_j$ für jedes $j = 1, \ldots, m$.

Eindeutigkeit: Seien f, $g \in \mathrm{Hom}(V, W)$ mit $f(v_j) = w_j = g(v_j)$ für jedes j. Sei $v \in V$; es gibt also $\lambda_1, \ldots, \lambda_m \in K$, so dass $v = \lambda_1 v_1 + \cdots + \lambda_m v_m$, and dann ist

$$\begin{aligned}f(v) &= f(\lambda_1 v_1 + \cdots + \lambda_m v_m) = \lambda_1 f(v_1) + \cdots + \lambda_m f(v_m) \\ &= \lambda_1 g(v_1) + \cdots + \lambda_m g(v_m) = g(\lambda_1 v_1 + \cdots + \lambda_m v_m) = g(v)\ ;\end{aligned}$$

d.h., $f = g$. □

Wenn es einen Isomorphismus $f : V \to W$ gibt, so heißen V und W *isomorph*. (Da für jeden Isomorphismus $f : V \to W$ auch $f^{-1} : W \to V$ ein Isomorphismus ist, sind die Rollen von V und W in dieser Definition symmetrisch.)

Satz 5.4 *Ist V endlichdimensional, so sind V und W isomorph genau dann, wenn W endlichdimensional ist und* $\dim V = \dim W$.

Beweis Sind V und W isomorph, dann folgt unmittelbar aus Lemma 5.10 (3), dass W endlichdimensional ist und $\dim V = \dim W$. Sei umgekehrt W endlichdimensional mit $\dim V = \dim W = n$, und sei $(v_1, \ldots, v_m)$ eine Basis von V und $(w_1, \ldots, w_m)$ eine Basis von W. Nach Lemma 5.11 gibt es eine lineare Abbildung $f : V \to W$ mit $f(v_j) = w_j$ für jedes $j = 1, \ldots, m$ und nach Lemma 5.10 (4) ist dann f ein Isomorphismus, d.h., V und W sind isomorph. □

Ist V endlichdimensional mit $\dim V = n$, so ist nach Satz 5.4 V isomorph zu K^n.

Sei $f : V \to W$ eine lineare Abbildung. Ist der Untervektorraum $\mathrm{Bild}\, f$ von W endlichdimensional, dann nennt man $\dim \mathrm{Bild}\, f$ den *Rang* von f und bezeichnet diese Zahl mit $\mathrm{rang}\, f$.

Lemma 5.12 *Seien U, V und W endlichdimensionale Vektorräume über K und seien $f \in \mathrm{Hom}(U, V)$, $g \in \mathrm{Hom}(V, W)$ (also ist $g \circ f \in \mathrm{Hom}(U, W)$). Dann gilt:*

(1) $\mathrm{rang}\, g \circ f \leq \min\{\mathrm{rang}\, f, \mathrm{rang}\, g\}$.

(2) Ist f surjektiv, so ist $\mathrm{rang}\, g \circ f = \mathrm{rang}\, g$.

(3) Ist g injektiv, so ist $\mathrm{rang}\, g \circ f = \mathrm{rang}\, f$.

Beweis (1) Es gilt $\operatorname{Bild}(g \circ f) = (g \circ f)(U) = g(f(U)) \subset g(V) = \operatorname{Bild} g$ und damit ist $\operatorname{rang} g \circ f = \dim \operatorname{Bild}(g \circ f) \leq \dim \operatorname{Bild} g = \operatorname{rang} g$. Andererseits gilt nach Lemma 5.9, dass

$$\operatorname{rang} g \circ f = \dim(g \circ f)(U) = \dim g(f(U)) \leq \dim f(U) = \operatorname{rang} f \ .$$

(2) Ist f surjektiv, so ist $(g \circ f)(U) = g(f(U)) = g(V)$ und damit

$$\operatorname{rang} g \circ f = \dim(g \circ f)(U) = \dim g(V) = \operatorname{rang} g \ .$$

(3) Sei $(v_1, \ldots, v_m)$ eine Basis von $f(U)$. Nach Lemma 5.10 (1) sind die Vektoren $g(v_1), \ldots, g(v_m)$ linear unabhängig und nach Lemma 5.8 ist

$$(g \circ f)(U) = g(f(U)) = g(\mathrm{L}(v_1, \ldots, v_m)) = \mathrm{L}(g(v_1), \ldots, g(v_m)) \ ,$$

d.h., $(g(v_1), \ldots, g(v_m))$ ist eine Basis von $(g \circ f)(U)$. Insbesondere ist

$$\operatorname{rang} g \circ f = \dim(g \circ f)(U) = m = \dim f(U) = \operatorname{rang} f \ . \ \square$$

6 Matrizen

Im Folgenden sei K ein Körper. Sei $A = (a_{ij}) \in \mathrm{M}(\ell \times m, K)$ eine $\ell \times m$ Matrix und $B = (a_{jk}) \in \mathrm{M}(m \times n, K)$ eine $m \times n$ Matrix über K. Dann gibt es die linearen Abbildungen $\varphi_A \in \mathrm{Hom}(K^m, K^\ell)$ und $\varphi_B \in \mathrm{Hom}(K^n, K^m)$ und damit nach Lemma 5.1 die lineare Abbildung $\varphi_A \circ \varphi_B \in \mathrm{Hom}(K^n, K^\ell)$. Nach Satz 5.1 gibt es also eine eindeutige Matrix $C = (c_{ik}) \in \mathrm{M}(\ell \times n, K)$, so dass $\varphi_C = \varphi_A \circ \varphi_B$.

Lemma 6.1 *Für alle $i = 1, \ldots, \ell$, $k = 1, \ldots, n$ gilt $c_{ik} = \sum_{j=1}^{m} a_{ij} b_{jk}$.*

Beweis Sei $(e_1, \ldots, e_n)$ die kanonische Basis von K^n. Seien ferner $v_1, \ldots, v_m$ die Spalten von A, $v'_1, \ldots, v'_n$ die Spalten von B und $v''_1, \ldots, v''_n$ die Spalten von C. Für jedes $k = 1, \ldots, n$ gilt dann

$$\begin{aligned}(c_{1k}, \ldots, c_{\ell k}) &= v''_k = \varphi_C(e_k) = (\varphi_A \circ \varphi_B)(e_k) = \varphi_A(\varphi_B(e_k)) \\ &= \varphi_A(v'_k) = \varphi_A((b_{1k}, \ldots, b_{mk})) = b_{1k} v_1 + \cdots + b_{mk} v_m \\ &= b_{1k}(a_{11}, \ldots, a_{\ell,1}) + \cdots + b_{mk}(a_{1m}, \ldots, a_{\ell,m}) \\ &= \Big(\sum_{j=1}^{m} a_{1j} b_{jk}, \ldots, \sum_{j=1}^{m} a_{\ell j} b_{jk} \Big)\end{aligned}$$

und damit gilt $c_{ik} = \sum_{j=1}^{m} a_{ij} b_{jk}$ für alle $i = 1, \ldots, \ell$, $k = 1, \ldots, n$. □

Seien $A = (a_{ij}) \in \mathrm{M}(\ell \times m, K)$ und $B = (a_{jk}) \in \mathrm{M}(m \times n, K)$ und definiere eine $\ell \times n$ Matrix $C = (c_{ik}) \in \mathrm{M}(\ell \times n, K)$ über K durch

$$c_{ik} = \sum_{j=1}^{m} a_{ij} b_{jk}$$

für alle $i = 1, \ldots, \ell$, $k = 1, \ldots, n$. Diese Matrix C heißt das *Produkt* von A und B und wird mit AB bezeichnet. Nach Lemma 6.1 gilt dann

$$\varphi_A \circ \varphi_B = \varphi_{AB} \, .$$

Satz 6.1 *Seien $A \in \mathrm{M}(\ell \times m, K)$, $B \in \mathrm{M}(m \times n, K)$ und $C \in \mathrm{M}(n \times p, K)$. Dann gilt $(AB)C = A(BC)$. (Die Matrizenmultiplikation ist also assoziativ.)*

Beweis Nach Lemma 6.1 gilt

$$\varphi_{(AB)C} = \varphi_{AB} \circ \varphi_C = (\varphi_A \circ \varphi_B) \circ \varphi_C = \varphi_A \circ (\varphi_B \circ \varphi_C) = \varphi_A \circ \varphi_{BC} = \varphi_{A(BC)}$$

und daraus folgt nach der Eindeutigkeit in Satz 5.1, dass $(AB)C = A(BC)$. $\square$

Seien $A_1 \in \mathrm{M}(n_0 \times n_1, K)$, $A_2 \in \mathrm{M}(n_1 \times n_2, K)$, $\ldots$, $A_m \in \mathrm{M}(n_{m-1} \times n_m, K)$. Mit Hilfe von Satz 6.1 kann man durch Induktion zeigen, dass das 'Produkt' dieser Matrizen unabhängig von der Reihenfolge der einzelnen Multiplikationen ist; dieses 'Produkt' wird mit $A_1 A_2 \cdots A_m$ bezeichnet.

Sei $A = (a_{ij}) \in \mathrm{M}(m \times n, K)$ und sei $u = (\lambda_1, \ldots, \lambda_n) \in K^n$; dann gilt

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_m \end{pmatrix} ,$$

wobei $(\mu_1, \ldots, \mu_m) = \varphi_A(u)$. Folglich ist $\varphi_A(u) = Au$, wenn u bzw. $\varphi_A(u)$ als Elemente aus $\mathrm{M}(n \times 1, K)$ bzw. $\mathrm{M}(m \times 1, K)$ betrachtet wird. Aus diesem Grund schreibt man oft Au statt $\varphi_A(u)$. Seien

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \; B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \in \mathrm{M}(2 \times 2, K) .$$

Dann sind auch AB und BA Elemente von $\mathrm{M}(2 \times 2, K)$ und es gilt

$$AB = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} ,$$
$$BA = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} .$$

Da $AB \neq BA$, ist die Matrizenmultiplikation nicht kommutativ; ferner ist sie nicht nullteilerfrei, da $A \neq 0$, $B \neq 0$ aber $AB = 0$.

Für jedes $n \geq 1$ sei $E_n \in \mathrm{M}(n \times n, K)$ folgende Matrix:

$$E_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} .$$

Es gilt also $E_n = (\delta_{ij})$, wobei $\delta_{ij} = \begin{cases} 1 & \text{falls } i = j , \\ 0 & \text{sonst} . \end{cases}$ E_n heißt *Einheitsmatrix*.

Man sieht leicht, dass $AE_n = A$ und $E_n B = B$ für alle $A \in \mathrm{M}(m \times n, K)$ und alle $B \in \mathrm{M}(n \times p, K)$. Ferner gilt $\varphi_{E_n} = \mathrm{id}_{K^n}$. Man schreibt oft einfach E statt E_n, wenn es klar ist, was n ist.

Eine Matrix A heißt *quadratisch*, wenn sie genauso viele Zeilen wie Spalten hat, d.h., wenn $A \in \mathrm{M}(n \times n, K)$ für ein $n \geq 1$. Insbesondere ist E_n quadratisch.

Lemma 6.2 *Sei $A \in \mathrm{M}(n\times n, K)$. Dann gibt es höchstens ein $A' \in \mathrm{M}(n\times n, K)$, so dass $AA' = A'A = E_n$.*

Beweis Seien B, $C \in \mathrm{M}(n \times n, K)$ mit $AB = BA = E_n = AC = CA$. Dann gilt

$$B = BE_n = B(AC) = (BA)C = E_nC = C\ . \ \square$$

Eine quadratische Matrix $A \in \mathrm{M}(n\times n, K)$ heißt *invertierbar*, wenn es eine Matrix $A' \in \mathrm{M}(n\times n, K)$ gibt, so dass $AA' = A'A = E_n$. Nach Lemma 6.2 ist diese Matrix A' eindeutig und sie wird mit A^{-1} bezeichnet, d.h., A^{-1} ist die eindeutige Matrix mit $AA^{-1} = A^{-1}A = E_n$.

Lemma 6.3 *(1) Die Einheitsmatrix E_n ist invertierbar und $E_n^{-1} = E_n$.*

(2) Ist $A \in \mathrm{M}(n \times n, K)$ invertierbar, so ist auch A^{-1} invertierbar und es gilt $(A^{-1})^{-1} = A$.

(3) Sind A, $B \in \mathrm{M}(n \times n, K)$ invertierbar, so ist auch AB invertierbar und es gilt $(AB)^{-1} = B^{-1}A^{-1}$.

Beweis (1) Dies ist klar, da $E_nE_n = E_nE_n = E_n$.

(2) Dies ist auch klar, da $A^{-1}A = AA^{-1} = E_n$.

(3) Dies folgt aus der Assoziativität der Matrizenmultiplikation, da

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AE_nA^{-1} = AA^{-1} = E_n$$

und $(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}E_nB = B^{-1}B = E_n$. $\square$

Satz 6.2 *Seien A, $B \in \mathrm{M}(n \times n, K)$ mit $AB = E_n$. Dann sind A und B schon invertierbar und es gilt $B = A^{-1}$ und $A = B^{-1}$.*

Beweis Nach Lemma 6.1 gilt $\varphi_A \circ \varphi_B = \varphi_{AB} = \varphi_{E_n} = \mathrm{id}_{K^n}$ und daraus ergibt sich nach Lemma 5.1 (3), dass φ_A surjektiv und φ_B injectiv ist. Folglich sind nach Satz 5.3 φ_A und φ_B beide Automorphismen. Nach Satz 5.1 gibt es dann A', $B' \in \mathrm{M}(n \times n, K)$, so dass $\varphi_A^{-1} = \varphi_{A'}$ und $\varphi_B^{-1} = \varphi_{B'}$ und folglich ist

$$\begin{aligned}\varphi_{AA'} = \varphi_A \circ \varphi_{A'} = \varphi_A \circ \varphi_A^{-1} &= \mathrm{id}_{K^n} \\ &= \varphi_{E_n} = \mathrm{id}_{K^n} = \varphi_A^{-1} \circ \varphi_A = \varphi_{A'} \circ \varphi_A = \varphi_{A'A}\ .\end{aligned}$$

Nach der Eindeutigkeit in Satz 5.1 ist also $AA' = E_n = A'A$ und genauso gilt $BB' = E_n = B'B$. Daher sind A und B invertierbar. Ferner gilt

$$B = E_nB = (A^{-1}A)B = A^{-1}(AB) = A^{-1}E_n = A^{-1}$$

und $A = AE_n = A(BB^{-1}) = (AB)B^{-1} = E_nB^{-1} = B^{-1}$. $\square$

Bisher wurden unter elementaren Zeilenumformungen die folgenden Operationen verstanden:

(I) Addition eines Vielfachen einer Zeile zu einer anderen Zeile.

(II) Vertauschen zweier Zeilen.

Es erweist sich aber als nützlich, eine dritte Art von elementarer Zeilenumformung zuzulassen, und zwar:

(III) Multiplikation einer Zeile mit einem Skalar $\lambda \in K$, $\lambda \neq 0$.

Satz 2.4 gilt auch für Umformungen der dritten Art: Wird (A, b) durch eine elementare Zeilenumformung vom Typ III zu einer Matrix (A', b') verändert, so gilt $\text{Lös}(A, b) = \text{Lös}(A', b')$.

Satz 6.3 *Für eine quadratische Matrix $A \in \mathrm{M}(n \times n, K)$ sind äquivalent:*

(1) Die Matrix A ist invertierbar.

(2) Es gibt eine Matrix $B \in \mathrm{M}(n \times n, K)$, so dass $BA = E_n$.

(3) Es gibt eine Matrix $C \in \mathrm{M}(n \times n, K)$, so dass $AC = E_n$.

(4) Der Endomorphismus φ_A ist ein Automorphismus von K^n.

(5) $\text{Lös}(A, 0) = \{0\}$.

(6) Das zu (A, b) gehörige Gleichungssystem ist eindeutig lösbar für jedes $b \in K^n$.

(7) A läßt sich durch eine endliche Folge von elementaren Zeilenumformungen in eine Matrix $A' = (a'_{ij})$ folgender Gestalt überführen:

$$\begin{pmatrix} * & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & * & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & * & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & * & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & * & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & * & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & 0 & * & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & \cdot \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * \end{pmatrix}.$$

Es gilt also $a'_{ii} \neq 0$ für jedes $i = 1, \ldots, n$ und $a'_{ij} = 0$, falls $i > j$.

(8) A läßt sich durch eine endliche Folge von elementaren Zeilenumformungen in die Einheitsmatrix E_n überführen.

Beweis (5) $\Leftrightarrow$ (6) $\Leftrightarrow$ (7): Dies folgt aus Satz 2.5.

(4) $\Leftrightarrow$ (5): Dies folgt aus Satz 5.3, Lemma 5.6 und Lemma 5.7.

(1) $\Rightarrow$ (2) und (1) $\Rightarrow$ (3) sind klar.

(2) $\Rightarrow$ (1) und (3) $\Rightarrow$ (1): Dies folgt aus Satz 6.2.

(4) $\Rightarrow$ (1): Nach Satz 5.1 gibt es eine eindeutige Matrix $A' \in \mathrm{M}(n \times n, K)$ mit $\varphi_A^{-1} = \varphi_{A'}$. Aber $\varphi_{A'A} = \varphi_{A'} \circ \varphi_A = \varphi_A^{-1} \circ \varphi_A = \mathrm{id}_{K^n} = \varphi_{E_n}$ und daraus folgt zusammen mit der Eindeutigkeit in Satz 5.1, dass $A'A = E_n$. Genauso gilt $AA' = E_n$. Damit ist A invertierbar (und $\varphi_A^{-1} = \varphi_{A^{-1}}$).

(1) $\Rightarrow$ (4): Nach Lemma 6.1 ist

$$\varphi_{A^{-1}} \circ \varphi_A = \varphi_{A^{-1}A} = \varphi_{E_n} = \mathrm{id}_{K^n} = \varphi_{E_n} = \varphi_{AA^{-1}} = \varphi_A \circ \varphi_{A^{-1}}$$

und damit ist φ_A ein Automorphismus.

(8) $\Rightarrow$ (5) ist klar.

(7) $\Rightarrow$ (8): Übung. $\square$

Sei $A \in \mathrm{M}(n \times n, K)$ invertierbar. Nach Satz 6.3 ist dann φ_A ein Automorphismus von K^n und der Beweis für diesen Satz zeigt, dass $\varphi_A^{-1} = \varphi_{A^{-1}}$.

Sei $1 \leq i, j \leq n$ mit $i \neq j$ und $\lambda \in K$ mit $\lambda \neq 0$. Seien $S_n(i, \lambda)$, $P_n(i, j)$ und $Q_n(i, j, \lambda)$ folgende Elemente von $\mathrm{M}(n \times n, K)$:

$$S_n(i,\lambda) = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \lambda & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \leftarrow \text{i te Zeile}$$

$$P_n(i,j) = \begin{pmatrix} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & & 0 & & 1 & & \\ & & & & 1 & & & \\ & & & & & \ddots & & \\ & & & & & & 1 & \\ & & & 1 & & & 0 & \\ & & & & & & & 1 \\ & & & & & & & & \ddots \\ & & & & & & & & & 1 \end{pmatrix} \begin{matrix} \\ \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \\ \leftarrow j\text{-te Zeile} \\ \\ \end{matrix}$$

i-te Spalte $\uparrow$ $\qquad$ $\uparrow$ j-te Spalte

$$Q_n(i,j,\lambda) = \begin{pmatrix} 1 & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & 1 & & & & & & & \\ & & & 1 & & & \lambda & & & \\ & & & & 1 & & & & & \\ & & & & & \ddots & & & & \\ & & & & & & 1 & & & \\ & & & & & & & 1 & & \\ & & & & & & & & 1 & \\ & & & & & & & & & \ddots \\ & & & & & & & & & & 1 \end{pmatrix} \begin{matrix} \\ \\ \\ \leftarrow i\text{-te Zeile} \\ \\ \\ \\ \\ \\ \\ \end{matrix}$$

$$\uparrow j\text{-te Spalte}$$

(Außer der eingetragenen oder der durch Punkte angedeuteten Komponenten sind dabei alle Komponenten gleich Null.) Solche Matrizen werden *Elementarmatrizen* genannt.

Lemma 6.4 *Elementarmatrizen sind invertierbar mit* $S_n(i,\lambda)^{-1} = S_n(i,\lambda^{-1})$, $P_n(i,j)^{-1} = P_n(i,j)$ *und* $Q_n(i,j,\lambda)^{-1} = Q_n(i,j,-\lambda)$.

Beweis Übung. □

Lemma 6.5 *Sei* $A \in \mathrm{M}(m \times n, K)$ *und sei* $1 \leq i, j \leq m$ *mit* $i \neq j$ *und* $\lambda \in K$ *mit* $\lambda \neq 0$. *Dann gilt:*

(1) Die Matrix $S_m(i,\lambda)A$ *erhält man durch Multiplikation der* i*-ten Zeile von* A *mit dem Skalar* λ.

(2) Die Matrix $P_m(i,j)A$ *erhält man durch Vertauschung der* i*-ten und der* j*-ten Zeilen von* A.

(3) Die Matrix $Q_m(i,j,\lambda)A$ *erhält man durch Addition des* λ*-fachen der* j*-ten Zeile zu der* i*-ten Zeile von* A.

Beweis Übung. □

Multiplikation von links mit einer Elementarmatrix bewirkt also eine elementare Zeilenumformung. Ferner entsteht jede elementare Zeilenumformung auf diese Weise.

Satz 6.4 *Zu jeder Matrix* $A \in \mathrm{M}(m \times n, K)$ *gibt es eine invertierbare Matrix* $P \in \mathrm{M}(m \times m, K)$, *so dass* PA *Zeilen-Stufen-Form hat.*

Beweis Nach Satz 2.6 gibt es eine endliche Folge $\beta_1, \ldots, \beta_p$ von elementaren Zeilenumformungen, die A in eine Matrix A' mit Zeilen-Stufen-Form überführt. Für jedes $j = 1, \ldots, p$ sei $B_j \in \mathrm{M}(m \times m, K)$ die Elementarmatrix, die durch Multiplikation von links die Umformung β_j bewirkt. Dann gilt $A' = B_p \cdots B_2 B_1 A$. Setze $P = B_p \cdots B_2 B_1$; PA hat also Zeilen-Stufen-Form und nach Lemma 6.3 (3) und Lemma 6.4 ist P invertierbar. □

Satz 6.5 *Jede invertierbare Matrix ist Produkt von Elementarmatrizen.*

Beweis Sei $A \in \mathrm{M}(n \times n, K)$ invertierbar. Nach Satz 6.3 ((1) $\Leftrightarrow$ (8)) gibt es eine Folge $\beta_1, \ldots, \beta_p$ von elementaren Zeilenumformungen, die A in die Matrix E_n überführt. Für jedes j sei $B_j \in \mathrm{M}(m \times m, K)$ die Elementarmatrix, die durch Multiplikation von links die Umformung β_j bewirkt. Dann gilt $E_n = B_p \cdots B_2 B_1 A$ und folglich ist

$$\begin{aligned} A = E_n A &= (B_p \cdots B_2 B_1)^{-1} B_p \cdots B_2 B_1 A \\ &= (B_p \cdots B_2 B_1)^{-1} E_n = (B_p \cdots B_2 B_1)^{-1} = B_1^{-1} B_2^{-1} \cdots B_p^{-1} \ . \end{aligned}$$

Nach Lemma 6.5 sind aber $B_1^{-1}, B_2^{-1}, \ldots, B_p^{-1}$ auch Elementarmatrizen. □

Die Umkehrung von Satz 6.5 ist natürlich auch richtig: Nach Lemma 6.3 (3) und Lemma 6.4 ist jedes Produkt von Elementarmatrizen invertierbar.

Satz 6.6 *Sei $A \in \mathrm{M}(n \times n, K)$ invertierbar und sei $\beta_1, \ldots, \beta_p$ eine Folge von elementaren Zeilenumformungen, die A in die Einheitsmatrix E_n überführt. Dann überführt die Folge $\beta_1, \ldots, \beta_p$ die Einheitsmatrix E_n in die Matrix A^{-1}.*

Beweis Für jedes j sei B_j die Elementarmatrix, die durch Multiplikation von links die Umformung β_j bewirkt. Dann ist $E_n = B_p \cdots B_2 B_1 A$, und die Folge $\beta_1, \ldots, \beta_p$ überführt die Matrix E_n in die Matrix $B_p \cdots B_2 B_1 E_n$. Aber

$$B_p \cdots B_2 B_1 E_n = B_p \cdots B_2 B_1 = A^{-1} \ ,$$

da $(B_p \cdots B_2 B_1) A = E_n$. □

Im Folgenden sei V ein endlichdimensionaler K-Vektorraum mit $\dim V = n \geq 1$.

Sei $(u_1, \ldots, u_n)$ eine Basis von V. Sind $v_1, \ldots, v_n$ beliebige Vektoren aus V, so gibt es nach Satz 4.1 eine eindeutige Matrix $P = (p_{ij}) \in \mathrm{M}(n \times n, K)$, so dass

$$v_j = p_{1j} u_1 + \cdots + p_{nj} u_n$$

für jedes $j = 1, \ldots, n$. (Für jedes j ist also $p_{1j} u_1 + \cdots + p_{nj} u_n$ die eindeutige Darstellung von v_j als Linearkombination von $u_1, \ldots, u_n$.) Natürlich kommt jede Matrix $P = (p_{ij}) \in \mathrm{M}(n \times n, K)$ auf diese Weise vor, da umgekehrt Vektoren $v_1, \ldots, v_n \in V$ einfach durch $v_j = p_{1j} u_1 + \cdots + p_{nj} u_n$ für $j = 1, \ldots, n$ definiert werden können.

Lemma 6.6 *Sei $(u_1, \ldots, u_n)$ eine Basis von V und seien $v_1, \ldots, v_n \in V$. Sei $P = (p_{ij}) \in \mathrm{M}(n \times n, K)$ die eindeutige Matrix mit*

$$v_j = p_{1j}u_1 + \cdots + p_{nj}u_n$$

für jedes $j = 1, \ldots, n$. Dann ist P invertierbar genau, wenn $(v_1, \ldots, v_n)$ eine Basis von V ist.

Beweis Nach Satz 4.6 (2) ist $(v_1, \ldots, v_n)$ eine Basis von V genau dann, wenn $v_1, \ldots, v_n$ linear unabhängig sind. Seien $\lambda_1, \ldots, \lambda_n \in K$; dann ist

$$\begin{aligned}\sum_{j=1}^{n} \lambda_j v_j &= \sum_{j=1}^{n} \lambda_j \sum_{i=1}^{n} p_{ij}u_i = \sum_{j=1}^{n}\sum_{i=1}^{n} \lambda_j p_{ij}u_i \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_j p_{ij}u_i = \sum_{i=1}^{n}(\sum_{j=1}^{n} \lambda_j p_{ij})u_i = \sum_{i=1}^{n} \mu_i u_i \ ,\end{aligned}$$

wobei $(\mu_1 \ldots, \mu_n) = \varphi_P((\lambda_1, \ldots, \lambda_n))$. Damit ist $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$ genau dann, wenn $\varphi_P((\lambda_1, \ldots, \lambda_n)) = 0$, da die Vektoren $u_1, \ldots, u_n$ linear unabhängig sind. Daraus ergibt sich, dass $v_1, \ldots, v_n$ genau dann linear unabhängig sind, wenn $\operatorname{Kern} \varphi_P = \{0\}$. Folglich gilt nach Lemma 5.6, Satz 5.3 und Satz 6.3, dass $v_1, \ldots, v_n$ genau dann linear unabhängig sind, wenn die Matrix P invertierbar ist. □

Seien nun $(u_1, \ldots, u_n)$, $(v_1, \ldots, v_n)$ Basen von V und sei $P = (p_{ij}) \in \mathrm{M}(n \times n, K)$ die eindeutige Matrix mit $v_j = p_{1j}u_1 + \cdots + p_{nj}u_n$ für jedes j. Dann heißt P *die Matrix für den Wechsel von $(u_1, \ldots, u_n)$ nach $(v_1, \ldots, v_n)$*. Nach Lemma 6.6 ist P invertierbar.

Ist $(u_1, \ldots, u_n)$ eine Basis von V und ist $P = (p_{ij}) \in \mathrm{M}(n \times n, K)$ invertierbar, dann ist nach Lemma 6.6 P die Matrix für den Wechsel von $(u_1, \ldots, u_n)$ nach $(v_1, \ldots, v_n)$, wobei $(v_1, \ldots, v_n)$ die durch $v_j = p_{1j}u_1 + \cdots + p_{nj}u_n$ für $j = 1, \ldots, n$ definierte Basis von V ist.

Für Basen α, β von V sei $P_{\alpha,\beta} \in \mathrm{M}(n \times n, K)$ die Matrix für den Wechsel von α nach β.

Satz 6.7 *Für jede Basis α von V definiert die Abbildung $\beta \mapsto P_{\alpha,\beta}$ eine Bijektion zwischen der Menge der Basen von V und der Menge der $n \times n$ invertierbaren Matrizen über K.*

Beweis Dies folgt unmittelbar aus Lemma 6.6. □

Satz 6.8 *(1) Für jede Basis α von V ist $P_{\alpha,\alpha} = E_n$.*

(2) Für alle Basen α, β von V ist $P_{\beta,\alpha} = P_{\alpha,\beta}^{-1}$.

(3) Für alle Basen α, β, γ von V ist $P_{\alpha,\gamma} = P_{\alpha,\beta}P_{\beta,\gamma}$.

Beweis (1) Dies ist klar, da $v_j = \sum_{i=1}^{n} \delta_{ij} v_i$ für jedes j, wobei $E_n = (\delta_{ij})$.

(3) Seien $\alpha = (u_1, \ldots, u_n)$, $\beta = (v_1, \ldots, v_n)$ und $\gamma = (w_1, \ldots, w_n)$, und setze $P_{\alpha,\beta} = (p_{ij})$, $P_{\beta,\gamma} = (q_{ij})$. Ferner sei $P_{\alpha,\beta}P_{\beta,\gamma} = (r_{ij})$, also ist $r_{ij} = \sum_{k=1}^{n} p_{ik}q_{kj}$ für jedes $1 \leq i, j \leq n$. Für jedes $j = 1, \ldots, n$ ist nun

$$w_j = \sum_{k=1}^{n} q_{kj} v_k = \sum_{k=1}^{n} q_{kj} \sum_{i=1}^{n} p_{ik} u_i = \sum_{i=1}^{n} \sum_{k=1}^{n} p_{ik} q_{kj} u_i = \sum_{i=1}^{n} r_{ij} u_i \ ,$$

und folglich ist $P_{\alpha,\gamma} = (r_{ij}) = P_{\alpha,\beta}P_{\beta,\gamma}$.

(2) Nach (1) und (3) ist $P_{\alpha,\beta}P_{\beta,\alpha} = P_{\alpha,\alpha} = E_n = P_{\beta,\beta} = P_{\beta,\alpha}P_{\alpha,\beta}$ und daraus folgt, dass $P_{\beta,\alpha} = P_{\alpha,\beta}^{-1}$. □

Betrachte nun den speziellen Fall mit $V = K^n$. Für jede Basis $\alpha = (u_1, \ldots, u_n)$ von K^n sei $C_\alpha \in \mathrm{M}(n \times n, K)$ die Matrix, die die n-Tupel $u_1, \ldots, u_n$ als Spalten hat.

Satz 6.9 *Für alle Basen α, β von K^n ist $P_{\alpha,\beta} = C_\alpha^{-1} C_\beta$.*

Beweis Sei $\alpha = (u_1, \ldots, u_n)$ eine Basis von K^n und sei $C_\alpha = (c_{ij})$. Dann ist

$$u_j = (c_{1j}, \ldots, c_{nj}) = c_{1j}e_1 + \cdots + c_{nj}e_n$$

für jedes $j = 1, \ldots, n$, wobei $\epsilon = (e_1, \ldots, e_n)$ die kanonische Basis von K^n ist, und folglich ist C_α die Matrix für den Wechsel von ϵ nach α. Nach Satz 6.8 (2) ist dann C_α^{-1} die Matrix für den Wechsel von α nach ϵ und daraus ergibt sich nach Satz 6.8 (3), dass $C_\alpha^{-1} C_\beta$ die Matrix für den Wechsel von α nach β ist. □

7 Rang einer Matrix

Im folgenden sei K ein Körper und sei $A = (a_{ij}) \in \mathrm{M}(m \times n, K)$ eine $m \times n$ Matrix über K. Seien $u_1, \ldots, u_m \in K^n$ die Zeilen und $v_1, \ldots, v_n \in K^m$ die Spalten von A; es gilt also $u_i = (a_{i1}, \ldots, a_{in})$ für jedes $i = 1, \ldots, m$ und $v_j = (a_{1j}, \ldots, a_{mj})$ für jedes $j = 1, \ldots, n$.

Der Untervektorraum $\mathrm{L}(u_1, \ldots, u_m)$ von K^n heißt der *Zeilenraum* von A und wird mit $\mathrm{ZR}(A)$ bezeichnet. Analog heißt der Untervektorraum $\mathrm{L}(v_1, \ldots, v_n)$ von K^m der *Spaltenraum* von A und wird mit $\mathrm{SR}(A)$ bezeichnet.

Setze $\mathrm{Rang}_Z(A) = \dim \mathrm{ZR}(A)$ und $\mathrm{Rang}_S(A) = \dim \mathrm{SR}(A)$; $\mathrm{Rang}_Z(A)$ heißt der *Zeilenrang* und $\mathrm{Rang}_S(A)$ der *Spaltenrang* von A. Also gilt $0 \leq \mathrm{Rang}_Z(A) \leq n$ und $0 \leq \mathrm{Rang}_S(A) \leq m$, da $\mathrm{ZR}(A)$ bzw. $\mathrm{SR}(A)$ ein Untervektorraum von K^n bzw. von K^m ist.

Ist $A = 0$, so ist $\mathrm{Rang}_Z(A) = \mathrm{Rang}_S(A) = 0$. Ist umgekehrt $\mathrm{Rang}_Z(A) = 0$ oder $\mathrm{Rang}_S(A) = 0$, so ist nach Lemma 3.4 $A = 0$.

Satz 7.1 *Es gilt* $\mathrm{Rang}_Z(A) = \mathrm{Rang}_S(A)$.

Beweis Dies ist trivial richtig, wenn $A = 0$; es kann also angenommen werden, dass $A \neq 0$ und damit ist $\mathrm{ZR}(A) \neq \{0\}$ und $\mathrm{SR}(A) \neq \{0\}$. Das folgende Lemma wird benötigt:

Lemma 7.1 *(1) Gibt es Matrizen $B \in \mathrm{M}(m \times s, K)$ und $C \in \mathrm{M}(s \times n, K)$ mit $A = BC$, so gilt* $\mathrm{Rang}_Z(A) \leq s$.

(2) Ist $(v'_1, \ldots, v'_s)$ eine Basis von $\mathrm{SR}(A) = \mathrm{L}(v_1, \ldots, v_n)$, so gibt es eine Matrix $C \in \mathrm{M}(s \times n, K)$ mit $A = BC$, wobei $B \in \mathrm{M}(m \times s, K)$ die Matrix ist, die $v'_1, \ldots, v'_s$ als Spalten hat.

Beweis (1) Sei $B = (b_{ij})$ und seien $w_1, \ldots, w_s$ die Zeilen von C; da $A = BC$, gilt $u_i = b_{i1}w_1 + \cdots + b_{is}w_s$ für jedes $i = 1, \ldots, m$. Damit ist $u_i \in \mathrm{L}(w_1, \ldots, w_s)$ für jedes i und folglich gilt nach Satz 3.2, dass $\mathrm{L}(u_1, \ldots, u_m) \subset \mathrm{L}(w_1, \ldots, w_s)$. Daraus ergibt sich nach Satz 4.3 und Satz 4.4, dass

$$\mathrm{Rang}_Z(A) = \dim \mathrm{L}(u_1, \ldots, u_m) \leq \dim \mathrm{L}(w_1, \ldots, w_s) \leq s \; .$$

(2) Da $(v'_1, \ldots, v'_s)$ eine Basis von $\mathrm{L}(v_1, \ldots, v_n)$ ist, gibt es für jedes $j = 1, \ldots, n$ eindeutige Elemente $c_{1j}, \ldots, c_{sj} \in K$, so dass $v_j = c_{1j}v'_1 + \cdots + c_{sj}v'_s$ und dann gilt $A = BC$. □

Sei $s = \mathrm{Rang}_S(A)$; nach Lemma 7.1 (2) gibt es dann Matrizen $B \in \mathrm{M}(m \times s, K)$, $C \in \mathrm{M}(s \times n, K)$ mit $A = BC$, und daraus ergibt sich nach Lemma 7.1 (1), dass $\mathrm{Rang}_Z(A) \leq s = \mathrm{Rang}_S(A)$.

Sei nun $A^t = (a^t_{ji}) \in \mathrm{M}(n \times m, K)$ die Matrix mit $a^t_{ji} = a_{ij}$. Dann sind $v_1, \ldots, v_n$ die Zeilen und $u_1, \ldots, u_m$ die Spalten von A^t und folglich ist $\mathrm{ZR}(A^t) = \mathrm{SR}(A)$ und $\mathrm{SR}(A^t) = \mathrm{ZR}(A)$. Daher gilt

$$\mathrm{Rang}_S(A) = \mathrm{Rang}_Z(A^t) \leq \mathrm{Rang}_S(A^t) = \mathrm{Rang}_Z(A) \ .$$

Dies zeigt also, dass $\mathrm{Rang}_Z(A) = \mathrm{Rang}_S(A)$. □

Setze nun $\operatorname{rang} A = \mathrm{Rang}_Z(A)$ $(= \mathrm{Rang}_S(A))$. Man nennt $\operatorname{rang} A$ einfach den *Rang* von A.

Lemma 7.2 *Es gilt* $\operatorname{Bild} \varphi_A = \mathrm{SR}(A)$ *und insbesondere* $\operatorname{rang} \varphi_A = \operatorname{rang} A$.

Beweis Da $\varphi_A((\lambda_1, \ldots, \lambda_n)) = \lambda_1 v_1 + \cdots + \lambda_n v_n$ für alle $(\lambda_1, \ldots, \lambda_n) \in K^n$, gilt

$$\operatorname{Bild} \varphi_A = \varphi_A(K^n) = \mathrm{L}(v_1, \ldots, v_n) = \mathrm{SR}(A)$$

und damit auch $\operatorname{rang} \varphi_A = \dim \operatorname{Bild} \varphi_A = \dim \mathrm{SR}(A) = \operatorname{rang} A$. □

Satz 7.2 *Es gilt* $\dim \mathrm{Lös}(A, 0) = n - \operatorname{rang} A$.

Beweis Nach Satz 5.2, Lemma 5.7 und Lemma 7.2 gilt

$$\dim \mathrm{Lös}(A, 0) = \dim \operatorname{Kern} \varphi_A = \dim K^n - \dim \operatorname{Bild} \varphi_A = n - \operatorname{rang} A \ . \ \square$$

Lemma 7.3 *Wird A durch eine elementare Zeilenumformung zu einer Matrix A' verändert, so gilt* $\mathrm{ZR}(A) = \mathrm{ZR}(A')$ *(und damit* $\operatorname{rang} A = \operatorname{rang} A'$*).*

Beweis Nach Lemma 3.6 gilt $\mathrm{ZR}(A) = \mathrm{ZR}(A')$, wenn A' aus A durch Vertauschen zweier Zeilen entsteht. Entsteht A' aus A durch eine Umformung vom Typ III, so gilt $\mathrm{ZR}(A) = \mathrm{ZR}(A')$ nach Lemma 3.5. Es kann also angenommen weden, dass A' durch Addition des λ-fachen der k-ten Zeile zu der ℓ-ten Zeile von A entsteht, (wobei $\lambda \in K$ und $k \neq \ell$). Seien $u_1, \ldots, u_m$ die Zeilen von A und $u'_1, \ldots, u'_m$ die Zeilen von A'. Dann gilt $u'_i = u_i$ für jedes $i \neq \ell$ und $u'_\ell = u_\ell + \lambda u_k$. Daraus folgt nach Satz 3.1 und Satz 3.2, dass

$$\mathrm{ZR}(A) = \mathrm{L}(u_1, \ldots, u_m) = \mathrm{L}(u'_1, \ldots, u'_m) = \mathrm{ZR}(A') \ ,$$

da $u_\ell = u'_\ell + (-\lambda)u'_k$. □

Satz 7.3 *Sei $A \neq 0$ eine Matrix, die Zeilen-Stufen-Form hat und sei q der Index der letzten von Null verschiedenen Zeile von A. Dann gilt* $\operatorname{rang} A = q$.

Beweis Sei $A = (a_{ij})$ und für $i = 1, \ldots, m$ sei

$$k_i = \begin{cases} \min\{1 \le j \le n : a_{ij} \neq 0\} & \text{falls } (a_{i1}, \ldots, a_{in}) \neq 0 \ , \\ 0 & \text{sonst} \ . \end{cases}$$

Da die Matrix A Zeilen-Stufen-Form hat, gilt $k_1 < k_2 < \cdots < k_q$.

Seien $u_1, \ldots, u_m$ die Zeilen von A. Dann ist

$$\mathrm{ZR}(A) = \mathrm{L}(u_1, \ldots, u_m) = \mathrm{L}(u_1, \ldots, u_q) \ ,$$

da $u_i = 0$, wenn $i > q$. Aber die Vektoren $u_1, \ldots, u_q$ sind linear unabhängig: Seien $\lambda_1, \ldots, \lambda_q \in K$ mit $\lambda_1 u_1 + \cdots + \lambda_q u_q = 0$. Dann gilt $\lambda_1 a_{1j} + \cdots + \lambda_q a_{qj} = 0$ für jedes $j = 1, \ldots, n$ und insbesondere gilt $\lambda_1 a_{1k_j} + \cdots + \lambda_q a_{qk_j} = 0$ für jedes $j = 1, \ldots, q$. Nun ist $a_{1k_1} \neq 0$ und $a_{ik_1} = 0$ für $i > 1$ und damit ist $\lambda_1 = 0$. Nehme an, dass $\lambda_\ell = 0$ für $\ell = 1, \ldots, j-1$ für ein j mit $1 < j < q$. Dann ist auch $\lambda_j = 0$, da $a_{jk_j} \neq 0$ und $a_{ik_j} = 0$ für $i > j$. Es gilt also $\lambda_j = 0$ für jedes $j = 1, \ldots, q$ und daher sind $u_1, \ldots, u_q$ linear unabhängig. Daraus folgt, dass $(u_1, \ldots, u_q)$ eine Basis von $\mathrm{ZR}(A)$ ist und insbesondere ist $\mathrm{Rang}_Z(A) = q$.

Bemerkung: Sind $v_1, \ldots, v_n$ die Spalten von A, so ist $(v_{k_1}, \ldots, v_{k_q})$ eine Basis von $\mathrm{SR}(A)$:. Sei A' die $m \times q$ Matrix über K, die $v_{k_1}, \ldots, v_{k_q}$ als Spalten hat. Dann hat A' Zeilen-Stufen-Form und 1, 2, …, q ist die Treppen-Folge von A'. Daraus folgt nach Satz 2.5, dass das zu A' und 0 gehörige lineare Gleichungssystem eindeutig lösbar ist, und damit sind nach Satz 3.5 $v_{k_1}, \ldots, v_{k_q}$ linear unabhängig. Aber für jedes j ist das zu A' und v_j gehörige lineare Gleichungssystem lösbar und daraus ergibt sich nach Satz 3.3, dass $v_j \in \mathrm{L}(v_{k_1}, \ldots, v_{k_q})$. Nach Satz 3.2 ist dann $\mathrm{SR}(A) = \mathrm{L}(v_{k_1}, \ldots, v_{k_q})$. Daher ist $(v_{k_1}, \ldots, v_{k_q})$ eine Basis von $\mathrm{SR}(A)$. □

Lemma 7.4 *Seien $A \in \mathrm{M}(\ell \times m, K)$, $B \in \mathrm{M}(m \times n, K)$. Dann gilt:*

(1) $\operatorname{rang} AB \le \min\{\operatorname{rang} A, \operatorname{rang} B\}$.

(2) $\operatorname{rang} AB = \operatorname{rang} A$, *falls* $\operatorname{rang} B = m$.

(3) $\operatorname{rang} AB = \operatorname{rang} B$, *falls* $\operatorname{rang} A = m$.

Beweis Nach Lemmas 6.1 und 7.2 gilt $\operatorname{rang} A = \operatorname{rang} \varphi_A$, $\operatorname{rang} B = \operatorname{rang} \varphi_B$ und $\operatorname{rang} AB = \operatorname{rang} \varphi_{AB} = \operatorname{rang} \varphi_A \circ \varphi_B$.

(1) Nach Lemma 5.12 (1) gilt also

$$\operatorname{rang} AB = \operatorname{rang} \varphi_A \circ \varphi_B \le \min\{\operatorname{rang} \varphi_A, \operatorname{rang} \varphi_B\} = \min\{\operatorname{rang} A, \operatorname{rang} B\} \ .$$

(2) Ist $\operatorname{rang} B = m$, so gilt $\dim \varphi_B(K^n) = \operatorname{rang} \varphi_B = \operatorname{rang} B = m$ und damit ist φ_B surjektiv. Daraus ergibt sich nach Lemma 5.12 (2), dass

$$\operatorname{rang} AB = \operatorname{rang} \varphi_A \circ \varphi_B = \operatorname{rang} \varphi_A = \operatorname{rang} A \ .$$

(3) Ist $\operatorname{rang} A = m$, so gilt nach Satz 5.2, dass

$$\dim \operatorname{Kern} \varphi_B = \dim K^m - \operatorname{rang} \varphi_A = m - \operatorname{rang} A = 0$$

und damit ist φ_A injektiv. Daraus folgt nach Lemma 5.10 (3), dass

$$\operatorname{rang} AB = \operatorname{rang} \varphi_A \circ \varphi_B = \operatorname{rang} \varphi_B = \operatorname{rang} B \; . \; \square$$

Satz 7.4 *Eine Matrix $A \in \mathrm{M}(n \times n, K)$ ist invertierbar genau dann, wenn sie vollen Rang hat, d.h., wenn* $\operatorname{rang} A = n$.

Beweis Dies folgt aus Satz 6.3 ((1) $\Leftrightarrow$ (5)), da nach Satz 7.2 $\operatorname{rang} A = n$ genau dann gilt, wenn $\mathrm{Lös}(A, 0) = \{0\}$. $\square$

Erinnerung: Unter elementaren Zeilenumformungen werden nun seit Kapitel 6 die folgenden Operationen verstanden:

(I) Addition eines Vielfachen einer Zeile zu einer anderen Zeile.

(II) Vertauschen zweier Zeilen.

(III) Multiplikation einer Zeile mit einem Skalar $\lambda \in K$, $\lambda \neq 0$.

Es gibt auch natürlich entsprechende Spaltenumformungen. Unter *elementaren Spaltenumformungen* einer Matrix versteht man die folgenden Operationen:

(I) Addition eines Vielfachen einer Spalte zu einer anderen Spalte.

(II) Vertauschen zweier Spalten.

(III) Multiplikation einer Spalte mit einem Skalar $\lambda \in K$, $\lambda \neq 0$.

Unter einer *elementaren Umformung* einer Matrix versteht man eine elementare Zeilenumformung oder eine elementare Spaltenumformung.

Lemma 7.5 *Elementare Umformungen ändern den Rang einer Matrix nicht.*

Beweis Dies folgt aus Lemmas 7.2 und 3.5 und den entsprechenden Ergebnissen für elementare Spaltenumformungen. $\square$

Sei q mit $0 \leq q \leq \min\{m, n\}$; die folgende $m \times n$ Matrix (die Rang q hat)

$$E^q_{m,n} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \leftarrow q\text{-te Zeile}$$

wird die *Sub-Einheitsmatrix vom Rang q* genannt. Also ist $E^q_{m,n} = (\delta_{ij})$, wobei $\delta_{ij} = 0$ für alle $i \neq j$, $\delta_{ii} = 1$ für $i = 1, \ldots, q$ und $\delta_{ii} = 0$ für alle $i > q$.

Lemma 7.6 *Sei $B \in \mathrm{M}(m \times n, K)$ eine Matrix, die Zeilen-Stufen-Form hat. Dann gibt es eine Folge von elementaren Spaltenumformungen, die die Matrix B in die Sub-Einheitsmatrix vom Rang q überführt, wobei $q = \operatorname{rang} B$.*

Beweis Übung. □

Satz 7.5 *Jede Matrix $A \in \mathrm{M}(m \times n, K)$ vom Rang q läßt sich durch eine Folge von elementaren Umformungen in die Sub-Einheitsmatrix vom Rang q überführen.*

Beweis Nach Satz 2.6 gibt es eine Folge von elementaren Zeilenumformungen, die A in eine Matrix B mit Zeilen-Stufen-Form überführt und nach Lemma 7.5 ist dann $\operatorname{rang} B = q$. Nun gibt es nach Lemma 7.6 eine Folge von elementaren Spaltenumformungen, die B in die Sub-Einheitsmatrix vom Rang q überführt. □

Erinnerung: Multiplikation von links mit einer Elementarmatrix bewirkt eine elementare Zeilenumformung. Analog bewirkt nun Multiplikation von rechts mit einer Elementarmatrix eine elementare Spaltenumformung und jede elementare Spaltenumformung entsteht auf diese Weise.

Satz 7.6 *Sei $A \in \mathrm{M}(m \times n, K)$ eine $m \times n$ Matrix mit $\operatorname{rang} A = q$. Dann gibt es invertierbare Matrizen $P \in \mathrm{M}(m \times m, K)$ und $Q \in \mathrm{M}(n \times n, K)$, so dass PAQ die Sub-Einheitsmatrix vom Rang q ist, d.h.,*

$$PAQ = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \leftarrow q\text{-te Zeile}$$

Beweis Nach Satz 6.4 gibt es eine invertierbare Matrix $P \in \mathrm{M}(m \times m, K)$, so dass $B = PA$ Zeilen-Stufen Form hat, und nach Lemma 7.4 ist $\operatorname{rang} B = q$. Ferner gibt es nach Lemma 7.5 eine Folge $\gamma_1, \ldots, \gamma_p$ von elementaren Spaltenumformungen, die B in die Sub-Einheitsmatrix $E^q_{m,n}$ vom Rang q überführt. Für jedes j sei $C_j \in \mathrm{M}(n \times n, K)$ die Elementarmatrix, die durch Multiplikation von rechts die Umformung γ_j bewirkt. Dann gilt $E^q_{m,n} = BC_1 \cdots C_p$. Setze $Q = C_1 \cdots C_p$; nach Lemmas 6.3 (3) und 6.4 ist also Q invertierbar und $PAQ = BQ = E^q_{m,n}$. □

8 Lineare Abbildungen und Matrizen

Im folgenden sei K ein Körper und seien V, W endlichdimensionale Vektorräume über K mit $\dim V = n \geq 1$ und $\dim W = m \geq 1$. Sei auch $(v_1, \ldots, v_n)$ eine Basis von V und $(w_1, \ldots, w_m)$ eine Basis von W.

Sei nun $f \in \mathrm{Hom}(V, W)$. Für jedes $j = 1, \ldots, n$ ist $f(v_j)$ ein Element von W; nach Satz 4.1 gibt es also genau ein m-Tupel $(a_{1j}, a_{2j}, \ldots, a_{mj}) \in K^m$, so dass

$$f(v_j) = a_{1j}w_1 + a_{2j}w_2 + \cdots + a_{mj}w_m \ .$$

Auf diese Weise wird eine Matrix $A = (a_{ij}) \in \mathrm{M}(m \times n, K)$ definiert, die die *Matrix von f bezüglich der Basen $(v_1, \ldots, v_n)$ und $(w_1, \ldots, w_m)$* heißt.

Lemma 8.1 *Für jedes $A \in \mathrm{M}(m \times n, K)$ ist A die Matrix von φ_A bezüglich der kanonischen Basen von K^n und K^m.*

Beweis Sei $(e_1, \ldots, e_n)$ bzw. $(e'_1, \ldots, e'_m)$ die kanonische Basis von K^n bzw. von K^m. Für jedes $j = 1, \ldots, n$ gilt dann

$$\varphi_A(e_j) = (a_{1j}, \ldots, a_{mj}) = a_{1j}e'_1 + \cdots + a_{mj}e'_m \ ,$$

und daraus ergibt sich, dass A die Matrix von φ_A bezüglich der Basen $(e_1, \ldots, e_n)$ und $(e'_1, \ldots, e'_m)$ ist. □

Lemma 8.2 *Sei $A \in \mathrm{M}(m \times n, K)$. Dann gibt es ein eindeutiges $f \in \mathrm{Hom}(V, W)$, so dass A die Matrix von f bezüglich der Basen $(v_1, \ldots, v_n)$ und $(w_1, \ldots, w_m)$ ist.*

Beweis Nach Lemma 5.11 gibt es ein eindeutiges $f \in \mathrm{Hom}(V, W)$, so dass

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

für jedes $j = 1, \ldots, n$, wobei $A = (a_{ij})$. Offensichtlich ist f dann die eindeutige lineare Abbildung, so dass A die Matrix von f bezüglich der Basen $(v_1, \ldots, v_n)$ und $(w_1, \ldots, w_m)$ ist. □

Sei U ein weiterer endlichdimensionaler Vektorraum von V und sei $(u_1, \ldots, u_p)$ eine Basis von U.

Satz 8.1 *Seien $f \in \mathrm{Hom}(V, W)$, $g \in \mathrm{Hom}(U, V)$, A die Matrix von f bezüglich der Basen $(v_1, \ldots, v_n)$ und $(w_1, \ldots, w_m)$ und B die Matrix von g bezüglich der Basen $(u_1, \ldots, u_p)$ und $(v_1, \ldots, v_n)$. Dann ist AB die Matrix von $f \circ g$ bezüglich der Basen $(u_1, \ldots, u_p)$ und $(w_1, \ldots, w_m)$.*

Beweis Seien $A = (a_{ij})$, $B = (b_{jk})$. Für jedes $k = 1, \ldots, p$ gilt dann

$$\begin{aligned}(f \circ g)(u_k) = f(g(u_k)) = f\Big(\sum_{j=1}^{n} b_{jk} v_j\Big) &= \sum_{j=1}^{n} b_{jk} f(v_j) \\ = \sum_{j=1}^{n} b_{jk} \sum_{i=1}^{m} a_{ij} w_i = \sum_{i=1}^{m} \Big(\sum_{j=1}^{n} a_{ij} b_{jk}\Big) w_i &= \sum_{i=1}^{m} c_{ik} w_i \ ,\end{aligned}$$

wobei $AB = (c_{ik})$. Damit ist AB die Matrix von $f \circ g$ bezüglich der Basen $(u_1, \ldots, u_p)$ und $(w_1, \ldots, w_m)$. □

Setze nun $\alpha = (v_1, \ldots, v_n)$ und $\beta = (w_1, \ldots, w_m)$.

Satz 8.2 *Sei $\alpha' = (v'_1, \ldots, v'_n)$ eine weitere Basis von V und $\beta' = (w'_1, \ldots, w'_m)$ eine weitere Basis von W, sei $f \in \mathrm{Hom}(V, W)$, sei A die Matrix von f bezüglich α und β und B die Matrix von f bezüglich α' und β'. Dann gilt*

$$B = Q^{-1} A P \ ,$$

wobei P die Matrix für den Wechsel von α nach α' und Q die Matrix für den Wechsel von β nach β' ist.

Beweis Sei $P = (p_{kj})$, $Q = (q_{i\ell})$, $A = (a_{ik})$ und $B = (b_{\ell j})$. Für jedes $j = 1, \ldots, n$ gilt dann

$$f(v'_j) = f\Big(\sum_{k=1}^{n} p_{kj} v_k\Big) = \sum_{k=1}^{n} p_{kj} f(v_k) = \sum_{k=1}^{n} p_{kj} \Big(\sum_{i=1}^{m} a_{ij} w_i\Big) = \sum_{i=1}^{m} \Big(\sum_{k=1}^{n} a_{ik} p_{kj}\Big) w_i$$

sowie auch

$$f(v'_j) = \sum_{\ell=1}^{m} b_{\ell j} w'_\ell = \sum_{\ell=1}^{m} b_{\ell j} \Big(\sum_{i=1}^{m} q_{i\ell} w_i\Big) = \sum_{\ell=1}^{m} \sum_{i=1}^{m} b_{\ell j} q_{i\ell} w_i = \sum_{i=1}^{m} \Big(\sum_{\ell=1}^{m} q_{i\ell} b_{\ell j}\Big) w_i \ .$$

Daraus ergibt sich, dass für jedes $j = 1, \ldots, n$

$$\begin{aligned}\sum_{i=1}^{m} \Big(\sum_{\ell=1}^{m} q_{i\ell} b_{\ell j} - \sum_{\ell=1}^{m} a_{ik} p_{kj}\Big) w_i &= \sum_{i=1}^{m} \Big(\sum_{\ell=1}^{m} q_{i\ell} b_{\ell j}\Big) w_i - \sum_{i=1}^{m} \sum_{\ell=1}^{m} \Big(a_{ik} p_{kj}\Big) w_i \\ &= f(v'_j) - f(v'_j) = 0 \ ,\end{aligned}$$

und für jedes $i = 1, \ldots, m$ gilt also

$$\sum_{\ell=1}^{m} q_{i\ell} b_{\ell j} = \sum_{\ell=1}^{m} a_{ik} p_{kj} \ ,$$

da $w_1, \ldots, w_m$ linear unabhängig sind. Dies bedeutet aber genau, dass $QB = AP$, und damit ist $B = P^{-1} A P$. □

Satz 8.3 *Sei $f \in \mathrm{Hom}(V,W)$ und sei A die Matrix von f bezüglich der Basen α und β. Dann gilt* $\operatorname{rang} f = \operatorname{rang} A$.

Beweis Sei $\psi : K^m \to W$ die Abbildung, die definiert ist durch

$$\psi((\lambda_1, \ldots, \lambda_m)) = \lambda_1 w_1 + \cdots + \lambda_m w_m \ .$$

Man sieht leicht, dass ψ linear ist. Ist $(e_1, \ldots, e_m)$ die kanonische Basis von K^m, so gilt $\psi(e_j) = w_j$ für jedes j und daraus folgt nach Lemma 5.10 (4), dass ψ ein Isomorphismus ist. Für jedes $(\lambda_1, \ldots, \lambda_n) \in K^n$ gilt nun

$$\begin{aligned} f\Big(\sum_{j=1}^{n} \lambda_j v_j\Big) = \sum_{j=1}^{n} \lambda_j f(v_j) = \sum_{j=1}^{n} \lambda_j \sum_{i=1}^{m} a_{ij} w_i = \sum_{i=1}^{m} \Big(\sum_{j=1}^{n} a_{ij}\lambda_j\Big) w_i \\ = \psi(\varphi_A((\lambda_1, \ldots, \lambda_n))) \ , \end{aligned}$$

und daraus ergibt sich, dass $\operatorname{Bild} f = \psi(\operatorname{Bild} \varphi_A)$, da $(v_1, \ldots, v_n)$ eine Basis von V ist. Nach Lemma 5.8 und Lemma 5.10 (1) gilt aber $\dim U = \dim \psi(U)$ für jeden Untervektorraum U von K^m und folglich ist nach Lemma 7.1

$$\operatorname{rang} A = \operatorname{rang} \varphi_A = \dim \operatorname{Bild} \varphi_A = \dim \psi(\operatorname{Bild} \varphi_A) = \dim \operatorname{Bild} f = \operatorname{rang} f \ . \quad \square$$

Satz 8.4 *Sei $f \in \mathrm{Hom}(V,W)$ und $A \in \mathrm{M}(m \times n, K)$. Dann sind äquivalent:*

(1) $\operatorname{rang} f = \operatorname{rang} A$.

(2) Es gibt eine Basis α von V und eine Basis β von W, so dass A die Matrix von f bezüglich α und β ist.

Beweis (2) $\Rightarrow$ (1): Dies folgt unmittelbar aus Satz 8.3.

(1) $\Rightarrow$ (2): Sei α' eine beliebige Basis von V, sei β' eine beliebige Basis von W und sei B die Matrix von f bezüglich α' und β'. Nach Satz 8.3 ist dann $\operatorname{rang} f = \operatorname{rang} B$, d.h., $\operatorname{rang} A = \operatorname{rang} B$. Nun gibt es nach Satz 7.6 invertierbare Matrizen $P_1 \in \mathrm{M}(m \times m, K)$ und $Q_1 \in \mathrm{M}(n \times n, K)$, so dass $P_1 A Q_1 = E^q_{m,n}$, wobei $q = \operatorname{rang} A$ und $E^q_{m,n} \in \mathrm{M}(m \times n, K)$ die Sub-Einheitsmatrix vom Rang q ist. Genauso gibt es invertierbare Matrizen $P_2 \in \mathrm{M}(m \times m, K)$ und $Q_2 \in \mathrm{M}(n \times n, K)$, so dass $P_2 B Q_2 = E^q_{m,n}$. Damit ist

$$\begin{aligned} A = E_m A E_n &= (P_1^{-1}P_1)A(Q_1Q_1^{-1}) = P_1^{-1}(P_1AQ_1)Q_1^{-1} \\ &= P_1^{-1}E^q_{m,n}Q_1^{-1} = P_1^{-1}(P_2BQ_2)Q_1^{-1} \\ &= (P_1^{-1}P_2)B(Q_2Q_1^{-1}) = (P_2P_1^{-1})^{-1}B(Q_2Q_1^{-1}) = P^{-1}BQ \ , \end{aligned}$$

wobei $P = P_2^{-1}P_1$ und $Q = Q_2Q_1^{-1}$. Es gibt aber nach Lemma 6.6 eine Basis α von V, so dass P die Matrix für den Wechsel von α' nach α ist und es gibt eine Basis β von W, so dass Q die Matrix für den Wechsel von β' nach β ist. Nach Satz 8.2 ist $A = P^{-1}BQ$ dann die Matrix von f bezüglich der Basen α und β. $\square$

Satz 8.5 *Ist $f \in \mathrm{Hom}(V,W)$ mit $\operatorname{rang} f = q$, so gibt es eine Basis $(v_1,\dots,v_n)$ von V und eine Basis $(w_1,\dots,w_m)$ von W mit $f(v_j) = w_j$ für $j = 1,\dots,q$ und $f(v_j) = 0$ für $j = q+1,\dots,n$.*

Beweis Sei $E^q_{m,n} \in \mathrm{M}(m\times n,K)$ die Sub-Einheitsmatrix vom Rang q; dann ist $\operatorname{rang} E^q_{m,n} = q$ und folglich gibt es nach Satz 8.4 eine Basis $(v_1,\dots,v_n)$ von V und eine Basis $(w_1,\dots,w_m)$ von W, so dass $E^q_{m,n}$ die Matrix von f bezüglich der Basen $(v_1,\dots,v_n)$ und $(w_1,\dots,w_m)$ ist. Dies bedeutet aber genau, dass $f(v_j) = w_j$ für $j = 1,\dots,q$ und $f(v_j) = 0$ für $j = q+1,\dots,n$. □

Seien V, W Vektorräume über K; für f, $g \in \mathrm{Hom}(V,W)$ und $\lambda \in K$ sind die Abbildungen $f+g$ und λf auch linear, d.h., $f+g$, $\lambda f \in \mathrm{Hom}(V,W)$. Mit diesen Verknüpfungen ist $\mathrm{Hom}(V,W)$ ein Vektorraum über K.

Seien nun $A = (a_{ij})$, $B = (b_{ij}) \in \mathrm{M}(m\times n,K)$ und $\lambda \in K$. Definiere Matrizen $A+B$, $\lambda A \in \mathrm{M}(m\times n,K)$ durch $A+B = (a_{ij}+b_{ij})$ und $\lambda A = (\lambda a_{ij})$. Mit diesen Verknüpfungen ist $\mathrm{M}(m\times n,K)$ auch ein Vektorraum über K.

Satz 8.6 *$\mathrm{M}(m\times n,K)$ ist endlichdimensional und $\dim \mathrm{M}(m\times n,K) = mn$.*

Beweis Übung. □

Seien V, W endlichdimensionale Vektorräume über K, sei $\alpha = (v_1,\dots,v_n)$ eine Basis von V und $\beta = (w_1,\dots,w_m)$ eine Basis von W. Für jede lineare Abbildung $f : V \to W$ sei $M_{\alpha,\beta}(f)$ die Matrix von f bezüglich der Basen α und β; es gibt also eine Abbildung $M_{\alpha,\beta} : \mathrm{Hom}(V,W) \to \mathrm{M}(m\times n,K)$.

Satz 8.7 *Die Abbildung $M_{\alpha,\beta} : \mathrm{Hom}(V,W) \to \mathrm{M}(m\times n,K)$ ist linear und eine Bijektion, d.h., $M_{\alpha,\beta}$ ist ein Isomorphismus.*

Beweis Man sieht leicht, dass $M_{\alpha,\beta}$ linear ist und nach Lemma 8.2 ist $M_{\alpha,\beta}$ eine Bijektion. □

Nach Satz 5.4, Satz 8.6 und Satz 8.7 ist $\mathrm{Hom}(V,W)$ endlichdimensional und

$$\dim \mathrm{Hom}(V,W) = \dim V \times \dim W\ .$$

Lemma 8.3 *(1) Seien U, V, W Vektorräume über K und f, $f' \in \mathrm{Hom}(U,V)$, g, $g' \in \mathrm{Hom}(V,W)$. Dann gilt $g\circ(f+f') = g\circ f + g\circ f'$, $(g+g')\circ f = g\circ f + g'\circ f$.*

(2) Es gilt $A(B+B') = AB+AB'$ und $(A+A')B = AB+A'B$ für alle Matrizen A, $A' \in \mathrm{M}(\ell\times m,K)$ und B, $B' \in \mathrm{M}(m\times n,K)$.

Beweis Übung. □

In der Linearen Algebra ist das Interesse hauptsächlich für Endomorphismen, d.h., für den Fall mit $W = V$. Im Folgenden sei V ein endlichdimensionaler Vektorraum über K mit $\dim V = n \geq 1$.

Sei $f : V \to V$ ein Endomorphismus von V und $(v_1, \ldots, v_n)$ eine Basis von V, und sei $A \in \mathrm{M}(n \times n, K)$ die Matrix von f bezüglich der Basen $(v_1, \ldots, v_n)$ und $(v_1, \ldots, v_n)$. Diese Matrix $A = (a_{ij})$ ist also durch die Bedingungen

$$f(v_j) = a_{1j}v_1 + a_{2j}v_2 + \cdots + a_{nj}v_n$$

für $j = 1, \ldots, n$ definiert und wird einfach die *Matrix von f bezüglich der Basis* $(v_1, \ldots, v_n)$ genannt.

Sei $A \in \mathrm{M}(n \times n, K)$; nach Lemma 8.1 ist A die Matrix von φ_A bezüglich der kanonischen Basis von K^n.

Satz 8.8 *Seien α, α' Basen von V, sei $f : V \to V$ ein Endomorphismus von V, sei A die Matrix von f bezüglich der Basis α und A' die Matrix von f bezüglich der Basis α'. Dann gilt*

$$A' = P^{-1}AP \,,$$

wobei P die Matrix für den Wechsel von α nach α' ist.

Beweis Dies ist ein Spezialfall von Satz 8.2. □

Sei X eine Menge; eine Relation $\equiv$ auf X heißt *Äquivalenzrelation*, wenn gilt:

(1) $x \equiv x$ für alle $x \in X$ ($\equiv$ ist *reflexiv*).

(2) $x_2 \equiv x_1$ für alle x_1, $x_2 \in X$ mit $x_1 \equiv x_2$ ($\equiv$ ist *symmetrisch*).

(3) Sind x_1, x_2, $x_3 \in X$ mit $x_1 \equiv x_2$ und $x_2 \equiv x_3$, so ist $x_1 \equiv x_3$ ($\equiv$ ist *transitiv*).

Seien nun A, $B \in \mathrm{M}(n \times n, K)$; man schreibt $A \sim B$, wenn es eine invertierbare Matrix $P \in \mathrm{M}(n \times n, K)$ gibt, so dass $B = P^{-1}AP$.

Lemma 8.4 *$\sim$ ist eine Äquivalenzrelation auf der Menge $\mathrm{M}(n \times n, K)$.*

Beweis (1) Für jedes $A \in \mathrm{M}(n \times n, K)$ gilt $A \sim A$, da $A = E_n^{-1}AE_n$.

(2) Seien A, $B \in \mathrm{M}(n \times n, K)$ mit $A \sim B$, es gibt also eine invertierbare Matrix $P \in \mathrm{M}(n \times n, K)$, so dass $B = P^{-1}AP$. Dann gilt $A = Q^{-1}BQ$ mit $Q = P^{-1}$ und damit ist $B \sim A$.

(3) Seien A, B, $C \in \mathrm{M}(n\times n, K)$ mit $A \sim B$ und $B \sim C$, es gibt also invertierbare Matrizen P, $Q \in \mathrm{M}(n \times n, K)$, so dass $B = P^{-1}AP$ und $C = Q^{-1}BQ$. Da

$$C = Q^{-1}BQ = Q^{-1}(P^{-1}AP)Q = (Q^{-1}P^{-1})A(PQ) = (PQ)^{-1}A(PQ) \ ,$$

ist dann $C = R^{-1}AR$ mit $R = PQ$ und damit ist $A \sim C$. □

Gilt $A \sim B$, so sagt man, dass die Matrizen A und B *ähnlich* sind.

Satz 8.9 *Sei $f : V \to V$ ein Endomorphismus von V und α eine Basis von V; sei ferner A die Matrix von f bezüglich der Basis α und sei $B \in \mathrm{M}(n \times n, K)$ eine beliebige Matrix. Dann gibt es eine Basis β von V, so dass B die Matrix von f bezüglich β ist, genau, wenn die Matrizen A und B ähnlich sind.*

Beweis Ist B die Matrix von f bezüglich einer Basis β, so gilt nach Satz 8.8, dass $B = P^{-1}AP$, wobei P die Matrix für den Wechsel von α nach β ist, und damit sind die Matrizen A und B ähnlich. Seien umgekehrt A und B ähnlich, es gibt also eine invertierbare Matrix $P \in \mathrm{M}(n \times n, K)$, so dass $B = P^{-1}AP$. Nach Lemma 6.6 gibt es nun eine Basis β von V, so dass P die Matrix für den Wechsel von α nach β ist, und nach Satz 8.8 ist $B = P^{-1}AP$ dann die Matrix von f bezüglich der Basis β. □

Sei $f : V \to V$ ein Endomorphismus. Ein Hauptziel der Linearen Algebra ist es, eine vernüftige Antwort zur folgenden Frage zu finden:

— Wie wählt man eine Basis von V, so dass die Matrix von f bezüglich dieser Basis 'so einfach wie möglich' ist?

— Sei $B \in \mathrm{M}(n \times n, K)$. Wie prüft man, ob es eine Basis von V gibt, so dass B die Matrix von f bezüglich dieser Basis ist?

Nach Satz 8.9 hätte man eine Antwort zu diesen Fragen, wenn die folgenden Fragen beantwortet werden könnten:

— Sei $A \in \mathrm{M}(n \times n, K)$. Was ist die 'einfachste' Matrix, die zu A ähnlich ist?

— Seien A, $B \in \mathrm{M}(n \times n, K)$. Wie prüft man, ob A und B ähnlich sind, und wie findet man gegebenenfalls eine invertierbare Matrix P mit $B = P^{-1}AP$?

(Wähle irgendeine Basis $\alpha = (u_1, \ldots, u_n)$ von V und sei $A \in \mathrm{M}(n \times n, K)$ die Matrix von f bezüglich α. Sei A' die 'einfachste' Matrix, die zu A ähnlich ist, und finde eine invertierbare Matrix $P = (p_{ij}) \in \mathrm{M}(n \times n, K)$ mit $A' = P^{-1}AP$. Für $j = 1, \ldots, n$ setze $v_j = \sum_{i=1}^{n} p_{ij}u_i$; nach Lemma 6.6 ist also $\beta = (v_1, \ldots, v_n)$ eine Basis von V und P ist die Matrix für den Wechsel von α nach β. Nach Satz 8.8 ist dann $A' = P^{-1}AP$ die Matrix von f bezüglich der Basis β; d.h., die Matrix von f bezüglich der Basis β ist 'so einfach wie möglich'. Sei nun $B \in \mathrm{M}(n \times n, K)$; nach Satz 8.9 gibt es eine Basis von V, so dass B die Matrix von f bezüglich dieser Basis ist, genau dann, wenn A und B ähnlich sind.)

9 Direkte Summen

Sei V ein Vektorraum über einem Körper K, sei $m \geq 2$ und seien $U_1, \ldots, U_m$ Untervektoräume von V.

Satz 9.1 *Äquivalent sind:*

(1) Für jedes $j = 1, \ldots, m$ ist $U_j \cap (U_1 + \cdots + U_{j-1} + U_{j+1} + \cdots + U_m) = \{0\}$.

(2) Gilt $u_1 + \cdots + u_m = 0$, wobei $u_j \in U_j$ für $j = 1, \ldots, m$, so ist $u_j = 0$ für jedes $j = 1, \ldots, m$.

(3) Gilt $u_1 + \cdots + u_m = u'_1 + \cdots + u'_m$, wobei $u_j, u'_j \in U_j$ für $j = 1, \ldots, m$, so ist $u_j = u'_j$ für jedes $j = 1, \ldots, m$.

Beweis (1) $\Rightarrow$ (2): Für jedes $j = 1, \ldots, m$ sei $u_j \in U_j$. Gilt $u_1 + \cdots + u_m = 0$, dann ist für jedes $j = 1, \ldots, m$

$$u_j = -(u_1 + \cdots + u_{j-1} + u_{j+1} + \cdots + u_m) \in U_1 + \cdots + U_{j-1} + U_{j+1} + \cdots + U_m$$

und daraus folgt nach (1), dass $u_j = 0$.

(2) $\Rightarrow$ (3): Seien $u_j, u'_j \in U_j$ mit $u_1 + \cdots + u_m = u'_1 + \cdots + u'_m$. Dann ist $u_j - u'_j \in U_j$ für jedes j und $(u_1 - u'_1) + \cdots + (u_m - u'_m) = 0$. Folglich gilt nach (2), dass $u_j - u'_j = 0$ und damit $u_j = u'_j$ für jedes $j = 1, \ldots, m$.

(3) $\Rightarrow$ (1): Sei $u_j \in U_j \cap (U_1 + \cdots + U_{j-1} + U_{j+1} + \cdots + U_m)$. Dann gibt es für jedes $k \neq j$ ein $u_k \in U_k$, so dass $u_k = u_1 + \cdots + u_{j-1} + u_{j+1} + \cdots + u_m$. Damit ist

$$0 + \cdots + 0 + u_k + 0 + \cdots + 0 = u_1 + \cdots + u_{j-1} + 0 + u_{j+1} + \cdots + u_m$$

und daraus folgt nach (3), dass $u_i = 0$ für alle $i = 1, \ldots, m$. Insbesondere ist $u_j = 0$, d.h., $U_j \cap (U_1 + \cdots + U_{j-1} + U_{j+1} + \cdots + U_m) = \{0\}$. □

Setze $U = U_1 + \cdots + U_m$; dann heißt U *direkte Summe* von $U_1, \ldots, U_m$, wenn $U_j \cap (U_1 + \cdots + U_{j-1} + U_{j+1} + \cdots + U_m) = \{0\}$ für jedes $j = 1, \ldots, m$, und in diesem Fall schreibt man $U = U_1 \oplus \cdots \oplus U_m$.

Ist $U = U_1 \oplus \cdots \oplus U_m$ eine direkte Summe, so hat nach Satz 9.1 ((1) $\Rightarrow$ (3)) jedes $u \in U$ eine eindeutige Darstellung $u = u_1 + \cdots + u_m$ mit $u_j \in U_j$ für $j = 1, \ldots, m$.

Ist $m \geq 3$, so ist es klar, dass $U = U_1 + (U_2 + \cdots + U_m)$.

Satz 9.2 *Sei $m \geq 3$, setze $W = U_2 + \cdots + U_m$ und nehme an, dass $U = U_1 \oplus W$ und $W = U_2 \oplus \cdots \oplus U_m$. Dann gilt auch $U = U_1 \oplus \cdots \oplus U_m$.*

Beweis Sei $u_j \in U_j$ für jedes $j = 1, \ldots, m$ mit $u_1 + \cdots + u_m = 0$. Dann ist $w = u_2 + \cdots + u_m \in W$ und $u_1 + w = 0$. Nach Satz 9.1 ((1) $\Rightarrow$ (2)) ist also $u_1 = 0$ und $w = 0$, da $U = U_1 \oplus W$, und wieder nach Satz 9.1 ((1) $\Rightarrow$ (2)) ist dann $u_j = 0$ für jedes $j = 2, \ldots, m$, da $W = U_2 \oplus \cdots \oplus U_m$. Daraus ergibt sich nach Satz 9.1 ((2) $\Rightarrow$ (1)), das $U = U_1 \oplus \cdots \oplus U_m$. □

Satz 9.3 *Sei $U_j \neq \{0\}$ für jedes j. Dann sind äquivalent:*

(1) Es gilt $U = U_1 \oplus \cdots \oplus U_m$.

(2) Ist $u_j \in U_j \setminus \{0\}$ für jedes j, so sind $u_1, \ldots, u_n$ linear unabhängig.

(3) Sind $u_1^j, \ldots, u_{k_j}^j \in U_j$ linear unabhängig für jedes $j = 1, \ldots, m$, so sind die Vektoren $u_1^1, \ldots, u_{k_1}^1, \ldots, u_1^m, \ldots, u_{k_m}^m$ linear unabhängig.

Beweis (1) $\Rightarrow$ (3): Für $j = 1, \ldots, m$ seien $u_1^j, \ldots, u_{k_j}^j \in U_j$ linear unabhängig und seien $\lambda_1^1, \ldots, \lambda_{k_1}^1, \ldots, \lambda_1^m, \ldots, \lambda_{k_m}^m \in K$ mit

$$\lambda_1^1 u_i^1 + \cdots + \lambda_{k_1}^1 u_{k_1}^1 + \cdots + \lambda_1^m u_1^m + \cdots + \lambda_{k_m}^m u_{k_m}^m = 0 \, .$$

Setze $u_j = \lambda_1^j u_1^j + \cdots + \lambda_{k_j}^j u_{k_j}^j$; dann ist $u_j \in U_j$ und $u_1 + \cdots + u_m = 0$. Nach Satz 1 ist also $u_j = \lambda_1^j u_1^j + \cdots + \lambda_{k_j}^j u_{k_j}^j = 0$ für jedes j und damit gilt $\lambda_i^j = 0$ für $i = 1, \ldots, k_j$, da $u_1^j, \ldots, u_{k_j}^j \in U_j$ linear unabhängig sind. Dies zeigt, dass die Vektoren $u_1^1, \ldots, u_{k_1}^1, \ldots, u_1^m, \ldots, u_{k_m}^m$ linear unabhängig sind.

(3) $\Rightarrow$ (2): Dies ist klar.

(2) $\Rightarrow$ (1): Für $j = 1, \ldots, m$ sei $u_j \in U_j$ mit $u_1 + \cdots + u_m = 0$, und setze

$$v_j = \begin{cases} u_j & \text{falls } u_j \neq 0 \, , \\ u_j' & \text{falls } u_j = 0 \, , \end{cases} \quad \text{und} \quad \lambda_j = \begin{cases} 1 & \text{falls } u_j \neq 0 \, , \\ 0 & \text{falls } u_j = 0 \, , \end{cases}$$

wobei u_j' ein beliebiges Element aus $U_j \setminus \{0\}$ ist. Dann ist $\lambda_1 v_1 + \cdots \lambda_m v_m = 0$. Aber $v_j \in U_j \setminus \{0\}$ für $j = 1, \ldots, m$, damit sind $v_1, \ldots, v_n$ linear unabhängig und folglich ist $\lambda_j = 0$ für jedes j, d.h., $u_j = 0$ für $j = 1, \ldots, m$. Daraus ergibt sich nach Satz 9.1, dass die Summe direkt ist. □

Satz 9.4 *Sei $U = U_1 \oplus \cdots \oplus U_m$ und für jedes $j = 1, \ldots, m$ sei $f_j : U_j \to U_j$ ein Endomorphismus von U_j. Dann gibt es einen eindeutigen Endomorphismus $f : U \to U$ von U mit $f(u) = f_j(u)$ für jedes $u \in U_j$, $j = 1, \ldots, m$.*

Beweis Nach Satz 9.1 ((1) $\Rightarrow$ (3)) hat jedes $u \in U$ eine eindeutige Darstellung $u = u_1 + \cdots + u_m$ mit $u_j \in U_j$ für $j = 1, \ldots, m$ und folglich gibt es eine eindeutige Abbildung $f : U \to U$ mit

$$f(u_1 + \cdots + u_m) = f_1(u_1) + \cdots + f_m(u_m)$$

für alle $u_j \in U_j$, $j = 1, \ldots, m$. Man sieht leicht, dass f ein Endomorphismus ist und per Definition gilt $f(u) = f_j(u)$ für jedes $u \in U_j$, $j = 1, \ldots, m$. Ferner ist es klar, dass f der eindeutige Endomorphismus von U mit dieser Eigenschaft ist. □

Der Endomorphismus in Satz 9.4 heißt die *direkte Summe* der Endomorphismen $f_1, \ldots, f_m$ und wird mit $f_1 \oplus \cdots \oplus f_m$ bezeichnet.

Lemma 9.1 *Sei $U = U_1 \oplus \cdots \oplus U_m$ und sei f ein Endomorphismus von U. Dann gibt es Endomorphismen $f_j : U_j \to U_j$, $j = 1, \ldots, m$, so dass $f = f_1 \oplus \cdots \oplus f_m$, genau dann, wenn $f(U_j) \subset U_j$ für jedes $j = 1, \ldots, m$.*

Beweis Dies ist klar. □

Satz 9.5 *Sei $U = U_1 \oplus \cdots \oplus U_m$ und seien $f_j, g_j : U_j \to U_j$ Endomorphismen von U_j für jedes $j = 1, \ldots, m$. Dann gilt*

$$(f_1 \oplus \cdots \oplus f_m) + (g_1 \oplus \cdots \oplus g_m) = (f_1 + g_1) \oplus \cdots \oplus (f_m + g_m)$$
$$\text{und} \quad (f_1 \oplus \cdots \oplus f_m) \circ (g_1 \oplus \cdots \oplus g_m) = (f_1 \circ g_1) \oplus \cdots \oplus (f_m \circ g_m) \, .$$

Beweis Setze $f = f_1 \oplus \cdots \oplus f_m$ und $g = g_1 \oplus \cdots \oplus g_m$. Dann gilt

$$\begin{aligned}(f+g)(u_1 + \cdots + u_m) &= f(u_1 + \cdots + u_m) + g(u_1 + \cdots + u_m) \\ &= f_1(u_1) + \cdots + f_m(u_m) + g_1(u_1) + \cdots + g_m(u_m) \\ &= (f_1(u_1) + g_1(u_1)) + \cdots + (f_m(u_m) + g_m(u_m)) \\ &= (f_1 + g_1)(u_1) + \cdots + (f_m + g_m)(u_m) \\ &= ((f_1 + g_1) \oplus \cdots \oplus (f_m + g_m))(u_1 + \cdots + u_m)\end{aligned}$$

und auch

$$\begin{aligned}(f \circ g)(u_1 + \cdots + u_m) &= f(g(u_1 + \cdots + u_m)) = f(g_1(u_1) + \cdots + g_m(u_m)) \\ &= f_1(g_1(u_1)) + \cdots + f_m(g_m(u_m)) = (f_1 \circ g_1)(u_1) + \cdots + (f_m \circ g_m)(u_m) \\ &= ((f_1 \circ g_1) \oplus \cdots \oplus (f_m \circ g_m))(u_1 + \cdots + u_m)\end{aligned}$$

für alle $u_j \in U_j$, $j = 1, \ldots, m$, und daraus ergibt sich, dass

$$f + g = (f_1 + g_1) \oplus \cdots \oplus (f_m + g_m) \quad \text{und} \quad f \circ g = (f_1 \circ g_1) \oplus \cdots \oplus (f_m \circ g_m) \, . \quad \square$$

Satz 9.6 *Sei $U = U_1 \oplus \cdots \oplus U_m$ und für $j = 1, \ldots, m$ sei $f_j : U_j \to U_j$ ein Endomorphismus von U_j. Dann ist $f_1 \oplus \cdots \oplus f_m$ ein Automorphismus von U genau, wenn f_j ein Automorphismus von U_j ist für jedes $j = 1, \ldots, m$.*

Beweis Dies folgt unmittelbar aus Satz 9.5, da $\mathrm{id}_U = \mathrm{id}_{U_1} \oplus \cdots \oplus \mathrm{id}_{U_m}$. □

Lemma 9.2 *Sei $U = U_1 \oplus \cdots \oplus U_m$ und für $j = 1, \ldots, m$ sei $f_j : U_j \to U_j$ ein Endomorphismus von U_j. Dann ist*

$$\operatorname{Kern}(f_1 \oplus \cdots \oplus f_m) = \operatorname{Kern} f_1 \oplus \cdots \oplus \operatorname{Kern} f_m \ .$$

Beweis Dies ist klar. □

Seien $U_1, \ldots, U_m$ endlichdimensional; da

$$U = U_1 + \cdots + U_m = (\cdots((U_1 + U_2) + U_3) + \cdots + U_m)$$

ist nach Satz 4.8 die Summe U endlichdimensional und es gilt

$$\dim U \leq \dim U_1 + \cdots + \dim U_m \ .$$

Satz 9.7 *Seien $U_1, \ldots, U_m$ endlichdimensional mit $\dim U_j \geq 1$ für jedes j. Dann sind äquivalent:*

(1) Es gilt $U = U_1 \oplus \cdots \oplus U_m$.

(2) Ist $(u_1^j, \ldots, u_{k_j}^j)$ eine Basis von U_j für jedes $j = 1, \ldots, m$, so ist

$$(u_1^1, \ldots, u_{k_1}^1, \ldots, u_1^m, \ldots, u_{k_m}^m)$$

eine Basis von U.

(3) Es gilt $\dim U = \dim U_1 + \cdots + \dim U_m$.

Beweis (2) ⇒ (3) ist klar und (1) ⇒ (2) folgt unmittelbar aus Satz 9.3, da

$$U = \mathrm{L}(v_1^1, \ldots, v_{k_1}^1, \ldots, v_1^m, \ldots, v_{k_m}^m) \ ,$$

falls $\mathrm{L}(v_1^j, \ldots, v_{k_j}^j) = U_j$ für jedes $j = 1, \ldots, m$.

(3) ⇒ (2): Für jedes $j = 1, \ldots, m$ sei $(u_1^j, \ldots, u_{k_j}^j)$ eine Basis von U_j. Da wieder

$$U = \mathrm{L}(u_1^1, \ldots, u_{k_1}^1, \ldots, u_1^m, \ldots, u_{k_m}^m)$$

und $\dim U = k_1 + \cdots + k_m$, ist nach Satz 4.6 (4) $(u_1^1, \ldots, u_{k_1}^1, \ldots, u_1^m, \ldots, u_{k_m}^m)$ eine Basis von U.

(2) ⇒ (1): Für jedes $j = 1, \ldots, m$ seien $u_1^j, \ldots, u_{k_j}^j \in U_j$ linear unabhängig. Nach dem Basisergänzungssatz gibt es eine Basis $(v_1^j, \ldots, v_{p_j}^j)$ von U_j mit $k_j \leq p_j$ und $v_i^j = u_i^j$ für $i = 1, \ldots, k_j$. Dann ist $(v_1^1, \ldots, v_{p_1}^1, \ldots, v_1^m, \ldots, v_{p_m}^m)$ eine Basis von U und insbesondere sind die Vektoren $v_1^1, \ldots, v_{p_1}^1, \ldots, v_1^m, \ldots, v_{p_m}^m$ linear unabhängig. Damit sind auch $u_1^1, \ldots, u_{k_1}^1, \ldots, u_1^m, \ldots, u_{k_m}^m$ linear unabhängig und daraus ergibt sich nach Satz 9.3, dass die Summe direkt ist. □

10 Diagonalisierbarkeit

Im Folgenden sei K ein Körper und sei V ein Vektorraum über K.

Sei $f : V \to V$ ein Endomorphismus von V. Ein Element $\lambda \in K$ heißt *Eigenwert* von f, wenn es einen Vektor $v \in V$, $v \neq 0$, mit $f(v) = \lambda v$ gibt. Ein Vektor $v \neq 0$ mit $f(v) = \lambda v$ heißt dann *Eigenvektor* von f zum Eigenwert λ.

Wichtige Bemerkung: 0 ist ein Eigenwert von f genau dann, wenn f nicht injektiv ist, d.h., genau dann, wenn $\operatorname{Kern} f \neq \{0\}$.

Nehme nun an, dass V stets endlichdimensional ist mit $\dim V = n \geq 1$.

Ein Endomorphismus $f : V \to V$ heißt *diagonalisierbar*, wenn es eine aus Eigenvektoren von f bestehende Basis von V gibt, d.h., wenn es eine Basis $(v_1, \ldots, v_n)$ von V gibt, so dass $f(v_j) = \lambda_j v_j$ für ein $\lambda_j \in K$ für jedes $j = 1, \ldots, n$.

Es gibt Endomorphismen, die keine Eigenwerte besitzen, und insbesondere sind solche Endomorphismen nicht diagonalisierbar.

Beispiel: Sei $f : \mathbb{R}^2 \to \mathbb{R}^2$ die durch $f((x, y)) = (-y, x)$ gegeben Abbildung; dann ist f ein Endomorphismus von $\mathbb{R}^2$, und $f = \varphi_A$ mit

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

Sei $(x, y) \in \mathbb{R}^2$ und $\lambda \in \mathbb{R}$ mit $f((x, y)) = \lambda(x, y)$; dann ist $(-y, x) = \lambda(x, y)$, d.h., $-y = \lambda x$ und $x = \lambda y$ und damit $x = -\lambda^2 x$, und dies ist nur möglich, wenn $(x, y) = (0, 0)$. Daraus folgt, dass f keinen Eigenwert besitzt.

Eine Matrix $D = (d_{ij}) \in \mathrm{M}(n \times n, K)$ heißt *Diagonalmatrix*, wenn $d_{ij} = 0$ für alle i, j mit $i \neq j$. Eine Diagonalmatrix hat also folgende Gestalt:

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

Lemma 10.1 *Sei $f : V \to V$ ein Endomorphismus und sei $\alpha = (v_1, \ldots, v_n)$ eine Basis von V. Dann sind äquivalent:*

(1) Die Basis α besteht aus Eigenvektoren von f.

(2) Die Matrix von f bezüglich α ist eine Diagonalmatrix.

Insbesondere ist f genau dann diagonalisierbar, wenn es eine Basis α von V gibt, so dass die Matrix von f bezüglich α eine Diagonalmatrix ist.

Beweis (1) $\Rightarrow$ (2): Da v_j ein Eigenvektor von f ist, ist $f(v_j) = \lambda_j v_j$ für den entsprechenden Eigenwert $\lambda_j \in K$. Dann aber ist die Diagonalmatrix

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

die Matrix von f bezüglich der Basis α.

(2) $\Rightarrow$ (1): Sei $D = (d_{ij})$ die Matrix von f bezüglich der Basis α. Sei $1 \leq j \leq n$; da D eine Diagonalmatrix ist, gilt

$$f(v_j) = d_{1j}v_1 + \cdots + d_{nj}v_n = d_{jj}v_j \ .$$

Ferner ist $v_j \neq 0$, da $(v_1, \ldots, v_n)$ eine Basis ist und folglich ist v_j ein Eigenvektor von f zum Eigenwert d_{jj}. Damit besteht $(v_1, \ldots, v_n)$ aus Eigenvektoren von f. □

Eine Matrix $A \in \mathrm{M}(n \times n, K)$ heißt *diagonalisierbar*, wenn A ähnlich zu einer Diagonalmatrix ist, d.h., wenn es Matrizen D, $P \in \mathrm{M}(n \times n, K)$ mit P invertierbar und D einer Diagonalmatrix gibt, so dass $A = P^{-1}DP$.

Satz 10.1 *Sei $f : V \to V$ ein Endomorphismus, sei α eine Basis von V und sei A die Matrix von f bezüglich der Basis α. Dann ist der Endomorphismus f diagonalisierbar genau, wenn die Matrix A diagonalisierbar ist.*

Beweis Nehme zunächst an, dass der Endomorphismus f diagonalisierbar ist. Nach Lemma 10.1 gibt es dann eine Basis β von V, so dass die Matrix D von f bezüglich β eine Diagonalmatrix ist. Aber nach Satz 8.8 gilt $A = P^{-1}DP$, wobei P die Matrix für den Wechsel von β nach α ist. Damit ist die Matrix A diagonalisierbar.

Nehme nun umgekehrt an, dass die Matrix A diagonalisierbar ist. Dann gibt es Matrizen D, $P \in \mathrm{M}(n \times n, K)$ mit D einer Diagonalmatrix und P invertierbar, so dass $A = P^{-1}DP$, und nach Lemma 6.6 gibt es eine Basis β von V, so dass P^{-1} die Matrix für den Wechsel von α nach β ist. Daraus ergibt sich nach Satz 8.8, das $(P^{-1})^{-1}AP^{-1} = PAP^{-1} = D$ die Matrix von f bezüglich der Basis β ist. Damit ist nach Lemma 10.1 der Endomorphismus f diagonalisierbar. □

Sei $A \in \mathrm{M}(n \times n, K)$; nach Satz 10.1 ist der Endomorphismus φ_A diagonalisierbar genau, wenn die Matrix A diagonalisierbar ist, da nach Lemma 8.1 A die Matrix von φ_A bezüglich der kanonischen Basis von K^n ist.

Sei $A \in \mathrm{M}(n \times n, K)$; Ein Element $\lambda \in K$ heißt *Eigenwert* von A, wenn λ Eigenwert von φ_A ist, d.h., wenn es einen Vektor $u \in K^n$, $u \neq 0$, mit $\varphi_A(u) = \lambda u$

gibt. Ein Vektor $u \neq 0$ mit $\varphi_A(u) = \lambda u$ heißt dann *Eigenvektor* von A zum Eigenwert λ, d.h., ein Eigenvektor von A zum Eigenwert λ ist nichts anderes als ein Eigenvektor von φ_A zum Eigenwert λ.

Es gilt aber $\varphi_A(u) = Au$, wenn u als Element von $\mathrm{M}(n \times 1, K)$ betrachtet wird, und folglich ist $\lambda \in K$ Eigenwert von A, genau wenn es einen Vektor $u \in K^n$, $u \neq 0$, mit $Au = \lambda u$ gibt, d.h., wenn

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} = \lambda \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} ,$$

wobei $u = (\mu_1, \ldots, \mu_n)$. Ein Vektor $u \in K^n$ mit $u \neq 0$ ist also Eigenvektor von A zum Eigenwert λ genau dann, wenn $Au = \lambda u$.

Lemma 10.2 *Seien A, $Q \in \mathrm{M}(n \times n, K)$ und sei*

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

eine Diagonalmatrix. Dann gilt $AQ = QD$ genau, wenn $Au_j = \lambda_j u_j$ für jedes j, wobei $u_1, \ldots, u_n$ die Spalten von Q sind. Ferner ist die Matrix Q invertierbar genau dann, wenn $(u_1, \ldots, u_n)$ eine Basis von K^n ist.

Beweis Die erste Aussage folgt aus der Definition von Matrizenmultiplikation. Nun gilt $\dim \mathrm{L}(u_1, \ldots, u_n) = \mathrm{Rang}_S(Q) = \operatorname{rang} Q$ und nach Satz 4.6 (4) ist $(u_1, \ldots, u_n)$ eine Basis von K^n genau dann, wenn $\dim \mathrm{L}(u_1, \ldots, u_n) = n$. Nach Satz 7.4 ist andererseits Q invertierbar genau dann, wenn $\operatorname{rang} Q = n$. Damit ist Q invertierbar genau dann, wenn $(u_1, \ldots, u_n)$ eine Basis von K^n ist. □

Satz 10.2 *Eine Matrix $A \in \mathrm{M}(n \times n, K)$ ist diagonalisierbar genau dann, wenn es eine aus Eigenvektoren von A bestehende Basis von K^n gibt.*

Beweis Dies folgt unmittelbar aus Satz 10.1, da nach Lemma 8.1 A die Matrix von φ_A bezüglich der kanonischen Basis von K^n ist. Es ist dennoch aufschlußreich, den folgenden direkten Beweis zu betrachten:

Nehme zuerst an, dass A diagonalisierbar ist; dann gibt es eine Diagonalmatrix

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

und eine invertierbare Matrix P, so dass $A = P^{-1}DP$. Sei $Q = P^{-1}$ und seien $u_1, \ldots, u_n$ die Spalten von Q. Dann ist $Q^{-1}AQ = D$ und damit auch $AQ = QD$. Nach Lemma 10.2 gilt also $Au_j = \lambda_j u_j$ für $j = 1, \ldots, n$ und $(u_1, \ldots, u_n)$ ist eine Basis von K^n. Damit ist $(u_1, \ldots, u_n)$ eine aus Eigenvektoren von A bestehende Basis von K^n.

Nehme nun umgekehrt an, dass es eine aus Eigenvektoren von A bestehende Basis $(u_1, \ldots, u_n)$ von K^n gibt. Für $j = 1, \ldots, n$ sei λ_j der entsprechende Eigenwert, d.h., es gilt $Au_j = \lambda_j u_j$. Sei $Q \in \mathrm{M}(n \times n, K)$ die Matrix, die $u_1, \ldots, u_n$ als Spalten hat. Nach Lemma 10.2 ist Q invertierbar und $AQ = QD$, wobei

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

Damit ist $A = P^{-1}DP$, wobei $P = Q^{-1}$, d.h., A ist diagonalisierbar. □

Für jede Basis $\alpha = (v_1, \ldots, v_n)$ von V sei $\psi_\alpha : K^n \to V$ die durch

$$\psi_\alpha((\lambda_1, \ldots, \lambda_n)) = \lambda_1 v_1 + \cdots + \lambda_n v_n$$

definierte Abbildung. Man sieht leicht, dass ψ_α linear ist. Ist $(e_1, \ldots, e_n)$ die kanonische Basis von K^n, so gilt $\psi_\alpha(e_j) = u_j$ für jedes $j = 1, \ldots, n$ und daraus folgt nach Lemma 5.10 (4), dass ψ_α ein Isomorphismus ist.

Lemma 10.3 *Sei $f : V \to V$ ein Endomorphismus und sei A die Matrix von f bezüglich einer Basis α von V. Dann gilt $f \circ \psi_\alpha = \psi_\alpha \circ \varphi_A$.*

Beweis Sei $\alpha = (v_1, \ldots, v_n)$ und $A = (a_{ij})$. Für alle $(\lambda_1, \ldots, \lambda_n) \in K^n$ ist

$$\begin{aligned}(f \circ \psi_\alpha)((\lambda_1, \ldots, \lambda_n)) &= f(\psi_\alpha((\lambda_1, \ldots, \lambda_n))) \\ = f\Big(\sum_{j=1}^n \lambda_j v_j\Big) &= \sum_{j=1}^n \lambda_j f(v_j) = \sum_{j=1}^n \lambda_j \sum_{i=1}^n a_{ij} v_i = \sum_{i=1}^n \Big(\sum_{j=1}^n a_{ij}\lambda_j\Big) v_i \\ &= \psi_\alpha(\varphi_A((\lambda_1, \ldots, \lambda_n))) = (\psi_\alpha \circ \varphi_A)((\lambda_1, \ldots, \lambda_n))\end{aligned}$$

und damit ist $f \circ \psi_\alpha = \psi_\alpha \circ \varphi_A$. □

Satz 10.3 *Sei $f : V \to V$ ein Endomorphismus, sei $\alpha = (v_1, \ldots, v_n)$ eine Basis von V und sei A die Matrix von f bezüglich der Basis α.*

(1) Ein Element $\lambda \in K$ ist ein Eigenwert von f genau dann, wenn λ ein Eigenwert von A ist.

(2) Sei λ ein Eigenwert von f und $u \in K^n$; dann ist u ein Eigenvektor von A zum Eigenwert λ genau, wenn $\psi_\alpha(u)$ ein Eigenvektor von f zum Eigenwert λ ist.

(3) Ist A diagonalisierbar und ist $(u_1, \ldots, u_n)$ eine aus Eigenvektoren von A bestehende Basis von K^n, so ist $(\psi_\alpha(u_1), \ldots, \psi_\alpha(u_n))$ eine aus Eigenvektoren von f bestehende Basis von V.

Beweis (1) und (2): Sei $\lambda \in K$ und $u \in K^n$; es gilt $f \circ \psi_\alpha = \psi_\alpha \circ \varphi_A$ nach Lemma 10.3, und ψ_α ist ein Isomorphismus, und daraus ergibt sich, dass

$$\begin{aligned} Au = \lambda u &\Leftrightarrow \varphi_A(u) = \lambda u \Leftrightarrow \varphi_A(u) - \lambda u = 0 \\ &\Leftrightarrow \psi_\alpha(\varphi_A(u) - \lambda u) = 0 \Leftrightarrow \psi_\alpha(\varphi_A(u)) - \lambda\psi_\alpha(u) = 0 \\ &\Leftrightarrow f(\psi_\alpha(u)) - \lambda\psi_\alpha(u) = 0 \Leftrightarrow f(\psi_\alpha(u)) = \lambda\psi_\alpha(u)\ . \end{aligned}$$

Ferner ist $u \neq 0$ genau dann, wenn $\psi_\alpha(u) \neq 0$. Damit ist λ ein Eigenwert von f genau dann, wenn λ ein Eigenwert von A ist, und u ist ein Eigenvektor von A zum Eigenwert λ genau dann, wenn $\psi_\alpha(u)$ ein Eigenvektor von f zum Eigenwert λ ist.

(3) Dies folgt unmittelbar aus (1) und (2) und Lemma 5.10 (3). □

Lemma 10.4 *Sei $f : V \to V$ ein Endomorphismus und für $j = 1, \ldots, m$ sei v_j ein Eigenvektor von f zum Eigenwert λ_j, wobei $\lambda_i \neq \lambda_j$, falls $i \neq j$. Dann sind die Vektoren $v_1, \ldots, v_m$ linear unabhängig.*

Beweis Nehme an, dass $v_1, \ldots, v_m$ linear abhängig sind. Da $v_1 \neq 0$, gibt es nach Lemma 3.13 ein k mit $2 \leq k \leq m$, so dass $v_k \in \mathrm{L}(v_1, \ldots, v_{k-1})$. Sei p der kleinste solche Index, es gilt also $2 \leq p \leq m$, $v_p \in \mathrm{L}(v_1, \ldots, v_{p-1})$, aber $v_i \notin \mathrm{L}(v_1, \ldots, v_{i-1})$ für alle $2 \leq i < p$. Nach Lemma 3.13 sind $v_1, \ldots, v_{p-1}$ linear unabhängig. Da $v_p \in \mathrm{L}(v_1, \ldots, v_{p-1})$, gibt es $\mu_1, \ldots, \mu_{p-1} \in K$, so dass $v_p = \mu_1 v_1 + \cdots + \mu_{p-1} v_{p-1}$. Dann ist

$$\begin{aligned} 0 = f(v_p) - \lambda_p v_p &= f(\mu_1 v_1 + \cdots + \mu_{p-1} v_{p-1}) - \lambda_p(\mu_1 v_1 + \cdots + \mu_{p-1} v_{p-1}) \\ &= \mu_1 f(v_1) + \cdots + \mu_{p-1} f(v_{p-1}) - \lambda_p(\mu_1 v_1 + \cdots + \mu_{p-1} v_{p-1}) \\ &= \mu_1\lambda_1 v_1 + \cdots + \mu_{p-1}\lambda_{p-1} v_{p-1} - (\mu_1\lambda_p v_1 + \cdots + \mu_{p-1}\lambda_p v_{p-1}) \\ &= \mu_1(\lambda_1 - \lambda_p)v_1 + \cdots + \mu_{p-1}(\lambda_{p-1} - \lambda_p)v_{p-1}\ , \end{aligned}$$

und daraus ergibt sich, dass $\mu_j(\lambda_j - \lambda_p) = 0$ für $j = 1, \ldots, p-1$, da $v_1, \ldots, v_{p-1}$ linear unabhängig sind. Aber $\lambda_j - \lambda_p \neq 0$ und also ist $\mu_j = 0$ für $j = 1, \ldots, p-1$. Damit ist $v_p = \mu_1 v_1 + \cdots + \mu_{p-1} v_{p-1} = 0$ und dies ist ein Widerspruch, da v_p ein Eigenvektor ist. Daher müssen $v_1, \ldots, v_m$ linear unabhängig sein. □

Satz 10.4 *(1) Ein Endomorphismus $f : V \to V$ hat höchstens n verschiedene Eigenwerte (wobei $n = \dim V$), und gibt es n verschiedene Eigenwerte, so ist f diagonalisierbar.*

(2) Eine Matrix $A \in \mathrm{M}(n \times n, K)$ hat höchstens n verschiedene Eigenwerte. Gibt es n verschiedene Eigenwerte von A, so ist A diagonalisierbar.

Beweis Dies folgt unmittelbar aus Lemma 10.4 und (1) und (2) von Satz 4.6. □

Sei $f : V \to V$ ein Endomorphismus und sei $\lambda \in K$. Für $v \in V$ gilt $f(v) = \lambda v$ genau dann, wenn $v \in \mathrm{Kern}\,(f - \lambda\,\mathrm{id}_V)$, da

$$f(v) = \lambda v \;\Leftrightarrow\; f(v) - \lambda v = 0 \;\Leftrightarrow\; (f - \lambda\,\mathrm{id}_V)(v) = 0 \;\Leftrightarrow\; v \in \mathrm{Kern}\,(f - \lambda\,\mathrm{id}_V)\,.$$

Folglich ist λ ein Eigenwert von f genau dann, wenn $\mathrm{Kern}\,(f - \lambda\,\mathrm{id}_V) \neq \{0\}$. Daraus ergibt sich nach Satz 5.3, dass λ genau dann ein Eigenwert von f ist, wenn der Endomorphismus $f - \lambda\,\mathrm{id}_V : V \to V$ kein Automorphismus ist.

Für jeden Eigenwert λ setze $\mathrm{E}(f, \lambda) = \mathrm{Kern}\,(f - \lambda\,\mathrm{id}_V)$. Dieser Untervektorraum von V heißt *Eigenraum von f zum Eigenwert λ*. Es gilt $\dim \mathrm{E}(f, \lambda) \geq 1$ und $\mathrm{E}(f, \lambda) \setminus \{0\}$ ist gerade die Menge aller Eigenvektoren von f zum Eigenwert λ.

Wichtige Bemerkung: Ist 0 ein Eigenwert von f, so ist $E(f, 0) = \mathrm{Kern}\, f$.

Satz 10.5 *Sei $f : V \to V$ ein Endomorphismus von V und seien $\lambda_1, \ldots, \lambda_m$ die verschiedenen Eigenwerte. Dann ist*

$$\mathrm{E}(f, \lambda_1) + \cdots + \mathrm{E}(f, \lambda_m) = \mathrm{E}(f, \lambda_1) \oplus \cdots \oplus \mathrm{E}(f, \lambda_m)\,.$$

(die Summe ist also direkt). Ferner ist f diagonalisierbar genau dann, wenn

$$V = \mathrm{E}(f, \lambda_1) \oplus \cdots \oplus \mathrm{E}(f, \lambda_m)\,.$$

Beweis Es folgt zunächst unmittelbar aus Satz 9.3 und Lemma 10.4, dass die Summe $\mathrm{E}(f, \lambda_1) + \cdots + \mathrm{E}(f, \lambda_m)$ direkt ist.

Gilt $V = \mathrm{E}(f, \lambda_1) \oplus \cdots \oplus \mathrm{E}(f, \lambda_m)$, so ist nach Satz 9.3 f diagonalisierbar. Ist umgekehrt f diagonalisierbar, so ist es klar, dass $V = \mathrm{E}(f, \lambda_1) + \cdots + \mathrm{E}(f, \lambda_m)$. Also ist f diagonalisierbar genau dann, wenn $V = \mathrm{E}(f, \lambda_1) \oplus \cdots \oplus \mathrm{E}(f, \lambda_m)$. □

Sei $f : V \to V$ ein diagonalisierbarer Endomorphismus und seien $\lambda_1, \ldots, \lambda_m$ die verschiedenen Eigenwerte von f. Für $j = 1, \ldots, m$ setze $U_j = \mathrm{E}(f, \lambda_j)$; also ist $V = U_1 \oplus \cdots \oplus U_m$ und ferner gilt

$$f = \lambda_1 \mathrm{id}_{U_1} \oplus \cdots \oplus \lambda_m \mathrm{id}_{U_m}\,,$$

d.h., f ist die direkte Summe der Endomorphismen $\lambda_1 \mathrm{id}_{U_1}, \ldots, \lambda_m \mathrm{id}_{U_m}$.

Satz 10.6 *Sei $f : V \to V$ ein Endomorphismus von V und seien $\lambda_1, \ldots, \lambda_m$ die verschiedenen Eigenwerte. Dann gilt* $\dim \mathrm{E}(f, \lambda_1) + \cdots + \dim \mathrm{E}(f, \lambda_m) \leq \dim V$. *Ferner ist f diagonalisierbar genau dann, wenn*

$$\dim \mathrm{E}(f, \lambda_1) + \cdots + \dim \mathrm{E}(f, \lambda_m) = \dim V \ .$$

Beweis Dies folgt unmittelbar aus Satz 10.5 und Satz 9.7. □

Beispiel Sei $f : K^2 \to K^2$ die Abbildung, die durch $f((x, y)) = (y, 0)$ gegeben ist; dann ist f ein Endomorphismus von K^2, und $f = \varphi_A$ mit

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

Sei $(x, y) \in K^2$ und $\lambda \in K$ mit $f((x, y)) = \lambda\,(x, y)$; dann ist $(y, 0) = \lambda(x, y)$, d.h., $\lambda y = 0$ und $y = \lambda x$. Damit ist 0 der einzige Eigenwert von f und

$$\mathrm{E}(f, 0) = \{v \in K^2 : f(v) = 0\} = \{(x, 0) : x \in K\} \ .$$

Insbesondere ist $\dim \mathrm{E}(f, 0) = 1 < 2 = \dim K^2$, und daraus folgt nach Satz 10.6, dass f nicht diagonalisierbar ist.

Sei nun $A \in \mathrm{M}(n \times n, K)$ und sei $\lambda \in K$. Für $u \in K^n$ gilt $Au = \lambda u$ genau dann, wenn $u \in \operatorname{Kern} \varphi_{A-\lambda E_n}$, da

$$\begin{aligned} Au = \lambda u \ \Leftrightarrow \ Au - \lambda E_n u = 0 \ &\Leftrightarrow \ (A - \lambda E_n)u = 0 \\ &\Leftrightarrow \ \varphi_{A-\lambda E_n}(u) = 0 \ \Leftrightarrow \ u \in \operatorname{Kern} \ \varphi_{A-\lambda E_n} \ . \end{aligned}$$

Folglich ist λ ein Eigenwert von A genau dann, wenn der Endomorphismus $\varphi_{A-\lambda E_n}$ kein Automorphismus von K^n ist. Nach Satz 6.3 ist also λ ein Eigenwert von A genau dann, wenn die Matrix $A - \lambda E_n$ nicht invertierbar ist.

Für jeden Eigenwert λ setze $\mathrm{E}(A, \lambda) = \operatorname{Kern} \varphi_{A-\lambda E_n}$. Dieser Untervektorraum von V heißt *Eigenraum von A zum Eigenwert λ*. Es gilt $\dim \mathrm{E}(A, \lambda) \geq 1$ und $\mathrm{E}(A, \lambda) \setminus \{0\}$ ist gerade die Menge aller Eigenvektoren von A zum Eigenwert λ.

Lemma 10.5 *Sei $A \in \mathrm{M}(n \times n, K)$; dann gilt*

$$\mathrm{E}(A, \lambda) = \mathrm{Lös}(A - \lambda E_n, 0)$$

für jeden Eigenwert λ von A.

Beweis Dies folgt unmittelbar aus Lemma 5.7. □

Sei nun $A \in \mathrm{M}(n \times n, K)$ und seien $\lambda_1, \ldots, \lambda_m$ die verschiedenen Eigenwerte von A. Nach Lemma 10.5 ist $\mathrm{E}(A, \lambda_j) = \mathrm{Lös}(A - \lambda_j E_n, 0)$, und also kann man mit

Hilfe von Satz 10.1 und des Gaußschen Algorithmus eine Basis $(u_1^j, \ldots, u_{p_j}^j)$ von $\mathrm{E}(A, \lambda_j)$ konstruieren. Nach Satz 10.6 ist A diagonalisierbar genau dann, wenn $p_1 + \cdots + p_m = n$. Nehme an, dass A diagonalisierbar ist und setze

$$(w_1, \ldots, w_n) = (u_1^1, \ldots, u_{p_1}^1, u_1^2, \ldots, u_{p_2}^2, \ldots, u_1^m, \ldots, u_{p_m}^m) \; .$$

Dann ist $(w_1, \ldots, w_n)$ eine aus Eigenvektoren von A bestehende Basis von K^n. Sei $Q \in \mathrm{M}(n \times n, K)$ die Matrix, die $w_1, \ldots, w_n$ als Spalten hat; wie im Beweis für Satz 10.2 gilt dann, dass Q invertierbar ist und $Q^{-1}AQ = D$, wobei D die Diagonalmatrix ist mit den Einträgen $\lambda_1, \ldots, \lambda_1, \lambda_2, \ldots, \lambda_2, \ldots, \lambda_m, \ldots, \lambda_m$ auf der Diagonalen, wobei für jedes j der Eigenwert λ_j genau p_j-mal vorkommt.

Diese Verfahren kann aber nur durchführt werden, wenn die Eigenwerte von A schon bekannt sind, und daher stellt sich die Frage: Wie findet man die Eigenwerte einer Matrix $A \in \mathrm{M}(n \times n, K)$?

Bevor aber diese Frage beantwortet werden kann, müssen einige Eigenschaften über Determinanten von Matrizen kennengelernt werden.

11 Trigonalisierbarkeit

Im Folgenden sei K ein Körper und V ein Vektorraum über K.

Sei $f : V \to V$ ein Endomorphismus von V. Ein Untervektorraum U von V heißt *f-invariant*, wenn $f(U) \subset U$. In diesem Fall kann eine Abbildung $f_{|U} : U \to U$ definiert werden durch $f_{|U}(u) = f(u)$ für alle $u \in U$. Man sieht leicht, dass $f_{|U}$ linear ist, d.h., $f_{|U}$ ist ein Endomorphismus von U. Diesen Endomorphismus $f_{|U}$ nennt man die *Einschränkung* von f auf U.

Sei $f : V \to V$ ein Endomorphismus von V. Für jedes $m \geq 0$ wird dann der Endomorphismus $f^m : V \to V$ definiert durch $f^0 = \mathrm{id}_V$, $f^1 = f$ und (für $m \geq 1$) $f^{m+1} = f \circ f^m$.

Der Endomorphismus f heißt *nilpotent*, wenn $f^m = 0$ für ein $m \geq 1$. Also ist f nilpotent genau dann, wenn Bild $f^m = \{0\}$ für ein $m \geq 1$, und man beachte, dass Bild $f^m = \{0\}$ genau dann gilt, wenn Kern $f^m = V$.

Nehme nun an, dass V endlichdimensional ist mit $\dim V = n \geq 1$.

Eine Matrix $A = (a_{ij}) \in \mathrm{M}(n \times n, K)$ heißt *obere Dreiecksmatrix*, wenn $a_{ij} = 0$ für alle $i > j$, und eine Matrix heißt *trigonalisierbar*, wenn sie ähnlich zu einer oberen Dreiecksmatrix ist. Jede diagonalisierbare Matrix ist trigonalisierbar, da eine Diagonalmatrix auch eine obere Dreiecksmatrix ist. Andererseits ist

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

eine obere Dreiecksmatrix und damit trigonalisierbar, aber diese Matrix ist nicht diagonalisierbar.

Eine Folge $V_0, V_1, \ldots, V_n$ von Untervektorräumen von V heißt eine *Fahne*, wenn

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

und $\dim V_j = j$ für jedes $j = 0, \ldots, n$. Ein Endomorphismus $f : V \to V$ von V heißt *trigonalisierbar*, wenn es eine f-invariante Fahne gibt, wobei eine Fahne $V_0, V_1, \ldots, V_n$ *f-invariant* heißt, wenn der Untervektorraum V_j f-invariant für jedes j ist.

Satz 11.1 *Sei $f : V \to V$ ein Endomorphismus von V und sei $A \in \mathrm{M}(n \times n, K)$ die Matrix von f bezüglich irgendeiner Basis von V. Dann ist f trigonalisierbar genau, wenn A trigonalisierbar ist.*

Beweis Nehme zuerst an, dass f trigonalisierbar ist. Es gibt also eine f-invariante Fahne $V_0, V_1, \ldots, V_n$. Nach wiederholter Anwendung des Basisergänzungssatzes kann man eine Basis $(u_1, \ldots, u_n)$ von V konstruieren, so dass $(u_1, \ldots, u_j)$ eine

Basis von V_j ist für jedes $j = 1, \ldots, n$. (Ist $j < n$ und ist $(u_1, \ldots, u_j)$ eine Basis von V_j, so gibt es nach dem Basisergänzungssatz ein $u_{j+1} \in V_{j+1}$, so dass $(u_1, \ldots, u_{j+1})$ eine Basis von V_{j+1} ist.) Sei $B = (b_{ij}) \in \mathrm{M}(n \times n, K)$ die Matrix von f bezüglich der Basis $(u_1, \ldots, u_n)$; dann ist $f(u_j) = b_{1j}u_1 + \cdots + b_{nj}u_n$ für $i = 1, \ldots, n$. Aber $f(u_j) \in V_j = \mathrm{L}(u_1, \ldots, u_j)$ und folglich ist $b_{ij} = 0$ für alle $i > j$, d.h., B ist eine obere Dreiecksmatrix. Damit ist A trigonalisierbar, da nach Satz 8.9 die Matrizen A und B ähnlich sind.

Nehme nun umgekehrt an, dass A trigonalisierbar ist. Dann ist A ähnlich zu einer oberen Dreiecksmatrix $B = (b_{ij}) \in \mathrm{M}(n \times n, K)$ und nach Satz 8.9 gibt es eine Basis $(u_1, \ldots, u_n)$ von V, so dass B die Matrix von f bezüglich $(u_1, \ldots, u_n)$ ist. Sei $V_0 = \{0\}$ und für $j = 1, \ldots, n$ setze $V_j = \mathrm{L}(u_1, \ldots, u_j)$; die Folge $V_0, V_1, \ldots, V_n$ ist also eine Fahne. Seien $1 \leq k \leq j \leq n$; dann ist

$$f(u_k) = b_{1k}u_1 + \cdots + b_{nk}u_n \in \mathrm{L}(u_1, \ldots, u_k) \subset \mathrm{L}(u_1, \ldots, u_j) = V_j \ ,$$

da $b_{ik} = 0$ für $i > k$. Damit ist nach Lemma 5.8 und Satz 3.2

$$f(V_j) = f(\mathrm{L}(u_1, \ldots, u_j)) = \mathrm{L}(f(u_1), \ldots, f(u_j)) \subset \mathrm{L}(u_1, \ldots, u_j) = V_j$$

für $j = 1, \ldots, n$, d.h., $V_0, V_1, \ldots, V_n$ ist eine f-invariante Fahne. Folglich ist f trigonalisierbar. □

Sei $f : V \to V$ trigonalisierbar und sei $V_0, V_1, \ldots, V_n$ eine f-invariante Fahne. Da $\dim V_1 = 1$ und $f(V_1) \subset V_1$, gibt es dann ein $\lambda \in K$, so dass $f(v) = \lambda v$ für jedes $v \in V$. Insbesondere besitzt jeder trigonalisierbare Endomorphismus einen Eigenwert. Nach Satz 11.1 besitzt also jede trigonalisierbare Matrix ebenfalls einen Eigenwert. Folglich ist zum Beispiel die reelle Matrix

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

nicht trigonalisierbar, da sie keinen Eigenwert besitzt. Es wird aber in Satz 11.8 gezeigt werden, dass jeder komplexe Endomorphismus und damit jede komplexe Matrix trigonalisierbar ist.

Hier ist das Hauptergebnis dieses Kapitels:

Satz 11.2 *Sei $f : V \to V$ ein Endomorphismus von V und seien $\lambda_1, \ldots, \lambda_m$ die verschiedenen Eigenwerte von f. Dann sind äquivalent:*

(1) f ist trigonalisierbar.

(2) Jeder f-invariante Untervektorraum U von V mit $\dim U \geq 1$ enthält einen Eigenvektor von f.

(3) Ist U ein f-invarianter Untervektorraum von V mit $\dim U \geq 1$, so besitzt die Einschränkung $f_{|U}$ von f auf U einen Eigenwert.

(4) Es gibt Endomorphismen $g, h : V \to V$ mit g diagonalisierbar, h nilpotent und $g \circ h = h \circ g$, so dass $f = g + h$.

(5) Es gibt eindeutige Endomorphismen $g, h : V \to V$ mit g diagonalisierbar, h nilpotent und $g \circ h = h \circ g$, so dass $f = g + h$.

(6) Es gibt f-invariante Untervektorräume $U_1, \ldots, U_m$ von V mit

$$V = U_1 \oplus \cdots \oplus U_m \ ,$$

so dass für jedes k der Endomorphismus $f_{|U_k} - \lambda_k \mathrm{id}_{U_k}$ von U_k nilpotent ist.

(7) Es gibt eindeutige f-invariante Untervektorräume $U_1, \ldots, U_m$ von V mit

$$V = U_1 \oplus \cdots \oplus U_m \ ,$$

so dass für jedes k der Endomorphismus $f_{|U_k} - \lambda_k \mathrm{id}_{U_k}$ von U_k nilpotent ist, und in der Tat gilt dann $U_k = \mathrm{Kern}\,(f - \lambda_k id_V)^n$ für jedes $k = 1, \ldots, m$.

Beweis Dieser erstreckt sich über die nächsten zehn Seiten. Man beachte aber an dieser Stelle, dass (2) und (3) trivial äquivalent sind: Sei U ein f-invarianter Untervektorraum von V mit $\dim U \geq 1$. Ist $u \in U$ ein Eigenvektor von f, so ist u auch Eigenvektor von $f_{|U}$ und damit besitzt $f_{|U}$ einen Eigenwert. Besitzt umgekehrt $f_{|U}$ einen Eigenwert λ, so gibt es $u \in U \setminus \{0\}$ mit $f_{|U}(u) = \lambda u$, und dann ist $f(u) = \lambda u$, d.h., u ist ein Eigenvektor von f, der in U enthalten ist. □

Bemerkung: In Kapitel 16 wird das charakteristische Polynom eingeführt und es wird gezeigt, dass ein Endomorphismus f genau dann trigonalisierbar ist, wenn sein charakteristisches Polynom χ_f in Linearfaktoren zerfällt.

Lemma 11.1 *Sei $f : V \to V$ trigonalisierbar. Dann ist die Einschränkung $f_{|U}$ von f auf U auch trigonalisierbar für jeden f-invarianten Untervektorraum U mit $\dim U \geq 1$.*

Beweis Sei $V_0, V_1, \ldots, V_n$ eine f-invariante Fahne und für jedes $k = 1, \ldots, n$ sei $U_k = U \cap V_k$. Dann ist U_k ein $f_{|U}$-invarianter Untervektorraum von U und

$$\{0\} = U_0 \subset U_1 \subset \cdots \subset \ U_n = U \ .$$

Ferner gilt nach Satz 4.8, dass für $k = 0, \ldots, n-1$

$$\begin{aligned}
\dim U_{k+1} - \dim U_k &= \dim(U \cap V_{k+1}) - \dim(U \cap V_k) \\
&= (\dim U + \dim V_{k+1} - \dim(U + V_{k+1})) \\
&\qquad\qquad - (\dim U + \dim V_k - \dim(U + V_k)) \\
&= \dim V_{k+1} - \dim V_k - \dim(U + V_{k+1}) + \dim(U + V_k)) \\
&= 1 - \dim(U + V_{k+1}) + \dim(U + V_k) \leq 1
\end{aligned}$$

und daraus ergibt sich, dass $\dim U_{k+1} = \dim U_k$ oder $\dim U_{k+1} = 1 + \dim U_k$ für $k = 0, \ldots, n-1$. Damit gibt es $0 < m_1 < \cdots < m_p = n$ mit $p = \dim U$, so dass $\dim U_{m_j} = j$ für $j = 0, \ldots, p$. Also ist $U_0, U_{m_1}, \ldots, U_{m_p}$ eine $f_{|U}$-invariante Fahne, und daher ist $f_{|U}$ trigonalisierbar. □

Wie schon erwähnt wurde, besitzt jeder trigonalisierbare Endomorphismus einen Eigenwert. Also liefert Lemma 11.1 insbesondere einen Beweis für (1) ⇒ (3) in Satz 11.2.

Bemerkung: Die nachfolgenden Ergebnisse bis einschließlich Lemma 11.7 sind auch richtig ohne die Annahme, dass V endlichdimensional ist.

Lemma 11.2 *Sei $f : V \to V$ ein Endomorphismus. Dann gilt:*

(1) $\text{Bild}\, f^{k+1} \subset \text{Bild}\, f^k$ *für jedes* $k \geq 0$.

(2) Gilt $\text{Bild}\, f^{m+1} = \text{Bild}\, f^m$ *für ein* $m \geq 0$, *so ist* $\text{Bild}\, f^k = \text{Bild}\, f^m$ *für alle* $k \geq m$.

(3) $\text{Kern}\, f^k \subset \text{Kern}\, f^{k+1}$ *für jedes* $k \geq 0$.

(4) Gilt $\text{Kern}\, f^m = \text{Kern}\, f^{m+1}$ *für ein* $m \geq 0$, *so ist* $\text{Kern}\, f^m = \text{Kern}\, f^k$ *für alle* $k \geq m$.

Beweis (1) Sei $k \geq 0$ und $v \in \text{Bild}\, f^{k+1}$; also gibt es $u \in V$ mit $v = f^{k+1}(u)$. Dann ist aber $v = f^k(w)$ mit $w = f(u)$ und folglich ist $v \in \text{Bild}\, f^k$. Daraus ergibt sich, dass $\text{Bild}\, f^{k+1} \subset \text{Bild}\, f^k$.

(2) Sei $m \geq 0$ und nehme an, dass $\text{Bild}\, f^{m+1} = \text{Bild}\, f^m$. Es wird gezeigt, dass dann $\text{Bild}\, f^{m+2} = \text{Bild}\, f^{m+1}$. Sei $v \in \text{Bild}\, f^{m+1}$, es gibt also ein $u \in V$ mit $v = f^{m+1}(u)$. Aber $\text{Bild}\, f^{m+1} = \text{Bild}\, f^m$ und $f^m(u) \in \text{Bild}\, f^m$, und folglich gibt es $w \in V$, so dass $f^m(u) = f^{m+1}(w)$. Damit ist

$$v = f^{m+1}(u) = f(f^m(u)) = f(f^{m+1}(w)) = f^{m+2}(w)\,,$$

d.h., $v \in \text{Bild}\, f^{m+1}$. Dies zeigt, dass $\text{Bild}\, f^{m+1} \subset \text{Bild}\, f^{m+2}$ und nach (1) ist dann $\text{Bild}\, f^{m+2} = \text{Bild}\, f^{m+1}$. Gilt nun $\text{Bild}\, f^{m+1} = \text{Bild}\, f^m$ für ein $m \geq 0$, so ist $\text{Bild}\, f^k = \text{Bild}\, f^m$ für alle $k \geq m$.

(3) Sei $k \geq 0$ und $v \in \text{Kern}\, f^k$; dann gilt $f^{k+1} = f(f^k(v)) = f(0) = 0$ und folglich ist $v \in \text{Kern}\, f^{k+1}$. Damit ist $\text{Kern}\, f^k \subset \text{Kern}\, f^{k+1}$ für jedes $k \geq 0$.

(4) Sei $m \geq 0$ und nehme an, dass $\text{Kern}\, f^m = \text{Kern}\, f^{m+1}$. Es wird gezeigt, dass $\text{Kern}\, f^{m+1} = \text{Kern}\, f^{m+2}$. Sei $v \in \text{Kern}\, f^{m+2}$; dann ist $f^{m+1}(f(v)) = f^{m+2}(v) = 0$, d.h., $f(v) \in \text{Kern}\, f^{m+1}$. Da $\text{Kern}\, f^m = \text{Kern}\, f^{m+1}$, ist dann $f(v) \in \text{Kern}\, f^m$, d.h., $v \in \text{Kern}\, f^{m+1}$. Dies zeigt, dass $\text{Kern}\, f^{m+2} \subset \text{Kern}\, f^{m+1}$ und nach (3) ist damit $\text{Kern}\, f^{m+1} = \text{Kern}\, f^{m+2}$. Gilt nun $\text{Kern}\, f^m = \text{Kern}\, f^{m+1}$ für ein $m \geq 0$, so ist $\text{Kern}\, f^m = \text{Kern}\, f^k$ für alle $k \geq n$. □

Lemma 11.3 *Sei $f : V \to V$ ein Endomorphismus.*

(1) Es gilt $\operatorname{Bild} f^2 = \operatorname{Bild} f$ *genau dann, wenn* $V = \operatorname{Kern} f + \operatorname{Bild} f$.

(2) Es gilt $\operatorname{Kern} f = \operatorname{Kern} f^2$ *genau dann, wenn* $\operatorname{Kern} f \cap \operatorname{Bild} f = \{0\}$.

Beweis (1) Nehme zunächst an, dass $V = \operatorname{Kern} f + \operatorname{Bild} f$. Sei $v \in \operatorname{Bild} f$, es gibt also $v' \in V$, so dass $v = f(v')$ und da $V = \operatorname{Kern} f + \operatorname{Bild} f$, gibt es dann $u \in \operatorname{Kern} f$ und $w \in V$, so dass $v' = u + f(w)$. Folglich ist

$$v = f(v') = f(u + f(w)) = f(u) + f(f(w)) = 0 + f(f(w)) = f^2(w)$$

und damit ist $v \in \operatorname{Bild} f^2$. Dies zeigt, dass $\operatorname{Bild} f \subset \operatorname{Bild} f^2$, und daher ist nach Lemma 11.2 (1) $\operatorname{Bild} f^2 = \operatorname{Bild} f$.

Nehme nun umgekehrt an, dass $\operatorname{Bild} f^2 = \operatorname{Bild} f$. Sei $v \in V$; dann gibt es $u \in V$, so dass $f(v) = f(f(u))$, da $\operatorname{Bild} f^2 = \operatorname{Bild} f$. Dann ist $v = (v - f(u)) + f(u)$ und $f(v - f(u)) = f(v) - f(f(u)) = 0$, d.h., $v - f(u) \in \operatorname{Kern} f$. Dies zeigt, dass $v \in \operatorname{Kern} f + \operatorname{Bild} f$ und damit ist $V = \operatorname{Kern} f + \operatorname{Bild} f$.

(2) Nehme zunächst an, dass $\operatorname{Kern} f \cap \operatorname{Bild} f = \{0\}$, und sei $v \in \operatorname{Kern} f^2$. Dann ist $f(f(v)) = 0$ und folglich ist $f(v) \in \operatorname{Kern} f \cap \operatorname{Bild} f$. Damit ist $f(v) = 0$, d.h., $v \in \operatorname{Kern} f$. Dies zeigt, dass $\operatorname{Kern} f^2 \subset \operatorname{Kern} f$ und nach Lemma 11.2 (3) ist dann $\operatorname{Kern} f = \operatorname{Kern} f^2$.

Nehme nun umgekehrt an, dass $\operatorname{Kern} f = \operatorname{Kern} f^2$, und sei $v \in \operatorname{Kern} f \cap \operatorname{Bild} f$. Da $v \in \operatorname{Kern} f$, gibt es $u \in V$, so dass $v = f(u)$, und dann ist $f(f(u)) = f(v) = 0$, da $v \in \operatorname{Kern} f$. Also ist $u \in \operatorname{Kern} f^2 = \operatorname{Kern} f$, d.h., $v = f(u) = 0$. Dies zeigt, dass $\operatorname{Kern} f \cap \operatorname{Bild} f = \{0\}$. □

Lemma 11.4 *Sei $f : V \to V$ ein nilpotenter Endomorphismus. Dann gilt:*

(1) 0 ist der einzige Eigenwert von f. (D.h: 0 ist ein Eigenwert und es gibt keinen Eigenwert λ mit $\lambda \neq 0$.)

(2) Für jedes $\lambda \in K$ ist λ der einzige Eigenwert von $f_\lambda = \lambda \mathrm{id}_V + f$, und ferner ist $\mathrm{E}(f_\lambda, \lambda) = \mathrm{E}(f, 0)$ *(und natürlich gilt* $\mathrm{E}(f, 0) = \operatorname{Kern} f$*).*

Beweis (1) Sei $m \geq 1$ mit $f^m = 0$. Sei $v \in V \setminus \{0\}$; da $f^0(v) = v \neq 0$ und $f^m(v) = 0$, gibt es ein p mit $0 \leq p < m$, so dass $f^p(v) \neq 0$ aber $f^{p+1}(v) = 0$. Setze $u = f^p(v)$; dann ist $u \neq 0$ und $f(u) = 0 = 0u$. Also ist 0 ein Eigenwert von f (und u ist ein Eigenvektor zum Eigenwert 0).

Sei nun $\lambda \in K$ ein Eigenwert von f und v ein Eigenvektor zum Eigenwert λ. Dann gilt $f(v) = \lambda v$ und daraus folgt durch Induktion nach k, dass $f^k(v) = \lambda^k v$ für alle $k \geq 1$: Gilt $f^k(v) = \lambda^k v$, so ist auch

$$f^{k+1}(v) = f(f^k(v)) = f(\lambda^k v) = \lambda^k f(v) = \lambda^k \lambda v = \lambda^{k+1} v \,.$$

Insbesondere ist $\lambda^m v = f^m(v) = 0$ und damit ist $\lambda^m = 0$, da $v \neq 0$. Also ist $\lambda = 0$. Dies zeigt, dass 0 der einzige Eigenwert von f ist.

(2) Dies folgt aus (1), da $f_\lambda(v) = \mu v$ genau dann gilt, wenn $f(v) = (\mu - \lambda)v$. □

Lemma 11.5 *Seien $f, g : V \to V$ Endomorphismen von V mit $f \circ g = g \circ f$.*

(1) Die Untervektorräume $\operatorname{Kern} g$ *und* $\operatorname{Bild} g$ *sind beide f-invariant.*

(2) Es gilt $f \circ g^m = g^m \circ f$ für jedes $m \geq 1$.

(3) Die Untervektorräume $\operatorname{Kern} g^m$ *und* $\operatorname{Bild} g^m$ *sind beide f-invariant für jedes $m \geq 1$.*

(4) Für jeden Eigenwert λ von g ist der Eigenraum $\mathrm{E}(g, \lambda)$ f-invariant.

Beweis (1) Sei $v \in \operatorname{Kern} g$; dann gilt $g(f(v)) = f(g(v)) = f(0) = 0$ und damit ist $f(v) \in \operatorname{Kern} g$; folglich ist $\operatorname{Kern} g$ f-invariant. Sei nun $v \in \operatorname{Bild} g$, es gibt also $w \in V$ mit $v = g(w)$ und dann ist $f(v) = f(g(w)) = g(f(w)) \in \operatorname{Bild} g$. Daher ist auch $\operatorname{Bild} g$ f-invariant.

(2) Dies folgt unmittelbar durch Induktion nach m.

(3) Dies folgt unmittelbar aus (1) und (2).

(4) Sei $v \in \mathrm{E}(g, \lambda)$; dann gilt $g(f(v)) = f(g(v)) = f(\lambda v) = \lambda f(v)$ und damit ist $f(v) \in \mathrm{E}(g, \lambda)$. □

Lemma 11.6 *Sei $f : V \to V$ ein Endomorphismus und sei U ein f-invarianter Untervektorraum von V. Dann ist jeder Eigenwert von $f_{|U}$ auch ein Eigenwert von V.*

Beweis Dies ist klar. □

Lemma 11.7 *Sei $f : V \to V$ ein Endomorphismus und seien U, W f-invariante Untervektorräume von V mit $V = U \oplus W$. Dann ist jeder Eigenwert von V ein Eigenwert von $f_{|U}$ oder ein Eigenwert von $f_{|W}$.*

Beweis Sei v ein Eigenvektor von f zum Eigenwert λ. Da $V = U \oplus W$, hat v eine eindeutige Darstellung $v = u + w$ mit $u \in U$ und $w \in W$, und da U und W f-invariant sind, ist $f(u) \in U$ und $f(w) \in W$. Damit ist $f(u) - \lambda u \in U$ und $f(w) - \lambda w \in W$ und es gilt

$$(f(u) - \lambda u) + (f(w) - \lambda w) = f(u + w) - \lambda(u + w) = f(v) - \lambda v = 0 \, .$$

Daraus egibt sich, dass $f(u) - \lambda u = f(w) - \lambda w = 0$, da $U \cap W = \{0\}$. Also ist λ ein Eigenwert von $f_{|U}$ oder ein Eigenwert von $f_{|W}$, da mindestens eines von u und w nicht gleich Null ist. □

Erinnerung: Es gilt $\dim V = n \geq 1$.

Lemma 11.8 *Sei $f : V \to V$ ein Endomorphismus. Dann gilt* $\operatorname{Bild} f^k = \operatorname{Bild} f^n$ *und* $\operatorname{Kern} f^k = \operatorname{Kern} f^n$ *für alle $k \geq n$. Insbesondere ist f genau dann nilpotent, wenn $f^n = 0$.*

Beweis Für $k \geq 0$ setze $d_k = \dim \operatorname{Bild} f^k$. Nach Lemma 11.2 (1) ist $d_{k+1} \leq d_k$ für jedes k und folglich gibt es $d \geq 0$ und $p \geq 0$, so dass $d_k = d$ für alle $k \geq p$. Sei $q = \min\{k \geq 0 : d_k = d\}$. Dann ist $d_k = d$ für alle $k \geq q$ und nach Lemma 11.2 (2) ist $d_{k+1} < d_k$ für alle $0 \leq k < q$. Aber $d_0 = n$ und $d \geq 0$ und daraus ergibt sich, dass $q \leq n$. Insbesondere ist $d_k = d$ für alle $k \geq n$ und also ist nach Satz 4.7 $\operatorname{Bild} f^k = \operatorname{Bild} f^n$ für alle $k \geq n$. Genauso gilt auch $\operatorname{Kern} f^k = \operatorname{Kern} f^n$ für alle $k \geq n$. □

Lemma 11.9 *Sei $f : V \to V$ ein Endomorphismus. Dann sind äquivalent:*

(1) $\operatorname{Bild} f^2 = \operatorname{Bild} f$.

(2) $V = \operatorname{Kern} f + \operatorname{Bild} f$.

(3) $\operatorname{Kern} f = \operatorname{Kern} f^2$.

(4) $\operatorname{Kern} f \cap \operatorname{Bild} f = \{0\}$.

Beweis (1) $\Leftrightarrow$ (2) und (3) $\Leftrightarrow$ (4) sind Lemma 11.3.

(2) $\Leftrightarrow$ (4): Nach Satz 4.8 ist

$$\dim(\operatorname{Kern} f + \operatorname{Bild} f) = \dim \operatorname{Kern} f + \dim \operatorname{Bild} f - \dim(\operatorname{Kern} f \cap \operatorname{Bild} f)$$

und nach der Dimensionsformel (Satz 5.2) ist $\dim V = \dim \operatorname{Kern} f + \dim \operatorname{Bild} f$. Daraus ergibt sich, dass

$$\dim V = \dim(\operatorname{Kern} f + \operatorname{Bild} f) + \dim(\operatorname{Kern} f \cap \operatorname{Bild} f)\ .$$

Außerdem folgt aus Satz 4.7, dass $\dim V = \dim(\operatorname{Kern} f + \operatorname{Bild} f)$ genau dann gilt, wenn $V = \operatorname{Kern} f + \operatorname{Bild} f$. Damit gilt $V = \operatorname{Kern} f + \operatorname{Bild} f$ genau dann, wenn $\dim(\operatorname{Kern} f \cap \operatorname{Bild} f) = 0$, d.h., genau dann wenn, $\operatorname{Kern} f \cap \operatorname{Bild} f = \{0\}$. □

Lemma 11.10 *Sei $f : V \to V$ ein Endomorphismus und setze $U = \operatorname{Kern} f^n$ und $W = \operatorname{Bild} f^n$. Dann gilt:*

(1) Die Untervektorräume U und W sind f-invariant.

(2) $V = U \oplus W$.

(3) Die Einschränkung $f_{|U}$ von f auf U ist nilpotent.

(4) Die Einschränkung $f_{|W}$ von f auf W ist ein Automorphismus.

Beweis (1) Dies folgt unmittelbar aus Lemma 11.5 (3), da $f \circ f = f \circ f$.

(2) Nach Lemma 11.8 gilt $\operatorname{Bild} f^k = \operatorname{Bild} f^n$ für alle $k \geq n$ und insbesondere ist $\operatorname{Bild} (f^n)^2 = \operatorname{Bild} f^{2n} = \operatorname{Bild} f^n$. Daraus ergibt sich nach Lemma 11.9, dass $V = \operatorname{Kern} f^n + \operatorname{Bild} f^n$ und $\operatorname{Kern} f^n \cap \operatorname{Bild} f^n = \{0\}$, d.h., $V = U \oplus W$.

(3) Für jedes $u \in U = \operatorname{Kern} f^n$ ist $f^n_{|U}(u) = f^n(u) = 0$. Daher ist $f^n_{|U} = 0$ und insbesondere ist $f_{|U}$ nilpotent.

(4) Sei $w \in W = \operatorname{Bild} f^n$; da $\operatorname{Bild} f^{n+1} = \operatorname{Bild} f^n$ gibt es dann $v \in V$, so dass $w = f^{n+1}(v)$. Aber $w = f(f^n(v))$ und $f^n(v) \in \operatorname{Bild} f^n = W$. Dies zeigt dass, $f(W) = W$ und damit ist der Endomorphismus $f_{|W} : W \to W$ surjektiv. Nach Satz 5.3 ist also $f_{|W}$ ein Automorphismus, da nach Satz 4.4 W endlichdimensional ist. □

Lemma 11.11 *Sei $f : V \to V$ ein Endomorphismus von V, der mindestens einen Eigenwert besitzt, und seien $\lambda_1, \ldots, \lambda_m$ die verschiedenen Eigenwerte. Dann gibt es f-invariante Untervektorräume U, W von V mit $V = U \oplus W$, so dass gilt:*

(1) Der Endomorphismus $f_{|U} - \lambda_1 \mathrm{id}_U$ von U ist nilpotent.

(2) $\lambda_2, \ldots, \lambda_m$ sind die verschiedenen Eigenwerte von $f_{|W}$.

Beweis Sei $g = f - \lambda_1 \mathrm{id}_V$ und setze $U = \operatorname{Kern} g^n$, $W = \operatorname{Bild} g^n$. Nach Lemma 11.5 sind die Untervektorräume U und W f-invariant, da

$$g \circ f = (f - \lambda_1 \mathrm{id}_V) \circ f = f^2 - \lambda_1 f = f \circ (f - \lambda_1 \mathrm{id}_V) = f \circ g \, .$$

Nach Lemma 11.10 (angewendet auf g) ist $V = U \oplus W$, die Untervektorräume U und W sind g-invariant, die Einschränkung $g_{|U}$ von g auf U ist nilpotent und die Einschränkung $g_{|W}$ von g auf W ist ein Automorphismus.

Nun ist $f_{|U} - \lambda_1 \mathrm{id}_U = g_{|U}$ und damit ist $f_{|U} - \lambda_1 \mathrm{id}_U$ nilpotent.

Ferner ist $f_{|W} - \lambda_1 \mathrm{id}_W = g_{|W}$, somit ist $f_{|W} - \lambda_1 \mathrm{id}_W$ ein Automorphismus und insbesondere ist λ_1 kein Eigenwert von $f_{|W}$. Da $f_{|U} = \lambda_1 \mathrm{id}_U + (f_{|U} - \lambda_1 \mathrm{id}_U)$ und $f_{|U} - \lambda_1 \mathrm{id}_U$ nilpotent ist, ist nach Lemma 11.4 (2) λ_1 der einzige Eigenwert von $f_{|U}$ und daraus ergibt sich nach Lemma 11.7, dass $\lambda_2, \ldots, \lambda_m$ Eigenwerte von $f_{|W}$ sind. Damit sind nach Lemma 11.6 $\lambda_2, \ldots, \lambda_m$ genau die verschiedenen Eigenwerte von $f_{|W}$. □

Satz 11.3 *Sei $f : V \to V$ ein Endomorphismus von V und seien $\lambda_1, \ldots, \lambda_m$ die verschiedenen Eigenwerte von f. Dann gibt es f-invariante Untervektorräume $U_1, \ldots, U_m$ und U_* von V mit $V = U_1 \oplus \cdots \oplus U_m \oplus U_*$, so dass gilt:*

(1) Der Endomorphismus $f_{|U_k} - \lambda_k \mathrm{id}_{U_k}$ von U_k ist nilpotent für jedes k.

(2) Der Endomorphismus $f_{|U_*}$ von U_* besitzt keinen Eigenwert.

Beweis Für $m \geq 0$ sei A_m die folgende Aussage: Ist V' ein endlichdimensionaler Vektorraum und ist $g : V' \to V'$ ein Endomorphismus von V' mit genau m verschiedenen Eigenwerten $\mu_1, \ldots, \mu_m$, so gibt es g-invariante Untervektorräume $U'_1, \ldots, U'_m$ und U'_* von V' mit $V' = U'_1 \oplus \cdots \oplus U'_m \oplus U'_*$, so dass gilt:

(1) Der Endomorphismus $g_{|U'_k} - \mu_k \mathrm{id}_{U'_k}$ von U'_k ist nilpotent für jedes k.

(2) Der Endomorphismus $g_{|U'_*}$ von U'_* besitzt keinen Eigenwert.

Es genügt also zu zeigen, dass A_m richtig ist für jedes $m \geq 0$.

Die Aussage A_0 ist trivial richtig (mit $U'_* = V'$). Sei $m \geq 1$ und nehme an, dass die Aussage A_{m-1} richtig ist. Sei V' ein endlichdimensionaler Vektorraum und $g : V' \to V'$ ein Endomorphismus mit genau m verschiedenen Eigenwerten $\mu_1, \ldots, \mu_m$. Nach Lemma 11.11 gibt es dann g-invariante Untervektorräume U'_1 und W' von V' mit $V' = U'_1 \oplus W'$ so dass gilt:

(1) Der Endomorphismus $g_{|U'_1} - \mu_1 \mathrm{id}_{U'_1}$ von U' ist nilpotent.

(2) $\mu_2, \ldots, \mu_m$ sind die verschiedenen Eigenwerte von $g_{|W'}$.

Da A_{m-1} richtig ist, gibt es $g_{|W'}$-invariante Untervektorräume $U'_2, \ldots, U'_m$ und U'_* von W' mit $W' = U'_2 \oplus \cdots \oplus U'_m \oplus U'_*$, so dass gilt:

(1) Der Endomorphismus $g_{|U'_k} - \mu_k \mathrm{id}_{U'_k}$ von U'_k ist nilpotent für $k = 2, \ldots, m$.

(2) Der Endomorphismus $g_{|U'_*}$ von U'_* besitzt keinen Eigenwert.

(Man beachte: $g_{|U'_k}$ ist die Einschränkung von $g_{|W'}$ auf U'_k für jedes k und $g_{|U'_*}$ ist die Einschränkung von $g_{|W'}$ auf U'_*.) Nach Satz 9.2 ist aber

$$V' = U'_1 \oplus W' = U'_1 \oplus (U'_2 \oplus \cdots \oplus U'_m \oplus U'_*) = U'_1 \oplus U'_2 \oplus \cdots \oplus U'_m \oplus U'_*$$

und dies zeigt, dass die Aussage A_m richtig ist. Durch Induktion nach m ist also A_m richtig für jedes $m \geq 0$. □

Lemma 11.12 *Jeder nilpotente Endomorphismus von V ist trigonalisierbar.*

Beweis Sei $h : V \to V$ nilpotent und setze $m = \min\{k \geq 1 : h^k = 0\}$. Für jedes $k = 0, \ldots, m$ sei $U_k = \operatorname{Kern} h^k$; nach Lemma 11.2 (3) und (4) ist dann

$$\{0\} = U_0 \subset U_1 \subset \cdots \subset U_m = V \ ,$$

und $U_k \neq U_{k+1}$ für jedes $k = 0, \ldots, m-1$. Für $k = 1, \ldots, m$ sei $p_k = \dim U_k$; nach wiederholter Anwendung des Basisergänzungssatzes kann man eine Basis $(u_1, \ldots, u_n)$ von V konstruieren, so dass $(u_1, \ldots, u_{p_k})$ eine Basis von U_k ist für jedes $k = 1, \ldots, m$. Für $j = 1, \ldots, n$ sei nun $V_j = \mathrm{L}(u_1, \ldots, u_j)$; mit $V_0 = \{0\}$ ist dann $V_0, \ldots, V_n$ eine Fahne, und da $h(U_k) \subset U_{k-1}$ für $k = 1, \ldots, m$, ist diese Fahne h-invariant. □

Satz 11.4 *Seien $g, h : V \to V$ Endomorphismen von V mit g diagonalisierbar, h nilpotent und $g \circ h = h \circ g$. Dann ist $f = g + h$ trigonalisierbar.*

Beweis Nach Satz 10.5 gilt $V = U_1 \oplus \cdots \oplus U_m$, wobei $U_k = \mathrm{E}(g, \lambda_k)$ und $\lambda_1, \ldots, \lambda_m$ die verschiedenen Eigenwerte von g sind. Für $k = 1, \ldots, m$ ist nach Lemma 11.5 (4) U_k h-invariant und es ist klar, dass $h_{|U_k}$ nilpotent ist. Damit ist nach Lemma 11.12 $h_{|U_k}$ trigonalisierbar und folglich gibt es nach Satz 11.1 und Satz 8.9 eine Basis $\beta_k = (u_1^k, \ldots, u_{n_k}^k)$ von U_k, so dass die Matrix A_k von $h_{|U_k}$ bezüglich β_k eine obere Dreiecksmatrix ist. Nun ist auch U_k f-invariant und die Matrix von $f_{|U_k}$ bezüglich β_k ist die obere Dreiecksmatrix $\lambda_k E_{n_k} + A_k$. Sei jetzt

$$\beta = (u_1^1, \ldots, u_{n_1}^1, \ldots, u_1^m, \ldots, u_{n_m}^m) \; ;$$

nach Satz 9.7 ist β eine Basis von V und die Matrix von f bezüglich β ist eine obere Dreiecksmatrix. Damit ist nach Satz 11.1 f trigonalisierbar. □

Seien $U_1, \ldots, U_m, U$ Untervektorräume von V mit $U = U_1 \oplus \cdots \oplus U_m$ und für jedes $k = 1, \ldots, m$ sei $f_k : U_k \to U_k$ ein Endomorphismus von U_k. Dann gibt es nach Satz 9.4 den Endomorphismus $f = f_1 \oplus \cdots \oplus f_m$ von U: f ist der eindeutige Endomorphismus mit $f(u) = f_k(u)$ für alle $u \in U_k$, $k = 1, \ldots, m$. Insbesondere ist dann U_k f-invariant und $f_k = f_{|U_k}$ für jedes $k = 1, \ldots, m$. Ist umgekehrt $f : U \to U$ ein Endomorphismus von U mit U_k f-invariant für jedes k, so ist nach Lemma 9.1 $f = f_{|U_1} \oplus \cdots \oplus f_{|U_m}$.

Gilt $f = f_1 \oplus \cdots \oplus f_m$, so ist nach Satz 9.5 auch $f^k = f_1^k \oplus \cdots \oplus f_m^k$ für jedes $k \geq 1$.

Lemma 11.13 *Sei $f : V \to V$ ein Endomorphismus und seien $\lambda_1, \ldots, \lambda_m$ die verschiedenen Eigenwerte. Nehme an, es gibt f-invariante Untervektorräume $U_1, \ldots, U_m$ von V mit $V = U_1 \oplus \cdots \oplus U_m$, so dass $f_{|U_k} - \lambda_k \mathrm{id}_{U_k}$ nilpotent ist für jedes k. Dann gilt $U_k = \mathrm{Kern}\,(f - \lambda_k \mathrm{id}_V)^n$ für jedes $k = 1, \ldots, m$.*

Beweis Für jedes k sei f_k die Einschränkung von f auf U_k und $h_k = f_k - \lambda_k id_{U_k}$; also gilt $f_k = \lambda_k \mathrm{id}_{U_k} + h_k$ und h_k ist nilpotent. Nun ist

$$\begin{aligned} f - \lambda_k \mathrm{id}_V &= (f_1 - \lambda_k \mathrm{id}_{U_1}) \oplus \cdots \oplus (f_m - \lambda_k \mathrm{id}_{U_m}) \\ &= ((\lambda_1 - \lambda_k)\mathrm{id}_{U_1} + h_1) \oplus \cdots \oplus ((\lambda_m - \lambda_k)\mathrm{id}_{U_m} + h_m) \; , \end{aligned}$$

und folglich gilt nach Satz 9.5 und Lemma 9.2, dass

$$\begin{aligned} &\mathrm{Kern}\,(f - \lambda_k \mathrm{id}_V)^n \\ &\quad = \mathrm{Kern}\,((\lambda_1 - \lambda_k)\mathrm{id}_{U_1} + h_1)^n \oplus \cdots \oplus ((\lambda_m - \lambda_k)\mathrm{id}_{U_m} + h_m)^n \\ &\quad = \mathrm{Kern}\,((\lambda_1 - \lambda_k)\mathrm{id}_{U_1} + h_1)^n \oplus \cdots \oplus \mathrm{Kern}\,((\lambda_m - \lambda_k)\mathrm{id}_{U_m} + h_m)^n \\ &\quad = \{0\} \oplus \cdots \oplus \{0\} \oplus \mathrm{Kern}\, h_k^n \oplus \{0\} \oplus \cdots \oplus \{0\} \\ &\quad = \mathrm{Kern}\, h_k^n = U_k \; , \end{aligned}$$

da für jedes $j \neq k$ der Endomorphismus $(\lambda_j - \lambda_k)\mathrm{id}_{U_j} + h_j$ und damit auch $((\lambda_j - \lambda_k)\mathrm{id}_{U_j} + h_j)^n$ ein Automorphismus ist. □

Beweis für Satz 11.2: Man erinnert sich, dass (2) ⇔ (3) trivial richtig ist und auch, dass (1) ⇒ (3) unmittelbar aus Lemma 11.1 folgt. Ferner ist (5) ⇒ (4) trivial richtig.

(3) ⇒ (6): Dies folgt unmittelbar aus Satz 11.3.

(6) ⇒ (7): Dies folgt unmittelbar aus Lemma 11.13.

(7) ⇒ (4): Seien $\lambda_1, \ldots, \lambda_m$ die verschiedenen Eigenwerte von f. Es gibt dann f-invariante Untervektorräume $U_1, \ldots, U_m$ mit $V = U_1 \oplus \cdots \oplus U_m$ so dass für jedes k der Endomorphismus $f_{|U_k} - \lambda_k \mathrm{id}_{U_k}$ von U_k nilpotent ist.

Setze $g = g_1 \oplus \cdots \oplus g_m$ und $h = h_1 \oplus \cdots \oplus h_m$, wobei $h_k = f_{|U_k} - \lambda_k \mathrm{id}_{U_k}$ und $g_k = \lambda_k \mathrm{id}_{U_k}$. Dann gilt $f_{|U_k} = g_k + h_k$ und

$$\begin{aligned} g_k \circ h_k = \lambda_k \mathrm{id}_{U_k} \circ (f_{|U_k} - \lambda_k \mathrm{id}_{U_k}) &= \lambda_k f_{|U_k} - \lambda_k^2 \mathrm{id}_{U_k} \\ &= (f_{|U_k} - \lambda_k \mathrm{id}_{U_k}) \circ \lambda_k \mathrm{id}_{U_k} = h_k \circ g_k \end{aligned}$$

für jedes k, und daraus ergibt sich nach Satz 9.5, dass $f = g + h$ und $g \circ h = h \circ g$. Ferner ist h nilpotent: Da h_k nilpotent ist und $\dim U_k \leq n$, ist nach Lemma 11.7 $h_k^n = 0$ für jedes k. Damit ist

$$\begin{aligned} h^n(u_1 + \cdots + u_m) &= (h_1 \oplus \cdots \oplus h_m)^n (u_1 + \cdots + u_m) \\ &= (h_1^n \oplus \cdots \oplus h_m^n)(u_1 + \cdots + u_m) = h_1^n(u_1) + \cdots + h_m^n(u_m) = 0 \end{aligned}$$

für alle $u_k \in U_k$, $k = 1, \ldots, m$, und folglich ist $h^n = 0$. Schließlich ist auch g diagonalisierbar: Dies folgt aus Satz 9.7 und Lemma 11.5, da jede Basis von U_k aus Eigenvektoren von $g_k = \lambda_k \mathrm{id}_{U_k}$ besteht.

(4) ⇒ (1): Dies ist Satz 11.4.

(4) ⇒ (5): Es gibt Endomorphismen $g, h : V \to V$ mit g diagonalisierbar, h nilpotent und $g \circ h = h \circ g$, so dass $f = g + h$. Seien $\lambda_1, \ldots, \lambda_m$ die verschiedenen Eigenwerte von g und für $k = 1, \ldots, m$ sei $U_k = \mathrm{E}(g, \lambda_k)$. Nach Satz 10.5 ist V die direkte Summe der Untervektorräume $U_1, \ldots, U_m$, d.h., $V = U_1 \oplus \cdots \oplus U_m$.

Da $g \circ h = h \circ g$, ist nach Lemma 11.5 (4) U_k h-invariant und damit ist U_k auch f-invariant (da U_k g-invariant ist und $f = g + h$). Für jedes k sei f_k (bzw. g_k bzw. h_k) die Einschränkung von f (bzw. g bzw. h) auf U_k. Also gilt $f_k = g_k + h_k$ für jedes k, $f = f_1 \oplus \cdots \oplus f_m$, $g = g_1 \oplus \cdots \oplus g_m$ und $h = h_1 \oplus \cdots \oplus h_m$. Ferner ist $g_k = \lambda_k \mathrm{id}_{U_k}$.

Nun ist $h_k = f_k - \lambda_k \mathrm{id}_{U_k}$ nilpotent; damit ist nach Lemma 11.4 (2) λ_k der einzige Eigenwert von $f_k = \lambda_k \mathrm{id}_{U_k} + h_k$ und insbesondere ist nach Lemma 11.6 λ_k ein Eigenwert von f. Sei $\lambda \notin \{\lambda_1, \ldots, \lambda_m\}$; nach Lemma 11.4 (2) ist 0 kein Eigenwert

von $f_k - \lambda \mathrm{id}_{U_k}$, also ist $f_k - \lambda \mathrm{id}_{U_k}$ ein Automorphismus von U_k und folglich ist nach Satz 9.6 $f - \lambda \mathrm{id}_V = (f_1 - \lambda \mathrm{id}_{U_1}) \oplus \cdots \oplus (f_m - \lambda \mathrm{id}_{U_m})$ ein Automorphismus von V, d.h., λ ist kein Eigenwert von f. Daraus ergibt sich, dass $\lambda_1, \ldots, \lambda_m$ genau die Eigenwerte von f sind und nach Lemma 11.13 ist dann $U_k = \mathrm{Kern}\,(f - \lambda_k \mathrm{id}_V)^n$ für jedes $k = 1, \ldots, m$.

Dies zeigt, dass $g = \lambda_1 \mathrm{id}_{U_1} \oplus \cdots \oplus \lambda_m \mathrm{id}_{U_m}$ und $h = f - g$, wobei $\lambda_1, \ldots, \lambda_m$ die Eigenwerte von f sind und $U_k = \mathrm{Kern}\,(f - \lambda_k \mathrm{id}_V)^n$. Insbesondere sind g und h eindeutig durch f bestimmt. □

Satz 11.5 *Sei $K = \mathbb{C}$; dann besitzt jeder Endomorphismus $f : V \to V$ einen Eigenwert.*

Beweis Nach Satz 8.7 hat der Vektorraum $\mathrm{End}(V)$ aller Endomorphismen von V die Dimension n^2 und folglich sind die Elemente $f^0, f^1, \ldots, f^{n^2}$ linear abhängig. Da $f^0 = \mathrm{id}_V \neq 0$, gibt es dann p mit $1 \leq p \leq n^2$, so dass $f^0, f^1, \ldots, f^{p-1}$ linear unabhängig sind aber $f^p \in \mathrm{L}(f^0, f^1, \ldots, f^{p-1})$ und also gibt es $c_0, \ldots, c_{p-1} \in \mathbb{C}$ mit $f^p = c_0 f^0 + \cdots + c_{p-1} f^{p-1}$. Nach dem Fundamentalsatz der Algebra (siehe Satz 16.4) gibt es nun $\lambda_0, \ldots, \lambda_{p-1} \in \mathbb{C}$, so dass

$$z^p - c_{p-1} z^{p-1} - \cdots - c_1 z - c_0 = (z - \lambda_0) \times \cdots \times (z - \lambda_{p-1})$$

für alle $z \in \mathbb{C}$ und daraus folgt (warum?), dass

$$(f - \lambda_0 \mathrm{id}_V) \circ \cdots \circ (f - \lambda_{p-1} \mathrm{id}_V) = f^p - c_{p-1} f^{p-1} - \cdots - c_1 f - c_0 \mathrm{id}_V = 0\,.$$

Aber $(f - \lambda_1 \mathrm{id}_V) \circ \cdots \circ (f - \lambda_{p-1} \mathrm{id}_V) \neq 0$, sonst wären $f^0, f^1, \ldots, f^{p-1}$ linear abhängig und folglich gibt es ein $v \in V$, so dass

$$u = (f - \lambda_1 \mathrm{id}_V) \circ \cdots \circ (f - \lambda_{p-1} \mathrm{id}_V)\,(v) \neq 0\,.$$

Dann ist $u \neq 0$ und $(f - \lambda_0 \mathrm{id}_V)(u) = 0$, d.h., λ_0 ist ein Eigenwert von f. □

Satz 11.6 *Sei $K = \mathbb{C}$; dann ist jeder Endomorphismus f von V trigonalisierbar.*

Beweis Sei U ein f-invarianter Untervektorraum von V mit $\dim U \geq 1$. Nach Satz 11.5 (angewendet auf $f_{|U}$) besitzt $f_{|U}$ einen Eigenwert und daraus ergibt sich nach Satz 11.2 (3) $\Rightarrow$ (1), dass f trigonalisierbar ist. □

Nach Satz 11.1 und Satz 11.6 ist jede komplexe Matrix trigonalisierbar.

12 Jordansche Normalform

Im Folgenden sei K ein Körper und sei V ein endlichdimensionaler Vektorraum über K mit $\dim V = n \geq 1$.

Sei $f : V \to V$ ein trigonalisierbarer Endomorphismus. Dieses Kapitel beschäftigt sich mit der folgenden Frage: Wie kann man eine Basis von V wählen, so dass die Matrix von f bezüglich dieser Basis so einfach wie möglich ist? Nach dem Beweis für Satz 11.1 gibt es eine Basis, so dass die Matrix eine obere Dreiecksmatrix ist. Satz 12.5 verfeinert diese Aussage: Es gibt eine Basis, so dass die Matrix von f folgende Block-Gestalt hat:

$$\begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s \end{pmatrix},$$

wobei $J_1, \ldots, J_s$ Jordanmatrizen sind. Eine *(elementare) Jordanmatrix* ist eine Matrix aus der Familie $\{J_k(\lambda) : k \geq 1,\ \lambda \in K\}$, wobei $J_k(\lambda)$ das folgende Element von $\mathrm{M}(k \times k, K)$ ist:

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}.$$

$J_k(\lambda)$ heißt Jordanmatrix *zum Eigenwert* λ. (Warum?)

Sei $f : V \to V$ trigonalisierbar und seien $\lambda_1, \ldots, \lambda_m$ die verschiedenen Eigenwerte von f. Nach Satz 11.2 gibt es f-invariante Untervektorräume $U_1, \ldots, U_m$ von V mit $V = U_1 \oplus \cdots \oplus U_m$, so dass für jedes k der Endomorphismus $f_{|U_k} - \lambda_k \mathrm{id}_{U_k}$ von U_k nilpotent ist. Setze $f_k = f_{|U_k}$ und $h_k = f_{|U_k} - \lambda_k \mathrm{id}_{U_k}$; also sind f_k und h_k Endomorphismen von U_k mit $f_k = \lambda_k \mathrm{id}_{U_k} + h_k$ und h_k ist nilpotent. Für jedes k sei $(u_1^k, \ldots, u_{p_k}^k)$ eine Basis von U_k und sei B_k die Matrix von h_k bezüglich dieser Basis. Dann ist $A_k = \lambda_k E_{p_k} + B_k$ die Matrix von f_k bezüglich $(u_1^k, \ldots, u_{p_k}^k)$. Sei

$$(v_1, \ldots, v_n) = (u_1^1, \ldots, u_{p_1}^1, \ldots, u_1^m, \ldots, u_{p_m}^m);$$

nach Satz 9.7 ist dann $(v_1, \ldots, v_n)$ eine Basis von V und die Matrix von f bezüglich dieser Basis ist

$$\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_m \end{pmatrix}.$$

Die Ausgangsfrage für trigonalisierbare Endomorphismen reduziert sich daher im Wesentlichen auf die analoge Frage für nilpotente Endomorphismen und zunächst werden also nilpotente Endomorphismen untersucht.

Sei $f : V \to V$ ein Endomorphismus von V und für jedes $v \in V$ setze

$$Z(v, f) = \bigcup_{k \geq 1} \mathrm{L}(v, f(v), \ldots, f^k(v)) \; .$$

Lemma 12.1 *Für jedes $v \in V$ ist $Z(v, f)$ der kleinste f-invariante Untervektorraum von V, der v enthält.*

Beweis Übung. □

Lemma 12.2 *Sei $v \in V \setminus \{0\}$ und nehme an, dass $f^k(v) = 0$ für ein $k \geq 1$; setze $m = \min\{k \geq 1 : f^k(v) = 0\}$. Dann ist $Z(v, f) = \mathrm{L}(v, f(v), \ldots, f^{m-1}(v))$ und $(v, f(v), \ldots, f^{m-1}(v))$ ist eine Basis von $Z(v, f)$.*

Beweis Für jedes $k \geq m$ ist $\mathrm{L}(v, f(v), \ldots, f^k(v)) = \mathrm{L}(v, f(v), \ldots, f^{m-1}(v))$, da $f^j(v) = 0$ für $j = m, \ldots, k$, und $\mathrm{L}(v, f(v), \ldots, f^k(v)) \subset \mathrm{L}(v, f(v), \ldots, f^{m-1}(v))$ für $k = 1, \ldots, m-1$. Damit ist $Z(v, f) = \mathrm{L}(v, f(v), \ldots, f^{m-1}(v))$. Es bleibt also zu zeigen, dass die Vektoren v, $f(v)$, $\ldots$, $f^{m-1}(v)$ linear unabhängig sind. Seien $\lambda_0, \ldots, \lambda_{m-1} \in K$ mit $\lambda_0 v + \lambda_1 f(v) + \cdots + \lambda_{m-1} f^{m-1}(v) = 0$ und nehme an, dass $\lambda_j \neq 0$ für ein j; setze $k = \min\{0 \leq j \leq m-1 : \lambda_j \neq 0\}$. Dann ist $\lambda_k f^k(v) + \cdots + \lambda_{m-1} f^{m-1}(v) = 0$ und daraus ergibt sich, dass

$$\begin{aligned} 0 &= f^{m-1-k}(\lambda_k f^k(v) + \cdots + \lambda_{m-1} f^{m-1}(v)) \\ &= \lambda_k f^{m-1-k}(f^k(v)) + \cdots + \lambda_{m-1} f^{m-1-k}(f^{m-1}(v)) \\ &= \lambda_k f^{m-1}(v) + \lambda_{k+1} f^m(v) + \cdots + \lambda_{m-1} f^{2m-2-k}(v) = \lambda_k f^{m-1}(v) \; , \end{aligned}$$

da $f^j(v) = 0$ für alle $j \geq m$. Aber dies ist nicht möglich, da $f^{m-1}(v) \neq 0$ und $\lambda_k \neq 0$, und folglich muss $\lambda_j = 0$ für jedes j sein, d.h., v, $f(v)$, $\ldots$, $f^{m-1}(v)$ sind linear unabhängig. □

Sei $v \in V \setminus \{0\}$ und m wie im Lemma 12.2 und sei g die Einschränkung von f auf $Z(v, f)$; dann ist natürlich $(f^{m-1}(v), \ldots, f(v), v)$ auch eine Basis von $Z(v, f)$ und die Matrix von g bezüglich dieser Basis ist

$$\begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix} \; ,$$

da $g(f^{m-1}(v)) = f^n(v) = 0$ und $g(f^k(v)) = f^{k+1}(v)$ für $k = 0, \ldots, m-2$. Diese Matrix ist eine Jordanmatrix zum Eigenwert 0 und wird auch *nilpotente Jordanmatrix* genannt.

Für einen nilpotenten Endomorphismus $f : V \to V$ setze

$$\mathrm{nil}(f) = \min\{k \geq 1 : f^k = 0\}$$

und für jedes $v \in V \setminus \{0\}$ setze $\mathrm{nil}(v, f) = \min\{k \geq 1 : f^k(v) = 0\}$. Also ist stets $\mathrm{nil}(v, f) \leq \mathrm{nil}(f)$ und es gibt mindestens ein $v \in V \setminus \{0\}$ mit $\mathrm{nil}(v, f) = \mathrm{nil}(f)$.

Lemma 12.3 *Sei f ein nilpotenter Endomorphismus von V und sei $v \in V \setminus \{0\}$ mit $\mathrm{nil}(v, f) = \mathrm{nil}(f)$. Dann gibt es einen f-invarianten Untervektorraum W von V, so dass $V = Z(v, f) \oplus W$. (Es gilt also $f(W) \subset W$, $V = Z(v, f) + W$ und $Z(v, f) \cap W = \{0\}$.)*

Beweis Zunächst eine Feststellung: Sind W, W_1 Untervektorräume eines endlichdimensionlen Vektorraumes mit $W_1 \subset W$, so gibt es einen Untervektorraum W_2, so dass $W = W_1 \oplus W_2$. (Im Allgemeinen ist aber W_2 nicht eindeutig.)

Für jedes $p \geq 1$ sei A_p die Aussage: Für jeden endlichdimensionalen Vektorraum $U \neq \{0\}$ über K, für jeden nilpotenten Endomorphismus h von U mit $\mathrm{nil}(h) = p$ und für jedes $u \in U \setminus \{0\}$ mit $\mathrm{nil}(u, h) = \mathrm{nil}(h)$ gibt es einen Untervektorraum W von U mit $h(W) \subset W$, so dass $U = Z(u, h) \oplus W$. Es wird durch Induktion nach p gezeigt, dass A_p richtig ist für alle $p \geq 1$.

Die Aussage A_1 ist richtig: Ist $\mathrm{nil}(h) = 1$, so ist $h = 0$. In diesem Fall gilt $h(W) \subset W$ für jeden Untervektorraum W von U und für jedes $u \in U \setminus \{0\}$ gibt es einen Untervektorraum W von U mit $V = Z(u, h) \oplus W$.

Sei nun $p \geq 2$ und nehme an, dass die Aussage A_{p-1} richtig ist. Sei $U \neq \{0\}$ ein endlichdimensionaler Vektorraum über K, sei h ein nilpotenter Endomorphismus von U mit $\mathrm{nil}(h) = p$ und sei $u \in U \setminus \{0\}$ mit $\mathrm{nil}(u, h) = \mathrm{nil}(h)$.

Setze $U_1 = \mathrm{Bild}\, h$; nach Lemma 11.3 gilt $h(U_1) \subset U_1$, und sei also $h_1 = h_{|U_1}$ die Einschränkung von h auf U_1. Dann ist der Endomorphismus $h_1 : U_1 \to U_1$ nilpotent mit $\mathrm{nil}(h_1) = p - 1$. (Für jedes $k \geq 1$ ist

$$h_1^{k-1}(U_1) = h^{k-1}(U_1) = h^{k-1}(h(U)) = h^k(U)$$

und insbesondere gilt $h_1^{p-1}(U_1) = h^p(U) = \{0\}$ und $h_1^{p-2}(U_1) = h^{p-1}(U) \neq \{0\}$, d.h., es gilt $h_1^{p-1} = 0$ aber $h_1^{p-2} \neq 0$.) Setze $u_1 = h(u)$; dann ist $u_1 \in U_1$ mit $\mathrm{nil}(u_1, h_1) = p - 1 = \mathrm{nil}(h_1)$, da $h_1^{p-2}(u_1) = h^{p-1}(u) \neq 0$. Da die Aussage A_{p-1} richtig ist, gibt es einen Untervektorraum W_1 von U_1 mit $h_1(W_1) \subset W_1$, so dass $U_1 = Z(u_1, h_1) \oplus W_1$.

Jetzt werden $Z(u_1, h_1)$ und W_1 als Untervektorräume von U angesehen und da $h_1 = h_{|U_1}$, gilt dann $Z(u_1, h_1) = Z(u_1, h)$ und $h(W_1) \subset W_1$. Dies bedeutet, dass W_1 ein Untervektorraum von U ist mit $h(W_1) \subset W_1$ und $U_1 = Z(u_1, h) \oplus W_1$. Man beachte, dass (nach Lemma 5.8 und da $h^p(u) = 0$)

$$\begin{aligned} Z(u_1, h) &= \mathrm{L}(u_1, h(u_1), \ldots, h^{p-2}(u_1)) = \mathrm{L}(h(u), h(h(u)), \ldots, h^{p-2}(h(u))) \\ &= \mathrm{L}(h(u), h^2(u), \ldots, h^{p-1}(u)) = \mathrm{L}(h(u), h^2(u), \ldots, h^{p-1}(u), 0) \\ &= \mathrm{L}(h(u), h^2(u), \ldots, h^{p-1}(u), h^p(u)) = \mathrm{L}(h(u), h(h(u)), \ldots, h(h^{p-1}(u))) \\ &= h(\mathrm{L}(u, h(u), \ldots, h^{p-1}(u))) = h(Z(u, h)) \; . \end{aligned}$$

Es gilt nun auch $Z(u, h) \cap W_1 = \{0\}$. (*Beweis:* Sei $v \in Z(u, h) \cap W_1$; dann ist $h(v) \in h(Z(u, h)) = Z(u_1, h)$ und $h(v) \in h(W_1) \subset W_1$. Damit ist $h(v) = 0$, weil $Z(u_1, h) \cap W_1 = \{0\}$. Da ferner $v \in Z(u, h) = \mathrm{L}(u, h(u), \ldots, h^{p-1}(u))$, gibt es $\lambda_0, \ldots, \lambda_{p-1} \in K$, so dass $v = \lambda_0 u + \cdots + \lambda_{p-1} h^{p-1}(u)$ und damit ist

$$\begin{aligned} 0 = h(v) &= h(\lambda_0 u + \cdots + \lambda_{p-1} h^{p-1}(u)) \\ &= \lambda_0 h(u) + \cdots + \lambda_{p-1} h^p(u) = \lambda_0 h(u) + \cdots + \lambda_{p-2} h^{p-1}(u) \; , \end{aligned}$$

da $h^p(u) = 0$. Nach Lemma 12.2 sind aber $h(u), \ldots, h^{p-1}(u)$ linear unabhängig und daraus ergibt sich, dass $\lambda_j = 0$ für $j = 0, \ldots, p-2$, d.h., $v = \lambda_{p-1} h^{p-1}(u)$. Insbesondere ist $v \in Z(u_1, h) = \mathrm{L}(h(u), \ldots, h^{p-1}(u))$. Also ist $v \in Z(u_1, h) \cap W_1$ und $Z(u_1, h) \cap W_1 = \{0\}$, d.h., $v = 0$.)

Setze $W' = \{v \in U : h(v) \in W_1\}$; dann ist W' ein Untervektorraum von U. (Der Beweis dafür ist eine Übung.) Da $h(W_1) \subset W_1$, ist $W_1 \subset W'$.

Ferner gilt $U = Z(u, h) + W'$. (*Beweis:* Sei $v \in U$; dann ist $h(v) \in U_1$ und da $U_1 = Z(u_1, h) + W_1$, gibt es $v_1 \in Z(u_1, h)$ und $w_1 \in W_1$, so dass $h(v) = v_1 + w_1$. Aber $Z(u_1, h) = h(Z(u, h))$ und damit gibt es $v_2 \in Z(u, h)$ mit $h(v_2) = v_1$. Folglich ist $w_1 = h(v) - h(v_2) = h(v - v_2)$ und also ist $v - v_2 \in W'$. Daher ist $v = v_2 + (v - v_2) \in Z(u, h) + W'$.)

Setze $W_2 = Z(u, h) \cap W'$; dann gilt $W_1 \subset W'$, $W_2 \subset W'$ und $W_1 \cap W_2 = \{0\}$ (da $W_1 \cap W_2 \subset W_1 \cap Z(u, h) = \{0\}$), d.h., $W_1 + W_2 = W_1 \oplus W_2$. Da $W_1 \oplus W_2 \subset W'$, gibt es nun einen Untervektorraum W_3 von W', so dass $W' = W_1 \oplus W_2 \oplus W_3$.

Setze $W = W_1 \oplus W_3$; da $W \subset W'$, ist $f(W) \subset f(W') \subset W_1 \subset W$ und auch

$$Z(u, h) \cap W = Z(u, h) \cap W' \cap W = W_2 \cap W = W_2 \cap (W_1 + W_3) = \{0\} \; .$$

Schließlich gilt $U = Z(u, h) + W$. (*Beweis:* Sei $v \in U$; da $U = Z(u, h) + W'$, gibt es $v' \in Z(u, h)$ und $w' \in W'$ mit $v = v' + w'$, und da $W' = W_1 \oplus W_2 \oplus W_3$, gibt es dann $w_i \in W_i$ für $i = 1, 2, 3$, so dass $w' = w_1 + w_2 + w_3$. Folglich ist $v = (v' + w_2) + (w_1 + w_3)$ und $v' + w_2 \in Z(u, h)$ und $w_1 + w_3 \in W$.)

Dies zeigt, dass die Aussage A_p richtig ist, und daraus folgt durch Induktion nach p, dass A_p richtig ist für alle $p \geq 1$. □

Satz 12.1 *Sei f ein nilpotenter Endomorphismus von V. Dann gibt es Vektoren $v_1, \ldots, v_m \in V \setminus \{0\}$ mit $\mathrm{nil}(f) = \mathrm{nil}(v_1, f) \geq \cdots \geq \mathrm{nil}(v_m, f)$, so dass*

$$V = Z(v_1, f) \oplus \cdots \oplus Z(v_m, f) \; .$$

Beweis Für jedes $n \geq 1$ sei A_n die Aussage: Für jeden Vektorraum U über K mit $\dim U \leq n$ und für jeden nilpotenten Endomorphismus h von U gibt es Vektoren $u_1, \ldots, u_p \in U \setminus \{0\}$ mit $\mathrm{nil}(h) = \mathrm{nil}(u_1, h) \geq \cdots \geq \mathrm{nil}(u_p, h)$, so dass

$$U = Z(u_1, h) \oplus \cdots \oplus Z(u_p, h) \; .$$

Es wird durch Induktion nach n gezeigt, dass A_n richtig ist für alle $n \geq 1$.

Die Aussage A_1 ist richtig: Ist $\dim U = 1$, so ist $h = 0$ der einzige nilpotente Endomorphimus von U und es gilt $U = \mathrm{L}(u) = Z(u, 0)$ für jedes $u \in U \setminus \{0\}$.

Sei nun $n \geq 2$ und nehme an, dass A_{n-1} richtig ist. Sei U ein Vektorraum über K mit $\dim U = n$ und sei h ein nilpotenter Endomorphismus von U. Wähle einen Vektor $u_1 \in U \setminus \{0\}$ mit $\mathrm{nil}(u_1, h) = \mathrm{nil}(h)$. Nach Lemma 12.4 gibt es also einen h-invarianten Untervektorraum W von U, so dass $U = Z(u_1, h) \oplus W$. Da $\dim Z(u_1, h) \geq 1$, ist $\dim W \leq n - 1$. Sei $g = h_{|W}$ die Einschränkung von h auf W; dann ist $g : W \to W$ ein nilpotenter Endomorphismus und $\mathrm{nil}(g) \leq \mathrm{nil}(h)$, weil $g^k = 0$, falls $h^k = 0$. Da A_{n-1} gilt, gibt es Vektoren $u_2, \ldots, u_p \in W \setminus \{0\}$ mit $\mathrm{nil}(g) = \mathrm{nil}(u_2, g) \geq \cdots \geq \mathrm{nil}(u_p, g)$, so dass

$$W = Z(u_2, g) \oplus \cdots \oplus Z(u_p, g) \; .$$

Aber $\mathrm{nil}(u_k, g) = \mathrm{nil}(u_k, h)$ und $Z(u_k, g) = Z(u_k, h)$ für $k = 2, \ldots, p$ und daraus folgt, dass $\mathrm{nil}(h) = \mathrm{nil}(u_1, h) \geq \mathrm{nil}(u_2, h) \geq \cdots \geq \mathrm{nil}(u_p, h)$ und

$$U = Z(u_1, h) \oplus W = Z(u_1, h) \oplus Z(u_2, h) \oplus \cdots \oplus Z(u_p, h) \; .$$

Dies zeigt, dass die Aussage A_n richtig ist, und daraus folgt durch Induktion nach n, dass A_n richtig ist für alle $n \geq 1$. □

Für jedes $k \geq 1$ sei N_k folgendes Element von $\mathrm{M}(k \times k, K)$:

$$N_k = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix} ,$$

also ist N_k die einzige $k \times k$ nilpotente Jordanmatrix, und man beachte, dass $N_1 = 0$ (als Element von $\mathrm{M}(1 \times 1, K)$).

Eine *Partition* von n ist eine Folge $[s_1, \ldots, s_m]$ mit $m \geq 1$, $s_1 \geq \cdots \geq s_m \geq 1$ und $s_1 + \cdots + s_m = n$. (Die Anzahl von Partitionen von n ist natürlich endlich.) Für jede Partition $\pi = [s_1, \ldots, s_m]$ von n wird ein Element $N_\pi \in \mathrm{M}(n \times n, K)$ definiert durch

$$N_\pi = \begin{pmatrix} N_{s_1} & & & \\ & N_{s_2} & & \\ & & \ddots & \\ & & & N_{s_m} \end{pmatrix} .$$

Für jede Partition $\pi = [s_1, \ldots, s_m]$ von n und jedes $k \geq 1$ sei $\#_k(\pi)$ die Anzahl der Elemente in der Menge $\{1 \leq j \leq m : s_j = k\}$.

Lemma 12.4 *Seien π und π' Partitionen von n; dann gilt $\pi = \pi'$ genau, wenn $\#_k(\pi) = \#_k(\pi')$ für alle $k \geq 1$.*

Beweis Übung. □

Lemma 12.5 *Für jede Partition π von n und jedes $k \geq 1$ gilt*

$$\#_k(\pi) = \operatorname{rang} N_\pi^{k-1} - 2 \operatorname{rang} N_\pi^k + \operatorname{rang} N_\pi^{k+1} .$$

Beweis Sei $\pi = [s_1, \ldots, s_m]$; für jedes $q \geq 0$ ist dann

$$N_\pi^q = \begin{pmatrix} N_{s_1}^q & & & \\ & N_{s_2}^q & & \\ & & \ddots & \\ & & & N_{s_m}^q \end{pmatrix}$$

und daraus ergibt sich, dass

$$\operatorname{rang} N_\pi^q = \sum_{j=1}^{m} \operatorname{rang} N_{s_j}^q = \sum_{k=1}^{n} \#_k(\pi) \operatorname{rang} N_k^q .$$

Sei $1 \leq q < k$; dann sind die Einträge in der Matrix N_k^q alle 0 außer denen auf der oberen Nebendiagonalen der Länge $k - q$ und da sind sie alle 1. Folglich ist $\operatorname{rang} N_k^q = k - q$. Damit ist $\operatorname{rang} N_k^q = \max\{k - q, 0\}$ für alle $k \geq 1$, $q \geq 0$, da $N_k^q = 0$, falls $q \geq k$ und $N_k^0 = E_k$. Also gilt

$$\operatorname{rang} N_\pi^q = \sum_{k=1}^{n} \max\{k - q, 0\} \#_k(\pi) = \sum_{k=q+1}^{n} (k - q) \#_k(\pi)$$

für alle $q \geq 0$. Mit anderen Worten:

$$\begin{aligned}
\operatorname{rang} N_\pi^q &= 0 \quad \text{für alle } q \geq n \ , \\
\operatorname{rang} N_\pi^{n-1} &= \#_k(\pi) \ , \\
\operatorname{rang} N_\pi^{n-2} &= \#_{n-1}(\pi) + 2\#_n(\pi) \ , \\
\operatorname{rang} N_\pi^{n-3} &= \#_{n-2}(\pi) + 2\#_{n-1}(\pi) + 3\#_n(\pi) \ , \\
&\vdots \\
\operatorname{rang} N_\pi^1 &= \#_2(\pi) + 2\#_3(\pi) + 3\#_4(\pi) + \cdots + (n-1)\#_n(\pi) \ , \\
\operatorname{rang} N_\pi^0 &= \#_1(\pi) + 2\#_2(\pi) + 3\#_3(\pi) + \cdots + n\#_n(\pi) \quad (= n) \ ,
\end{aligned}$$

und daraus sieht man leicht, dass für jedes $k \geq 1$

$$\#_k(\pi) = \operatorname{rang} N_\pi^{k-1} - 2\operatorname{rang} N_\pi^k + \operatorname{rang} N_\pi^{k+1} \ . \ \square$$

Lemma 12.6 *Seien π und π' Partitionen von n; dann sind die Matrizen N_π und $N_{\pi'}$ ähnlich genau, wenn $\pi = \pi'$.*

Beweis Sind Matrizen A und B ähnlich, so sind auch A^k und B^k ähnlich für jedes $k \geq 0$ und damit ist $\operatorname{rang} A^k = \operatorname{rang} B^k$ für jedes $k \geq 0$. Sind also N_π und $N_{\pi'}$ ähnlich, so folgt aus Lemma 12.5, dass $\#_k(\pi) = \#_k(\pi')$ für alle $k \geq 0$ und daher ist nach Lemma 12.4 $\pi = \pi'$. Die Umkehrung ist trivial richtig. $\square$

Satz 12.2 *Sei $f : V \to V$ ein nilpotenter Endomorphismus. Dann gibt es eine eindeutige Partition π von n und eine Basis $(u_1, \ldots, u_n)$ von V, so dass N_π die Matrix von f bezüglich $(u_1, \ldots, u_n)$ ist. Ferner gilt*

$$\#_k(\pi) = \operatorname{rang} f^{k-1} - 2\operatorname{rang} f^k + \operatorname{rang} f^{k+1}$$

für jedes $k \geq 1$.

Beweis Nach Satz 12.1 gibt es Vektoren $v_1, \ldots, v_m \in V \setminus \{0\}$, so dass

$$V = Z(v_1, f) \oplus \cdots \oplus Z(v_m, f)$$

und mit $\operatorname{nil}(f) = \operatorname{nil}(v_1, f) \geq \cdots \geq \operatorname{nil}(v_m, f)$. Sei $s_k = \operatorname{nil}(v_k)$; nach Lemma 12.2 ist $(f^{s_k-1}(v_k), \ldots, f(v_k), v_k)$ eine Basis von $Z(v_k, f)$ für jedes k und daraus folgt nach Satz 9.7, dass

$$(u_1, \ldots, u_n) = (f^{s_1-1}(v_1), \ldots, f(v_1), v_1, \ldots, f^{s_m-1}(v_m), \ldots, f(v_m), v_m)$$

eine Basis von V ist. Aber per Definition ist N_π die Matrix von f bezüglich dieser Basis, wobei $\pi = [s_1, \ldots, s_m]$.

Für jedes $k \geq 0$ ist nun N_π^k die Matrix von f^k beüglich $(u_1, \ldots, u_n)$ und damit ist nach Satz 8.3 $\operatorname{rang} f^k = \operatorname{rang} N_\pi^k$. Also ist nach Lemma 12.5

$$\#_k(\pi) = \operatorname{rang} N_\pi^{k-1} - 2\operatorname{rang} N_\pi^k + \operatorname{rang} N_\pi^{k+1} = \operatorname{rang} f^{k-1} - 2\operatorname{rang} f^k + \operatorname{rang} f^{k+1}$$

für jedes $k \geq 1$ und insbesondere ist (nach Lemma 12.4) π eindeutig durch f bestimmt. □

Satz 12.3 *Sei $A \in \mathrm{M}(n \times n, K)$ nilpotent; dann gibt es eine eindeutige Partition π von n, so dass A und N_π ähnlich sind. Ferner gilt*

$$\#_k(\pi) = \operatorname{rang} A^{k-1} - 2\operatorname{rang} A^k + \operatorname{rang} A^{k+1}$$

für jedes $k \geq 1$.

Beweis Der Endomorphismus $\varphi_A : K^n \to K^n$ ist nilpotent und folglich gibt es nach Satz 12.2 eine Basis $(u_1, \ldots, u_n)$ von K^n und eine Partition π von n, so dass N_π die Matrix von φ_A bezüglich $(u_1, \ldots, u_n)$ ist. Aber A ist die Matrix von φ_A bezüglich der kanonischen Basis von K^n und daraus ergibt sich nach Satz 8.8, dass A und N_π ähnlich sind. Für jedes $k \geq 0$ sind dann A^k und N_π^k auch ähnlich und damit gilt $\operatorname{rang} A^k = \operatorname{rang} N_\pi^k$. Nach Lemma 12.5 ist also

$$\#_k(\pi) = \operatorname{rang} N_\pi^{k-1} - 2\operatorname{rang} N_\pi^k + \operatorname{rang} N_\pi^{k+1} = \operatorname{rang} A^{k-1} - 2\operatorname{rang} A^k + \operatorname{rang} A^{k+1}$$

für jedes $k \geq 1$ und insbesondere ist π eindeutig durch A bestimmt. □

Satz 12.4 *Nilpotente Matrizen $A, B \in \mathrm{M}(n \times n, K)$ sind ähnlich genau dann, wenn $\operatorname{rang} A^k = \operatorname{rang} B^k$ für jedes $k = 1, \ldots, n-1$.*

Beweis Sind A und B ähnlich, so sind A^k und B^k ähnlich für jedes $k \geq 0$ und damit ist $\operatorname{rang} A^k = \operatorname{rang} B^k$ für alle k. Sei umgekehrt $\operatorname{rang} A^k = \operatorname{rang} B^k$ für jedes $k = 1, \ldots, n-1$; dann gilt $\operatorname{rang} A^k = \operatorname{rang} B^k$ für all $k \geq 0$, da $A^k = 0 = B^k$ für jedes $k \geq n$ und $A^0 = E_n = B^0$. Nach Satz 12.3 gibt es Partitionen π und π' von n, so dass A ähnlich zu N_π und B ähnlich zu $N_{\pi'}$ ist und ferner gilt

$$\begin{aligned}\#_k(\pi) &= \operatorname{rang} A^{k-1} - 2\operatorname{rang} A^k + \operatorname{rang} A^{k+1} \\ &= \operatorname{rang} B^{k-1} - 2\operatorname{rang} B^k + \operatorname{rang} B^{k+1} = \#_k(\pi')\end{aligned}$$

für alle $k \geq 1$. Nach Lemma 12.4 ist also $\pi' = \pi$. Folglich sind A und B beide ähnlich zu N_π und damit sind A und B ähnlich. □

Im folgenden Satz sei $f : V \to V$ ein trigonalisierbarer Endomorphismus und seien $\lambda_1, \ldots, \lambda_m$ die verschiedenen Eigenwerte von f; für $k = 1, \ldots, m$ setze $p_k = \dim \operatorname{Kern} (f - \lambda_k \mathrm{id}_V)^n$.

Satz 12.5 *Für jedes $k = 1, \ldots, m$ gibt es eine Partition π_k von p_k und eine Basis $(u_1, \ldots, u_n)$ von V, so dass*

$$J = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_m \end{pmatrix}$$

die Matrix von f bezüglich $(u_1, \ldots, u_n)$ ist, wobei $J_k = \lambda_k E_{p_k} + N_{\pi_k}$. Ferner ist die Matrix J eindeutig durch f bestimmt: Seien $q_k \geq 1$, $k = 1, \ldots, m$, mit $q_1 + \cdots + q_m = n$ und für jedes k sei $J'_k = \lambda_k E_{q_k} + N_{\tau_k}$, wobei τ_k eine Partition von q_k ist. Ist

$$J' = \begin{pmatrix} J'_1 & & & \\ & J'_2 & & \\ & & \ddots & \\ & & & J'_m \end{pmatrix}$$

die Matrix von f bezüglich irgendeiner Basis von V, so ist $q_k = p_k$ für jedes k und für alle $j \geq 1$, $k = 1, \ldots, m$ gilt

$$\#_j(\tau_k) = \operatorname{rang}(f - \lambda_k \mathrm{id}_V)^{j-1} - 2\operatorname{rang}(f - \lambda_k \mathrm{id}_V)^j + \operatorname{rang}(f - \lambda_k \mathrm{id}_V)^{j+1} .$$

Beweis Für $k = 1, \ldots, m$ setze $U_k = \operatorname{Kern}(f - \lambda_k \mathrm{id}_V)^n$. Nach Satz 11.2 sind $U_1, \ldots, U_m$ f-invariante Untervektorräume von V mit $V = U_1 \oplus \cdots \oplus U_m$ und für jedes k ist der Endomorphismus $f_{|U_k} - \lambda_k \mathrm{id}_{U_k}$ von U_k nilpotent. Setze $f_k = f_{|U_k}$ und $h_k = f_{|U_k} - \lambda_k \mathrm{id}_{U_k}$; also sind f_k und h_k Endomorphismen von U_k mit h_k nilpotent und $f_k = \lambda_k \mathrm{id}_{U_k} + h_k$. Für jedes k gibt es nach Satz 12.2 eine eindeutige Partition π_k von p_k und eine Basis $(u_1^k, \ldots, u_{p_k}^k)$ von U_k, so dass N_{π_k} die Matrix von h_k bezüglich dieser Basis ist. Dann ist $J_k = \lambda_k E_{p_k} + N_{p_k}$ die Matrix von f_k bezüglich $(u_1^k, \ldots, u_{p_k}^k)$. Sei

$$(u_1, \ldots, u_n) = (u_1^1, \ldots, u_{p_1}^1, \ldots, u_1^m, \ldots, u_{p_m}^m) ;$$

nach Satz 9.7 ist $(u_1, \ldots, u_n)$ eine Basis von V und J ist die Matrix von f bezüglich dieser Basis.

Nehme jetzt an, dass J' die Matrix von f bezüglich einer Basis $(v_1, \ldots, v_n)$ ist. Da $q_1 + \cdots + q_m = n$, kann man

$$(v_1, \ldots, v_n) = (v_1^1, \ldots, v_{q_1}^1, \ldots, v_1^m, \ldots, v_{q_m}^m)$$

schreiben. Für jedes k setze $W_k = \mathrm{L}(v_1^k, \ldots, v_{q_k}^k)$; dann sind die Untervektorräume $W_1, \ldots, W_m$ f-invariant und $V = W_1 \oplus \cdots \oplus W_m$. Nun ist N_{τ_k} die Matrix von $f_{|W_k} - \lambda_k \mathrm{id}_{W_k}$ bezüglich der Basis $(v_1^k, \ldots, v_{q_k}^k)$ von W_k und daraus ergibt sich, dass

$f_{|W_k} - \lambda_k \mathrm{id}_{W_k}$ nilpotent ist. Nach Satz 11.2 ist also $W_k = U_k$ und insbesondere ist $q_k = p_k$ für jedes k. Ferner ist nach Satz 12.2

$$\#_j(\tau_k) = \operatorname{rang} h_k^{j-1} - 2\operatorname{rang} h_k^j + \operatorname{rang} h_k^{j+1}$$

für alle $j \geq 1$, wobei $h_k = f_{|U_k} - \lambda_k \mathrm{id}_{U_k}$. Aber für jedes $\ell \geq 0$ ist

$$(f - \lambda_k \mathrm{id}_V)^\ell = (f_1 - \lambda_k \mathrm{id}_{U_1})^\ell \oplus \cdots \oplus (f_m - \lambda_k \mathrm{id}_{U_m})^\ell$$

und $(f_i - \lambda_k \mathrm{id}_{U_i})^\ell$ ist ein Automorphismus von U_i, wenn $i \neq k$. Folglich ist

$$\begin{aligned} \operatorname{rang}(f - \lambda_k \mathrm{id}_V)^\ell &= \operatorname{rang}(f_1 - \lambda_k \mathrm{id}_{U_1})^\ell + \cdots + \operatorname{rang}(f_m - \lambda_k \mathrm{id}_{U_m})^\ell \\ &= p_1 + \cdots + p_{k-1} + \operatorname{rang} h_k^\ell + p_{k+1} + \cdots + p_m \\ &= n - p_k + \operatorname{rang} h_k^\ell \end{aligned}$$

und daher ist

$$\#_j(\tau_k) = \operatorname{rang}(f - \lambda_k \mathrm{id}_V)^{j-1} - 2\operatorname{rang}(f - \lambda_k \mathrm{id}_V)^j + \operatorname{rang}(f - \lambda_k \mathrm{id}_V)^{j+1}$$

für alle $j \geq 1$, $k = 1, \ldots, m$. □

Im folgenden Satz sei $A \in \mathrm{M}(n \times n, K)$ trigonalisierbar und seien $\lambda_1, \ldots, \lambda_m$ die verschiedenen Eigenwerte von A; setze $p_k = n - \operatorname{rang}(A - \lambda_k E_n)^n$ für jedes $k = 1, \ldots, m$.

Satz 12.6 *(1) Es gilt $p_k \geq 1$ für jedes k und $p_1 + \cdots + p_m = n$.*

(2) Für jedes k gibt es eine Partition π_k von p_k, so dass A ähnlich zu der Matrix

$$J = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_m \end{pmatrix}$$

ist, wobei $J_k = \lambda_k E_{p_k} + N_{\pi_k}$.

(3) Die Matrix J ist eindeutig durch A bestimmt: Seien $q_k \geq 1$, $k = 1, \ldots, m$, mit $q_1 + \cdots + q_m = n$ und für jedes k sei $J'_k = \lambda_k E_{q_k} + N_{\tau_k}$, wobei τ_k eine Partition von q_k ist. Ist A ähnlich zu der Matrix

$$J' = \begin{pmatrix} J'_1 & & & \\ & J'_2 & & \\ & & \ddots & \\ & & & J'_m \end{pmatrix},$$

so ist $q_k = p_k$ für jedes k und für alle $j \geq 1$, $k = 1, \ldots, m$ gilt

$$\#_j(\tau_k) = \operatorname{rang}(A - \lambda_k E_n)^{j-1} - 2\operatorname{rang}(A - \lambda_k E_n)^j + \operatorname{rang}(A - \lambda_k E_n)^{j+1}.$$

Beweis Dies ist ein Spezialfall von Satz 12.5: Der Endomorphismus φ_A von K^n ist trigonalisierbar und $\lambda_1, \ldots, \lambda_m$ sind die verschiedenen Eigenwerte von φ_A. Für jedes $k = 1, \ldots, m$ ist

$$p_k = n - \operatorname{rang}(A - \lambda_k E_n)^n = n - \operatorname{rang}(\varphi_A - \lambda_k \mathrm{id}_{K^n})^n = \operatorname{Kern}(\varphi_A - \lambda_k \mathrm{id}_{K^n})^n$$

und ferner ist A ähnlich zu einer Matrix B genau dann, wenn es eine Basis von K^n gibt, so dass B die Matrix von φ_A bezüglich dieser Basis ist. □

Satz 12.7 *Seien $A, B \in \mathrm{M}(n \times n, K)$ trigonalisierbare Matrizen. Dann sind A und B ähnlich genau, wenn sie die gleichen Eigenwerte $\lambda_1, \ldots, \lambda_m$ besitzen und*

$$\operatorname{rang}(A - \lambda_k E_n)^j = \operatorname{rang}(B - \lambda_k E_n)^j$$

für jedes $j = 1, \ldots, n-1$ und jedes $k = 1, \ldots, m$.

Beweis Sind A und B ähnlich, so besitzen sie die gleichen Eigenwerte $\lambda_1, \ldots, \lambda_m$. Ferner sind dann $(A - \lambda_k E_n)^j$ und $(B - \lambda_k E_n)^j$ auch ähnlich und damit ist

$$\operatorname{rang}(A - \lambda_k E_n)^j = \operatorname{rang}(B - \lambda_k E_n)^j$$

für alle $j \geq 0$ und jedes $k = 1, \ldots, m$. Die Umkehrung folgt aus Satz 12.6. □

13 Gruppen

Ein Tripel $(G, \cdot, e)$ bestehend aus einer Menge G, einer Verknüpfung

$$\begin{aligned} \cdot : G \times G &\to G \\ (a, b) &\mapsto ab \end{aligned}$$

und einem Element $e \in G$ heißt *Gruppe*, wenn folgendes gilt:

(G1) *Assoziativität:* $(ab)c = a(bc)$ für alle a, b, $c \in G$.

(G2) *Das Element e ist linksneutral:* Für alle $a \in G$ gilt $ea = a$.

(G3) *Existenz eines Linksinversen:* Zu jedem $a \in G$ gibt es ein Element $b \in G$, so dass $ba = e$.

Lemma 13.1 *Sei $(G, \cdot, e)$ eine Gruppe.*

(1) Das Element e ist auch rechtsneutral: Für alle $a \in G$ gilt $ae = a$.

(2) Ein Linksinverses ist auch ein Rechtsinverses: Es gilt $ab = e$, falls $ba = e$.

Beweis (2) Seien a, $b \in G$ mit $ba = e$. Nach (G3) gibt es also ein Element $c \in G$ mit $cb = e$ und nach (G1) und (G2) gilt dann

$$e = cb = c(eb) = c((ba)b) = c(b(ab)) = (cb)(ab) = e(ab) = ab\ .$$

(1) Sei $a \in G$; nach (G3) gibt es also ein Element $b \in G$ mit $ba = e$. Nach (G1), (G2) und (2) gilt dann

$$ae = a(ba) = (ab)a = ea = a\ . \quad \square$$

Nach Lemma 13.1 können (G2) und (G3) in der Definition einer Gruppe durch (G2′) und (G3′) ersetzt werden, wobei

(G2′) *Das Element e ist neutral:* Für alle $a \in G$ gilt $ea = ae = a$.

(G3′) *Existenz eines Inversen:* Zu jedem $a \in G$ gibt es ein Element $b \in G$, so dass $ba = ab = e$.

Lemma 13.2 *Sei $(G, \cdot, e)$ eine Gruppe.*

(1) Das Element e ist das einzige neutrale Element: Gilt $e'a = ae' = a$ für alle $a \in G$ für ein $e' \in G$, so ist $e' = e$.

(2) Für jedes $a \in G$ ist das Inverse eindeutig: Gilt $ba = ab = e$ und $ca = ac = e$, so ist $b = c$.

Beweis (1) Dies ist klar, da $e' = ee' = e$.

(2) Hier ist $b = be = b(ac) = (ba)c = ec = c$. □

Nach Lemma 13.2 (1) wird e *das neutrale Element* der Gruppe $(G, \cdot, e)$ genannt. Für jedes $a \in G$ wird das Inverse von a meistens mit a^{-1} bezeichnet; a^{-1} ist also das eindeutige Element mit $a^{-1}a = aa^{-1} = e$. Es gilt $(a^{-1})^{-1} = a$ für jedes $a \in G$, da $aa^{-1} = a^{-1}a = e$, und $(ab)^{-1} = b^{-1}a^{-1}$ für alle $a, b \in G$, da $(b^{-1}a^{-1})ab = b^{-1}(a^{-1}a)b = b^{-1}eb = b^{-1}b = e$ und genauso gilt $ab(b^{-1}a^{-1}) = e$.

Sei $(G, \cdot, e)$ eine Gruppe; gilt $ab = ba$ für alle $a, b \in G$, so nennt man die Gruppe *abelsch*. Für eine abelsche Gruppe schreibt man meistens $(G, +, 0)$ statt $(G, \cdot, e)$; das neutrale Element wird also mit 0 bezeichnet und man schreibt $a + b$ statt ab. Ferner wird in diesem Fall das Inverse von a mit $-a$ statt a^{-1} bezeichnet.

Beispiele: (1) $(\mathbb{Z}, +, 0)$ ist eine abelsche Gruppe.

(2) $(\mathbb{R}, +, 0)$ ist eine abelsche Gruppe.

(3) Ist K ein Körper, so ist $(K, +, 0)$ eine abelsche Gruppe.

(4) Ist K ein Körper, so ist $(K^\times, \cdot, 1)$ eine abelsche Gruppe, wobei $K^\times = K \setminus \{0\}$.

(5) Für $n \geq 1$ ist $(\mathbb{Z}_n, +, 0)$ eine abelsche Gruppe, wobei $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ und $+ : \mathbb{Z}_n \times \mathbb{Z}_n \to \mathbb{Z}_n$ durch $\ell + m = (\ell + m) \operatorname{Rest} n$ definiert ist.

(6) Ist V ein Vektorraum über einem Körper K, so ist $(V, +, 0)$ eine abelsche Gruppe.

(7) Sei K ein Körper und sei $\mathrm{GL}(n, K)$ die Menge aller invertierbaren Elemente in $\mathrm{M}(n \times n, K)$. Nach Lemma 6.3 ist dann $(\mathrm{GL}(n, K), \cdot, E_n)$ eine Gruppe, die *general linear group*, wobei hier $\cdot$ Matrizenmultiplikation bezeichnet. Diese Gruppe ist nicht abelsch, wenn $n > 1$.

(8) Sei V ein Vektorraum über einem Körper K und sei $\mathrm{Aut}(V)$ die Menge aller Automorphismen von V. Nach Lemma 5.2 ist dann $(\mathrm{Aut}(V), \circ, \mathrm{id}_V)$ eine Gruppe. Diese Gruppe ist nicht abelsch, wenn $\dim V > 1$.

(9) Sei X eine Menge und sei $\mathrm{Bij}(X)$ die Menge aller Bijektionen $f : X \to X$. Dann ist $(\mathrm{Bij}(X), \circ, \mathrm{id}_X)$ eine Gruppe, die *Gruppe der Permutationen* von X oder die *symmetrische Gruppe* von X. Enthält X mehr als zwei Elemente, so ist die symmetrische Gruppe von X nicht abelsch. Im Spezialfall $X = \{1, 2, \ldots, n\}$ schreibt man S_n statt $\mathrm{Bij}(X)$. Die Menge S_n enthält $n!$ Elemente.

(10) Sei $G = \{e, a, b, c\}$ und sei $\cdot : G \times G \to G$ gegeben durch

$\cdot$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

Dann ist $(G, \cdot, e)$ eine abelsche Gruppe; sie heißt die *Kleinsche Vierergruppe*.

Bemerkung: Ist $(G, \cdot, e)$ eine Gruppe mit G endlich, so kann man die Verknüpfung $\cdot : G \times G \to G$ oft am Besten in einer *Gruppentafel* (wie in (10)) übersehen.

Ist $(G, \cdot, e)$ eine Gruppe, dann schreibt man meistens einfach G statt $(G, \cdot, e)$ (und geht davon aus, dass es klar ist, welche Verknüpfung $\cdot$ und welches neutrale Element e gemeint ist).

Sei G eine Gruppe und seien $a_1, \ldots, a_n \in G$ mit $n \geq 3$; dann ist das 'Produkt' von $a_1, \ldots, a_n$ unabhängig von der Reihenfolge der einzelnen Multiplikationen; dieses 'Produkt' wird mit $a_1 \cdots a_n$ bezeichnet.

Sei G eine Gruppe und sei $a \in G$; die Potenzen a^n, $n \geq 0$, werden rekursiv durch $a^0 = e$, $a^1 = a$ und (für $n \geq 1$) $a^{n+1} = aa^n$ definirt. (Ist G abelsch, so soll man aber na statt a^n schreiben.) Es ist leicht zu sehen, dass $a^{m+n} = a^m a^n$ für alle $m, n \geq 0$. Nehme an, dass es $0 \leq m < n$ mit $a^m = a^n$ gibt; dann gilt

$$\begin{aligned} e = (a^m)^{-1}a^m &= (a^m)^{-1}a^n \\ &= (a^m)^{-1}(a^m a^{n-m}) = ((a^m)^{-1}a^m)a^{n-m} = ea^{n-m} = a^{n-m} \,. \end{aligned}$$

Dies zeigt also, dass entweder

— alle Potenzen von a verschieden sind, und dann nennt man a ein Element *unendlicher Ordnung*, oder

— $a^m = e$ für ein $m \geq 1$, und dann nennt man a ein Element *endlicher Ordnung*.

Sei $a \in G$ ein Element endlicher Ordnung; die Zahl $\min\{m \geq 1 : a^m = e\}$ heißt dann die *Ordnung* von a und wird mit $\operatorname{ord} a$ bezeichnet. Ist $a \in G$ ein Element unendlicher Ordnung, so schreibt man auch $\operatorname{ord} a = \infty$.

Eine Gruppe G heißt *endlich*, wenn die Menge G endlich ist; in diesem Fall heißt die Anzahl der Elemente in G die *Ordnung* von G und wird mit $\operatorname{ord} G$ bezeichnet. Ist G eine endliche Gruppe, dann ist jedes Element von G ein Element endlicher Ordnung.

Sei G eine Gruppe; eine nichtleere Teilmenge H von G heißt *Untergruppe* von G, wenn H mit der Verknüpfung $\cdot$ und mit dem neutralen Element e aus G selbst eine Gruppe ist. D.h., $H \subset G$ ist eine Untergruppe von G, wenn gilt:

— $ab \in H$ für alle $a, b \in H$,

— $e \in H$,

— $a^{-1} \in H$ für jedes $a \in H$.

Beispiele: (1) Sei G eine Gruppe; dann sind $\{e\}$ und G stets Untergruppen von G.

(2) $\mathbb{Z}$ ist eine Untergruppe von $\mathbb{R}$.

(3) Ist V ein Vektorraum über einem Körper K, so ist jeder Untervektorraum von V auch eine Untergruppe von V.

(4) Sei V ein K-Vektorraum; dann ist $\mathrm{Aut}(V)$ eine Untergruppe von $\mathrm{Bij}(V)$.

(5) Sei Y Teilmenge einer Menge X und H eine Untergruppe von $\mathrm{Bij}(X)$. Setze

$$\mathrm{Fix}_H(Y) = \{f \in H : f(y) = y \text{ für alle } y \in Y\} ,$$
$$\mathrm{Inv}_H(Y) = \{f \in H : f(Y) = Y\} .$$

Dann sind $\mathrm{Fix}_H(Y)$ und $\mathrm{Inv}_H(Y)$ beide Untergruppen von $\mathrm{Bij}(X)$.

Ist H eine endliche Untergruppe von G, so heißt die Anzahl der Elemente in H die *Ordnung* von H und wird mit $\operatorname{ord} H$ bezeichnet.

Im Folgenden sei G eine Gruppe.

Lemma 13.3 *Sei H eine nichtleere Teilmenge von G. Gilt $a^{-1}b \in H$ für alle $a, b \in H$, so ist H eine Untergruppe von G.*

Beweis Da $H \neq \varnothing$, gibt es ein Element $c \in H$, und folglich ist $e = c^{-1}c \in H$. Seien nun $a, b \in H$; dann gilt $a^{-1} = a^{-1}e \in H$ und $ab = (a^{-1})^{-1}b \in H$. Damit ist H eine Untergruppe von G. □

Lemma 13.4 *Sei H eine endliche Teilmenge von G, für die gilt:*

— *$e \in H$,*

— *$ab \in H$ für alle $a, b \in H$.*

Dann ist H eine Untergruppe von G.

Beweis Es muss gezeigt werden, dass $a^{-1} \in H$ für jedes $a \in H$. Sei $a \in H$; dann liegen alle Potenzen a^n, $n \geq 0$, in H, da $bc \in H$ für alle $b, c \in H$. Aber H ist endlich und folglich können diese Potenzen nicht alle verschieden sein; also ist a ein Element endlicher Ordnung. Sei $n = \operatorname{ord} a$; dann ist $a^{n-1}a = a^n = e$, d.h., $a^{n-1} = a^{-1}$ und damit ist $a^{-1} \in H$. □

Sei S eine nichtleere Teilmenge von G und bezeichne mit $\langle S \rangle$ die Teilmenge von G, die aus allen endlichen Produkten von Elementen aus der Menge $S \cup S^{-1}$ besteht, wobei $S^{-1} = \{a^{-1} : a \in S\}$.

Lemma 13.5 *Sei S eine nichtleere Teilmenge von G. Dann ist $\langle S \rangle$ eine Untergruppe von G. Ferner ist $\langle S \rangle$ die kleinste Untergruppe von G, die S enthält: Ist H eine Untergruppe von G mit $S \subset H$, so ist $\langle S \rangle \subset H$.*

Beweis Es gilt $e \in \langle S\rangle$, da $e = cc^{-1}$ für jedes $c \in S$. Seien a, $b \in \langle S\rangle$; dann gibt es $a_1, \ldots, a_m, b_1, \ldots, b_n \in S \cup S^{-1}$, so dass $a = a_1 \cdots a_m$ und $b = b_1 \cdots b_n$ und folglich ist $ab = a_1 \cdots a_m b_1 \cdots b_n \in \langle S\rangle$. Ferner ist $a^{-1} = a_m^{-1} \cdots a_1^{-1} \in \langle S\rangle$. Damit ist $\langle S\rangle$ eine Untergruppe von G. Sei nun H eine Untergruppe von G mit $S \subset H$; dann ist $a_1 \cdots a_n \in H$ für alle $a_1, \ldots, a_n \in S \cup S^{-1}$ und daraus ergibt sich, dass $\langle S\rangle \subset H$. □

Man nennt $\langle S\rangle$ die *von S erzeugte Untergruppe von G*. Ist $S = \{a_1, \ldots, a_n\}$ endlich, so schreibt man einfach $\langle a_1, \ldots, a_n\rangle$ statt $\langle\{a_1, \ldots, a_n\}\rangle$.

Für $a \in G$, $m \geq 1$ schreibt man meistens a^{-m} statt $(a^{-1})^m$. Für alle m, $n \in \mathbb{Z}$ gilt nun $a^{m+n} = a^m a^n$.

Lemma 13.6 *(1) Ist $a \in G$ ein Element unendlicher Ordnung, dann sind die Potenzen a^m, $m \in \mathbb{Z}$, alle verschieden und $\langle a\rangle = \{\ldots, a^{-2}, a^{-1}, e, a, a^2, \ldots\}$.*

(2) Ist $a \in G$ ein Element endlicher Ordnung mit $n = \operatorname{ord} a$, dann sind die Potenzen a^m, $m = 0, \ldots, n-1$, alle verschieden und $\langle a\rangle = \{e, a, \ldots, a^{n-1}\}$.

Beweis (1) Nehme an, dass es m, $n \in \mathbb{Z}$ mit $m < n$ und $a^m = a^n$ gibt. Dann gilt $a^{m+\ell} = a^m a^\ell = a^n a^\ell = a^{n+\ell}$ für alle $\ell \geq 0$ und folglich gibt es p, q mit $0 \leq p < q$, so dass $a^p = a^q$. Da aber a ein Element unendlicher Ordnung ist, ist dies nicht möglich, und damit müssen die Potenzen a^m, $m \in \mathbb{Z}$, alle verschieden sein. Es folgt nun umittelbar aus der Definition von $\langle a\rangle$, dass $\langle a\rangle = \{\ldots, a^{-2}, a^{-1}, e, a, a^2, \ldots\}$.

(2) Nehme an, dass es $0 \leq k < \ell < n$ mit $a^k = a^\ell$ gibt. Dann ist

$$a^{\ell-k} = (a^k)^{-1} a^k a^{\ell-k} = (a^k)^{-1} a^\ell = (a^k)^{-1} a^k = e$$

und dies ist nicht möglich, da $1 \leq \ell - k < n$ und $n = \min\{m \geq 1 : a^m = e\}$. Die Potenzen a^m, $m = 0, \ldots, n-1$ sind also alle verschieden. Nach Lemma 13.4 ist $H = \{e, a, \ldots, a^{n-1}\}$ eine Untergruppe von G, da $a^k a^\ell = a^p$ mit $p = (k+\ell) \operatorname{Rest} n$ für alle k, $\ell \in \{0, 1, \ldots, n-1\}$, und daraus folgt nach Lemma 13.5, dass $\langle a\rangle \subset H$. Damit ist $H = \langle a\rangle$. □

Eine Untergruppe H von G heißt *zyklisch*, wenn $H = \langle a\rangle$ für ein $a \in H$. Ist H eine unendliche zyklische Untergruppe von G, dann ist nach Lemma 13.6 (1)

$$H = \{\ldots, a^{-2}, a^{-1}, e, a, a^2, \ldots\}$$

für ein $a \in H$. Ist dagegen H eine endliche zyklische Untergruppe von G, so ist nach Lemma 13.6 (2) $\langle a\rangle = \{e, a, \ldots, a^{n-1}\}$ für ein $a \in H$ mit $n = \operatorname{ord} a = \operatorname{ord} H$.

Die Gruppe G heißt *zyklisch*, wenn $G = \langle a\rangle$ für ein $a \in G$, d.h., wenn G zyklisch ist als Untergruppe von sich selbst. Jede zyklische Gruppe ist abelsch, da

$$a^m a^n = a^{m+n} = a^{n+m} = a^n a^m$$

für alle m, $n \in \mathbb{Z}$. Für jedes $n \geq 1$ ist $\mathbb{Z}_n$ eine zyklische Gruppe der Ordnung n; $\mathbb{Z}$ ist eine unendliche zyklische Gruppe.

Satz 13.1 *Ist G zyklisch, so ist jede Untergruppe H von G auch zyklisch.*

Beweis Man kann annehmen, dass $H \neq \{e\}$, da die Untergruppe $\{e\} = \langle e \rangle$ trivial zyklisch ist. Nun ist $G = \langle a \rangle$ für ein $a \in G$, und folglich gibt es $n \in \mathbb{Z} \setminus \{0\}$, so dass $a^n \in H$. Damit gibt es $k \geq 1$, so dass $a^k \in H$, da $a^{-n} \in H$, falls $a^n \in H$. Setze $m = \min\{k \geq 1 : a^k \in H\}$; es gilt also $\langle a^m \rangle \subset H$. Nehme an, dass $\langle a^m \rangle \neq H$, und sei $a^n \in H \setminus \langle a^m \rangle$. Es gibt dann $q \in \mathbb{Z}$ und $0 \leq \ell < m$, so dass $n = mq + \ell$, und tatsächlich ist $\ell \geq 1$, sonst wäre $a^n = a^{mq} = (a^m)^q \in \langle a^m \rangle$. Aber

$$a^\ell = (a^{mq})^{-1} a^{mq} a^\ell = (a^{mq})^{-1} a^{mq+\ell} = (a^{mq})^{-1} a^n = ((a^m)^q)^{-1} a^n \in H \ ,$$

im Widerspruch zur Wahl von m. Daraus ergibt sich, dass $H = \langle a^m \rangle$, d.h., H ist zyklisch. □

Sei H eine Untergruppe von G und für $a, b \in G$ schreibe $a \equiv b \bmod H$, wenn $ab^{-1} \in H$. Auf diese Weise wird eine Relation $\equiv \bmod H$ auf G definiert.

Lemma 13.7 *Die Relation $\equiv \bmod H$ ist eine Äquivalenzrelation auf G.*

Beweis Reflexivität: Es gilt $aa^{-1} = e \in H$ und damit ist $a \equiv a \bmod H$ für jedes $a \in G$.

Symmetrie: Seien $a, b \in G$ mit $a \equiv b \bmod H$. Dann ist $ab^{-1} \in H$ und damit auch $ba^{-1} = (ab^{-1})^{-1} \in H$, d.h., $b \equiv a \bmod H$.

Transitivität: Seien $a, b, c \in G$ mit $a \equiv b \bmod H$ und $b \equiv c \bmod H$. Dann ist $ab^{-1} \in H$ und $bc^{-1} \in H$ und damit auch $ac^{-1} = a(b^{-1}b)c^{-1} = (ab^{-1})(bc^{-1}) \in H$, d.h., $a \equiv c \bmod H$. □

Die folgende allgemeine Konstruktion für Äquivalenzrelationen wird benötigt: Sei X eine Menge und sei $\equiv$ eine Äquivalenzrelation auf X. Eine Teilmenge A von X heißt dann *Äquivalenzklasse (bezüglich $\equiv$)*, wenn es ein Element $x \in X$ gibt, so dass $A = \{x' \in X : x' \equiv x\}$.

Lemma 13.8 *Zu jedem $x \in X$ gibt es eine eindeutige Äquivalenzklasse, die x enthält.*

Beweis Für jedes $x \in X$ sei $A_x = \{x' \in X : x' \equiv x\}$. Dann gilt $x \in A_x$, da $x \equiv x$, und folglich gibt es mindestens eine Äquivalenzklasse (nämlich A_x), die x enthält. Sei nun A eine beliebige Äquivalenzklasse, die x enthält. Dann gilt $A = A_x$: Es gibt $y \in X$ mit $A = A_y$; dann ist $x \equiv y$ und damit auch $y \equiv x$, da $\equiv$ symmetrisch ist. Sei $z \in A$; dann ist $z \equiv y$ und also $z \equiv x$, da $y \equiv x$ und $\equiv$ transitiv ist. Damit ist $z \in A_x$ und daraus ergibt sich, dass $A \subset A_x$. Sei umgekehrt $z' \in A_x$; dann ist

$z' \equiv x$ und also $z' \equiv y$, da $x \equiv y$ und $\equiv$ transitiv ist. Daher ist $z' \in A_y = A$ und folglich ist $A_x \subset A$. d.h., $A = A_x$. □

Seien A, A' Äquivalenzklassen; dann nach Lemma 13.8 ist entweder $A = A'$ oder $A \cap A' = \varnothing$.

Betrachte nun wieder die Äquivalenzrelation $\equiv \mod H$. Für jedes $a \in G$ setze $Ha = \{ba : b \in H\}$; insbesondere ist $He = H$. Die Mengen Ha, $a \in G$, heißen die *Rechtsnebenklassen* von H.

Lemma 13.9 *Sei $a \in G$; dann ist*

$$Ha = \{b \in G : b \equiv a \bmod H\}$$

und die Abbildung $b \mapsto ba$ bildet die Menge H auf die Menge Ha bijektiv ab.

Beweis Sei $b \in G$; dann gilt

$$\begin{aligned} b \in Ha \;\Leftrightarrow\; b = ca \text{ für ein } c \in H \;&\Leftrightarrow\; ba^{-1} = c \text{ für ein } c \in H \\ &\Leftrightarrow\; ba^{-1} \in H \;\Leftrightarrow\; b \equiv a \bmod H \end{aligned}$$

und folglich gilt $Ha = \{b \in G : b \equiv a \bmod H\}$. Es ist klar, dass die Abbildung $b \mapsto ba$ die Menge H auf Ha surjektiv abbildet. Ferner ist diese Abbildung injektiv, da $b = baa^{-1} = caa^{-1} = c$, falls $ba = ca$. □

Seien a, $b \in G$; dann ist nach Lemma 13.8 und Lemma 13.9 entweder $Ha = Hb$ oder $Ha \cap Hb = \varnothing$. Ferner gilt $Ha = Hb$ genau dann, wenn $a \equiv b \bmod H$.

Satz 13.2 (Satz von Lagrange) *Sei G eine endliche Gruppe. Dann teilt die Ordnung jeder Untergruppe von G die Ordnung der Gruppe G.*

Beweis Sei H eine Untergruppe und seien $A_1, \ldots, A_m$ die Äquivalenzklassen bezüglich der Äquivalenzrelation $\equiv \mod H$. (Da die Menge G endlich ist, gibt es nur endlich viele Äquivalenzklassen.) Für jedes $j = 1, \ldots, m$ gibt es nach Lemma 13.9 ein $a_j \in G$, so dass $A_j = Ha_j$ und daraus folgt nach der zweiten Aussage in Lemma 13.9, dass A_j genau $\operatorname{ord} H$ Elemente enthält. Damit gilt nach Lemma 13.8, dass $\operatorname{ord} G = m \times \operatorname{ord} H$. Insbesondere ist $\operatorname{ord} H$ ein Teiler von $\operatorname{ord} G$. □

Sei H eine Untergruppe von G; die Anzahl der verschiedenen Rechtsnebenklassen von H heißt der *Index* von H in G und wird mit $[G : H]$ bezeichnet. Ist G endlich, dann folgt aus dem Beweis für Satz 13.2, dass $\operatorname{ord} G = [G : H] \times \operatorname{ord} H$.

Satz 13.3 *Sei G eine endliche Gruppe der Ordnung n. Für jedes $a \in G$ ist dann* $\operatorname{ord} a$ *ein Teiler von n und insbesondere ist $a^n = e$.*

Beweis Sei $a \in G$ und setze $H = \langle a \rangle$; nach Lemma 13.4 (2) ist H eine Untergruppe von G mit $\operatorname{ord} H = \operatorname{ord} a$ und nach Satz 13.2 ist $\operatorname{ord} H$ ein Teiler von $n = \operatorname{ord} G$, d.h., $\operatorname{ord} a$ ist ein Teiler von n. Sei nun $m = \operatorname{ord} H$, es gibt also $\ell \geq 1$ mit $\ell m = n$, und damit ist $a^n = (a^m)^\ell = e^\ell = e$. □

Satz 13.4 *Sei G eine endliche Gruppe der Ordnung p, wobei p Primzahl ist. Dann ist G zyklisch.*

Beweis Sei $a \in G$ mit $a \neq e$; dann ist $\operatorname{ord} a \neq 1$ und nach Satz 13.3 ist $\operatorname{ord} a$ ein Teiler von p. Die einzige Möglichkeit ist also $\operatorname{ord} a = p$ und folglich ist $\langle a \rangle$ eine Untergruppe von G der Ordnung p. Damit ist $G = \langle a \rangle$, d.h., G ist zyklisch. □

Für $n > 1$ sei $\varphi(n)$ die Anzahl der zu n teilerfremden Zahlen m mit $1 \leq m < n$.

Satz 13.5 (Euler) *Sei $n > 1$, $m \in \mathbb{Z}$ teilerfremd; dann ist $m^{\varphi(n)} = 1 \bmod n$.*

Beweis Sei $\mathbb{Z}_n^* = \{m \in \mathbb{Z}_n : m \text{ und } n \text{ sind teilerfremd}\}$; sind k, $m \in \mathbb{Z}_n^*$, so ist $km \operatorname{Rest} n \in \mathbb{Z}_n^*$, und folglich kann eine Verknüpfung $\cdot : \mathbb{Z}_n^* \times \mathbb{Z}_n^* \to \mathbb{Z}_n^*$ durch $km = km \operatorname{Rest} n$ definiert werden. Dann ist $(\mathbb{Z}_n^*, \cdot, 1)$ eine Gruppe. (Der Beweis dafür ist eine Übung.) Da diese Gruppe der Ordnung $\varphi(n)$ ist, gilt nach Satz 13.3, dass $m^{\varphi(n)} = 1$ (in $\mathbb{Z}_n^*$) für jedes $m \in \mathbb{Z}_n^*$, d.h., $m^{\varphi(n)} = 1 \bmod n$ (in $\mathbb{Z}$) für jedes $m \in \mathbb{Z}_n^*$. Daraus ergibt sich, dass $m^{\varphi(n)} = 1 \bmod n$ für alle $m \in \mathbb{Z}$, die teilerfremd zu n sind. □

Sei p eine Primzahl; dann ist $\varphi(p) = p - 1$, und daraus folgt nach Satz 13.5, dass $m^{p-1} = 1 \bmod p$ und damit auch $m^p = m \bmod p$ für alle $m \in \mathbb{Z}$, die teilerfremd zu p sind. Ist aber $m \in \mathbb{Z}$ nicht teilerfremd zu p, so ist m ein Vielfaches von p und in diesem Fall ist $m^p = m = 0 \bmod p$. Damit wurde den folgenden Satz von Fermat bewiesen: Ist p eine Primzahl, so gilt $m^p = m \bmod p$ für alle $m \in \mathbb{Z}$.

Sei H eine Untergruppe einer Gruppe G; setze $aH = \{ab : b \in H\}$ für jedes $a \in G$. Die Mengen aH, $a \in G$, heißen die *Linksnebenklassen* von H. Es gilt $aH = Ha$ für jedes $a \in G$, wenn G abelsch ist, aber im Allgemeinen ist dies nicht richtig. (Übung: Man finde ein $b \in S_3$, so dass $bH \neq Hb$, wobei $H = \{e, a\}$ und $a : \{1, 2, 3\} \to \{1, 2, 3\}$ durch $a(1) = 2$, $a(2) = 1$ und $a(3) = 3$ gegeben ist.)

Für jedes $a \in G$ setze $aHa^{-1} = \{aba^{-1} : b \in H\}$ und für Teilmengen S und T Teilmengen von G setze auch $ST = \{ab : a \in S, b \in T\}$.

Lemma 13.10 *Sei H eine Untergruppe von G; dann sind äquivalent:*

(1) $aba^{-1} \in H$ *für alle* $a \in G$, $b \in H$, *(d.h.,* $aHa^{-1} \subset H$ *für jedes* $a \in G$*).*

(2) $aHa^{-1} = H$ *für jedes* $a \in G$*.*

(3) $aH = Ha$ *für jedes* $a \in G$*.*

(4) $H(ab) = (Ha)(Hb)$ *für alle* a, $b \in G$*.*

Beweis (1) $\Rightarrow$ (2): Da $aba^{-1} \in H$ für alle $a \in G$, $b \in H$ und $(a^{-1})^{-1} = a$, ist auch $a^{-1}ba \in H$ für alle $a \in G$, $b \in H$. Sei $b \in H$ und $a \in G$; dann ist $b' = a^{-1}ba \in H$ und $ab'a^{-1} = b$. Daraus folgt, dass $H \subset aHa^{-1}$, d.h. $H = aHa^{-1}$ für jedes $a \in G$.

(2) $\Rightarrow$ (3): Seien $a \in G$, $b \in H$ und setze $b_1 = aba^{-1}$. Dann ist $b_1 \in H$ und damit $ab = b_1a \in Ha$; d.h., $aH \subset Ha$. Ferner gibt es $b_2 \in H$, so dass $b = ab_2a^{-1}$ und also ist $ba = ab_2 \in aH$; d.h., $Ha \subset aH$.

(3) $\Rightarrow$ (4): Seien a, $b \in G$; für jedes $c \in G$ ist $cab = (ca)(eb) \in (Ha)(Hb)$ und damit gilt $H(ab) \subset (Ha)(Hb)$ immer. Sei nun $c \in (Ha)(Hb)$, es gibt also c_1, $c_2 \in H$ mit $c = c_1ac_2b$. Aber $aH = Ha$ und folglich gibt es $c_3 \in H$, so dass $ac_2 = c_3a$. Damit ist $c = c_1c_3ab \in H(ab)$; d.h., $(Ha)(Hb) \subset H(ab)$.

(4) $\Rightarrow$ (1): Seien $a \in G$, $b \in H$; dann ist

$$aba^{-1} = eaba^{-1} \in (Ha)(Ha^{-1}) = H(aa^{-1}) = He = H \ . \ \square$$

Eine Untergruppe H von G heißt *Normalteiler* von G, wenn $aba^{-1} \in H$ für alle $a \in G$, $b \in H$.

Beispiele: (1) $\{e\}$ und G sind stets Normalteiler von G.

(2) Jede Untergruppe einer abelschen Gruppe ist ein Normalteiler.

(3) Jede Untergruppe H einer endlichen Gruppe G mit Index 2 (d.h., $[G : H] = 2$) ist ein Normalteiler von G. (Für jedes $a \in H$ gilt $Ha = H = aH$ und für jedes $a \notin H$ gilt $Ha = G \setminus H = aH$.)

Lemma 13.11 *Sei S eine Teilmenge von G mit $aba^{-1} \in S$ für alle $a \in G$, $b \in S$. Dann ist $\langle S \rangle$ ein Normalteiler von G.*

Beweis Da $ab^{-1}a^{-1} = (aba^{-1})^{-1}$, ist $aba^{-1} \in S \cup S^{-1}$ für alle $a \in G$, $b \in S \cup S^{-1}$. Seien nun $a \in G$, $b \in \langle S \rangle$; dann gibt es $b_1, \ldots, b_n \in S \cup S^{-1}$, so dass $b = b_1 \cdots b_n$, und folglich ist

$$aba^{-1} = ab_1 \cdots b_n a^{-1} = (ab_1a^{-1}) \cdots (ab_na^{-1}) \in \langle S \rangle \ .$$

Damit ist $\langle S \rangle$ ein Normalteiler. $\square$

Sei $S = \{aba^{-1}b^{-1} : a, b \in G\}$. Dann ist

$$c(aba^{-1}b^{-1})c^{-1} = (cac^{-1})(cbc^{-1})(cac^{-1})^{-1}(cbc^{-1})^{-1}$$

ein Element von S für alle a, b, $c \in G$, und daraus folgt nach Lemma 13.11, dass $\langle S \rangle$ ein Normalteiler von G ist. Dieser Normalteiler wird mit $K(G)$ bezeichnet und heißt die *Kommutatorgruppe* von G. Es gilt $K(G) = \{e\}$ genau dann, wenn G abelsch ist.

Sei H ein Normalteiler von G; mit G/H wird die Menge aller Äquivalenzklassen bezüglich der Äquivalenzrelation $\equiv \mod H$ bezeichnet. Nach Lemma 13.9 ist also G/H die Menge der verschiedenen Rechtsnebenklassen von H. Nach Lemma 13.10 ist $XY \in G/H$ für alle X, $Y \in G/H$. (Es gibt a, $b \in G$ mit $X = Ha$ und $Y = Hb$ und dann ist $XY = (Ha)(Hb) = H(ab) \in G/H$.) Damit gibt es eine Verknüpfung

$$\begin{aligned} \cdot : G/H \times G/H &\to G/H \\ (X, Y) &\mapsto XY \ . \end{aligned}$$

Lemma 13.12 *Sei H ein Normalteiler von G; dann ist $(G/H, \cdot, H)$ eine Gruppe, die Faktorgruppe von G nach H.*

Beweis Assoziativität: Sind X, Y, Z beliebige Teilmengen von G, dann gilt

$$(XY)Z = X(YZ) = \{abc : a \in X,\, b \in Y,\, c \in Z\} \ .$$

Das Element H ist linksneutral: Sei $X \in G/H$; dann gibt es $a \in G$ mit $X = Ha$ und nach Lemma 13.10 ist $HX = (He)(Ha) = H(ea) = Ha = X$.

Existenz eines Linksinversen: Sei $X \in G/H$, es gibt also $a \in G$ mit $X = Ha$. Setze $Y = Ha^{-1}$; dann ist $Y \in G/H$ und nach Lemma 13.10 ist

$$YX = (Ha^{-1})(Ha) = H(a^{-1}a) = He = H \ . \ \square$$

Ist H Normalteiler einer endlichen Gruppe G, so ist $\operatorname{ord} G/H = [G : H]$.

Seien G, H Gruppen; eine Abbildung $\varphi : G \to H$ heißt ein *Homomorphismus* (oder ein *Gruppen-Homomorphismus*), wenn gilt:

$$\varphi(ab) = \varphi(a)\varphi(b)$$

für alle a, $b \in G$.

Lemma 13.13 *Sei $\varphi : G \to H$ ein Homomorphismus. Dann gilt $\varphi(e) = e$ und $\varphi(a^{-1}) = \varphi(a)^{-1}$ für jedes $a \in G$.*

Beweis Es gilt $\varphi(e) = \varphi(ee) = \varphi(e)\varphi(e)$ und daraus folgt, dass

$$\varphi(e) = e\varphi(e) = (\varphi(e)^{-1}\varphi(e))\varphi(e) = \varphi(e)^{-1}(\varphi(e)\varphi(e)) = \varphi(e)^{-1}\varphi(e) = e\ .$$

Ferner gilt $\varphi(a^{-1})\varphi(a) = \varphi(a^{-1}a) = \varphi(e) = e$ und damit ist $\varphi(a^{-1}) = \varphi(a)^{-1}$. □

Es ist klar, dass für jede Gruppe G die Identitätsabbildung $\mathrm{id}_G : G \to G$ ein Homomorphismus ist.

Lemma 13.14 *Seien G, H, K Gruppen und seien $\varphi : G \to H$, $\psi : H \to K$ Homomorphismen. Dann ist $\psi \circ \varphi : G \to K$ ein Homomorphismus.*

Beweis Für alle $a, b \in G$ ist

$$(\psi \circ \varphi)(ab) = \psi(\varphi(ab)) = \psi(\varphi(a)\varphi(b)) = \psi(\varphi(a))\psi(\varphi(b)) = (\psi \circ \varphi)(a)(\psi \circ \varphi)(b)$$

und damit ist $\psi \circ \varphi$ ein Homomorphismus. □

Seien G, H Gruppen; ein Homomorphismus $\varphi : G \to H$ heißt *Isomorphismus*, wenn es einen Homomorphismus $\psi : H \to G$ gibt, so dass $\psi \circ \varphi = \mathrm{id}_G$ und $\varphi \circ \psi = \mathrm{id}_H$. In diesem Fall ist nach Lemma 5.1 (2) φ eine bijektive Abbildung und ψ ist die Umkehrabbildung φ^{-1} von φ. Die Umkehrung ist auch richtig:

Lemma 13.15 *Sei $\varphi : G \to H$ ein Homomorphismus. Ist φ bijektiv, dann ist die Umkehrabbildung $\varphi^{-1} : H \to G$ ein Homomorphismus und damit is φ ein Isomorphismus.*

Beweis Seien $a, b \in H$; dann ist $\varphi(\varphi^{-1}(a)\varphi^{-1}(b)) = \varphi(\varphi^{-1}(a))\varphi(\varphi^{-1}(b)) = ab$ und daraus folgt, dass $\varphi^{-1}(ab) = \varphi^{-1}(a)\varphi^{-1}(b)$. □

Sind $\varphi : G \to H$ und $\psi : H \to K$ Isomorphismen, so sind auch $\psi \circ \varphi : G \to K$ und $\varphi^{-1} : H \to G$ Isomorphismen. Ferner ist $\mathrm{id}_G : G \to G$ ein Isomorphismus.

Für einen Homomorphismus $\varphi : G \to H$ setze

$$\operatorname{Kern}\varphi = \{a \in G : \varphi(a) = e\} \quad \text{und} \quad \operatorname{Bild}\varphi = \varphi(G)\ .$$

Lemma 13.16 *(1)* $\operatorname{Kern}\varphi$ *ist ein Normalteiler von G.*

(2) $\operatorname{Bild}\varphi$ *ist eine Untergruppe von H.*

Beweis (1) Da $e \in \operatorname{Kern}\varphi$, ist $\operatorname{Kern}\varphi \neq \varnothing$. Seien a, $b \in \operatorname{Kern}\varphi$; dann ist nach Lemma 13.13 $\varphi(a^{-1}b) = \varphi(a^{-1})\varphi(b) = \varphi(a)^{-1}\varphi(b) = e^{-1}e = e$ und damit ist $a^{-1}b \in \operatorname{Kern}\varphi$. Daraus folgt nach Lemma 13.3, dass $\operatorname{Kern}\varphi$ eine Untergruppe von G ist. Sei nun $b \in \operatorname{Kern}\varphi$; für jedes $a \in G$ ist

$$\varphi(aba^{-1}) = \varphi(a)\varphi(b)\varphi(a^{-1}) = \varphi(a)e\varphi(a)^{-1} = \varphi(a)\varphi(a)^{-1} = e\ ,$$

d.h., $\varphi(aba^{-1}) \in \operatorname{Kern}\varphi$. Damit ist Kern φ Normalteiler von G.

(2) Seien a, $b \in \operatorname{Bild}\varphi$; es gibt also c, $d \in G$ mit $a = \varphi(c)$ und $b = \varphi(d)$, und folglich ist $a^{-1}b = \varphi(c)^{-1}\varphi(d) = \varphi(c^{-1})\varphi(d) = \varphi(c^{-1}d)$, d.h., $a^{-1}b \in$ Bild φ. Ferner ist $\operatorname{Bild}\varphi \neq \varnothing$, da $e = \varphi(e) \in \operatorname{Bild}\varphi$. Nach Lemma 13.3 ist also $\operatorname{Bild}\varphi$ eine Untergruppe von H. □

Beispiele: (1) Seien G, H Gruppen und definiere $\varphi : G \to H$ durch $\varphi(a) = e$ für alle $a \in G$. Dann ist φ ein Homomorphismus mit $\operatorname{Kern}\varphi = G$ und $\operatorname{Bild}\varphi = \{e\}$.

(2) Sei $m \geq 1$ und definiere $\varphi : \mathbb{Z} \to \mathbb{Z}_m$ durch $\varphi(n) = n \operatorname{Rest} m$. Dann ist φ ein Homomorphismus mit $\operatorname{Kern}\varphi = \{nm : n \in \mathbb{Z}\}$ und $\operatorname{Bild}\varphi = \mathbb{Z}_m$.

(3) Sei G eine Gruppe und $a \in G$; definiere eine Abbildung $\varphi : \mathbb{Z} \to G$ durch $\varphi(n) = a^n$ für jedes $n \in \mathbb{Z}$. Dann ist φ ein Homomorphismus mit $\operatorname{Bild}\varphi = \langle a\rangle$. Es gilt $\operatorname{Kern}\varphi = \{0\}$, falls a ein Element unendlicher Ordnung ist, und wenn a ein Element endlicher Ordnung ist, so ist $\operatorname{Kern}\varphi = \{nm : n \in \mathbb{Z}\}$, wobei $m = \operatorname{ord} a$.

(4) Man betrachte die zwei Gruppen $(\mathbb{R}, +, 0)$ und $(\mathbb{R}\setminus\{0\}, \cdot, 1)$ und definiere eine Abbildung $\varphi : \mathbb{R} \to \mathbb{R} \setminus \{0\}$ durch $\varphi(x) = e^x$. Dann ist φ ein Homomorphismus mit $\operatorname{Kern}\varphi = \{0\}$ und $\operatorname{Bild}\varphi = \{x \in \mathbb{R} \setminus \{0\} : x > 0\}$.

(5) Man betrachte die zwei Gruppen $(\mathbb{R}, +, 0)$ und $(\mathbb{C}^\times, \cdot, 1)$ (mit $\mathbb{C}^\times = \mathbb{C} \setminus \{0\}$) und definiere eine Abbildung $\varphi : \mathbb{R} \to \mathbb{C}^\times$ durch $\varphi(x) = e^{2\pi i x}$. Dann ist φ ein Homomorphismus mit $\operatorname{Kern}\varphi = \mathbb{Z}$ und $\operatorname{Bild}\varphi = \{z \in \mathbb{C}^\times : |z| = 1\}$.

Lemma 13.17 *Sei $\varphi : G \to H$ ein Homomorphismus. Dann ist φ surjektiv genau, wenn* $\operatorname{Bild}\varphi = H$*, und injektiv genau, wenn* $\operatorname{Kern}\varphi = \{e\}$.

Beweis Es ist klar, dass φ genau dann surjektiv ist, wenn $\operatorname{Bild}\varphi = H$. Ferner ist $\operatorname{Kern}\varphi = \{e\}$, wenn φ injektiv ist, da $\varphi(e) = e$. Sei also umgekehrt $\operatorname{Kern}\varphi = \{e\}$ und seien a, $b \in G$ mit $\varphi(a) = \varphi(b)$. Dann ist

$$\varphi(a^{-1}b) = \varphi(a^{-1})\varphi(b) = \varphi(a)^{-1}\varphi(b) = \varphi(b)^{-1}\varphi(b) = e\ ,$$

d.h., $a^{-1}b \in \operatorname{Kern}\varphi$. Damit ist $a^{-1}b = e$, d.h., $a = b$. Folglich ist φ injektiv. □

Sei $\varphi : G \to H$ ein Homomorphismus. Nach Lemmas 13.15 und 13.17 ist φ ein Isomorphismus genau dann, wenn $\operatorname{Kern}\varphi = \{e\}$ und $\operatorname{Bild}\varphi = H$. Die Gruppen G und H heißen *isomorph*, wenn es einen Isomorphismus von G auf H gibt. Sind G und H isomorph, so schreibt man $G \cong H$. Es gilt $G \cong G$ für jede Gruppe G, ferner ist $H \cong G$, wenn $G \cong H$, und $G \cong K$, falls $G \cong H$ und $H \cong K$.

Satz 13.6 *(1) Jede unendliche zyklische Gruppe ist isomorph zu* $\mathbb{Z}$.

(2) Jede endliche zyklische Gruppe der Ordnung m *ist isomorph zu* $\mathbb{Z}_m$.

Beweis (1) Sei G eine unendliche zyklische Gruppe, sei $a \in G$ mit $G = \langle a \rangle$ und sei $\varphi : \mathbb{Z} \to G$ durch $\varphi(n) = a^n$ gegeben. Dann ist φ ein Homomorphismus und nach Lemma 13.6 (1) ist φ bijektiv. Folglich ist nach Lemma 13.15 φ ein Isomorphismus.

(2) Sei G eine endliche zyklische Gruppe der Ordnung m. Nach Lemma 13.6 (2) gibt es dann $a \in G$ mit $\operatorname{ord} a = m$, so dass $G = \{e, a, \ldots, a^{m-1}\}$. Sei $\varphi : \mathbb{Z}_m \to G$ gegeben durch $\varphi(n) = a^n$ für $n = 0, \ldots, m-1$. Man sieht leicht, dass φ ein Homomorphismus ist und damit ist nach Lemma 13.15 φ ein Isomorphismus. □

Sei p eine Primzahl. Nach Satz 13.4 und Satz 13.6 (2) ist jede endliche Gruppe der Ordnung p isomorph zu $\mathbb{Z}_p$.

Satz 13.7 *Sei* G *eine Gruppe und* H *ein Normalteiler von* G*, sei* $\pi : G \to G/H$ *die durch* $\pi(a) = Ha$ *definierte Abbildung. Dann ist* π *ein Homomorphismus mit* $\operatorname{Kern} \varphi = H$ *und* $\operatorname{Bild} \varphi = G/H$.

Beweis Nach Lemma 13.10 ist $\pi(ab) = H(ab) = (Ha)(Hb) = \pi(a)\pi(b)$ für alle $a, b \in G$, und damit ist π ein Homomorphismus. Ferner ist

$$\operatorname{Kern} \pi = \{a \in G : Ha = H\} = H$$

und es ist klar, dass $\operatorname{Bild} \pi = G/H$. □

Der Homomorphismus $\pi : G \to G/H$ heißt der *kanonische Homomorphismus* von G auf G/H.

Sei S eine Teilmenge einer Gruppe G; nach Satz 13.7 und Lemma 13.16 (1) gibt es eine Gruppe H und einen Homomorphismus $\varphi : G \to H$ mit $\operatorname{Kern} \varphi = S$ genau dann, wenn S Normalteiler von G ist.

Satz 13.8 (Homomorphiesatz) *Seien* G, H *Gruppen und sei* $\varphi : G \to H$ *ein surjektiver Homomorphismus; setze* $K = \operatorname{Kern} \varphi$ *und sei* $\pi : G \to G/K$ *der kanonische Homomorphismus von* G *auf* G/K*. Dann gibt es einen Isomorphismus* $\psi : G/K \to H$*, so dass* $\varphi = \psi \circ \pi$*. Insbesondere ist* $G/K \cong H$.

Beweis Seien $a, b \in G$ mit $Ka = Kb$; da $a \in Ka = Kb$, gibt es dann $c \in K$, so dass $a = cb$, und folglich ist $\varphi(a) = \varphi(cb) = \varphi(c)\varphi(b) = e\varphi(b) = \varphi(b)$, da $c \in K = \operatorname{Kern} \varphi$. Sind also $a, b \in G$ mit $Ka = Kb$, so ist $\varphi(a) = \varphi(b)$, und daher

kann eine Abbildung $\psi : G/K \to H$ durch $\psi(Ka) = \varphi(a)$ definiert werden. Für jedes $a \in G$ ist dann $\varphi(a) = \psi(Ka) = \psi(\pi(a))$, d.h., $\varphi = \psi\pi$.

ψ ist ein Homomorphismus: Seien $X, Y \in G/K$; dann gibt es $a, b \in G$ mit $X = Ka$ und $Y = Kb$, und nach Lemma 13.10 ist

$$\begin{aligned}\psi(XY) = \psi((Ka)(Kb)) &= \psi(K(ab)) = \varphi(ab)\\ &= \varphi(a)\varphi(b) = \psi(Ka)\psi(Kb) = \psi(X)\psi(Y)\ .\end{aligned}$$

ψ ist ein Isomorphismus: Sei $a \in H$; da φ surjektiv ist, gibt es $b \in G$ mit $\varphi(b) = a$, und dann ist $\psi(Kb) = \varphi(b) = a$, d.h., ψ ist surjektiv. Sei nun $a \in G$; dann gilt $\psi(Ka) = e$, wenn $a \in \text{Kern } \varphi = K$, da $\psi(Ka) = \varphi(a)$. Folglich ist $\text{Kern}\,\psi = \{K\}$, und nach Lemma 13.17 ist dann ψ injektiv. Nach Lemma 13.15 is also ψ ein Isomorphismus. □

Seien G, H Gruppen und sei $\varphi : G \to H$ ein Homomorphismus. Dann kann φ als surjektiven Homomorphismus von G auf $\text{Bild}\,\varphi$ angesehen werden, und folglich ist nach Satz 13.8 $G/\text{Kern}\,\varphi \cong \text{Bild}\,\varphi$.

Für $m \geq 1$ sei $m\mathbb{Z} = \{mn : n \in \mathbb{Z}\}$, also ist $m\mathbb{Z}$ eine Untergruppe von $\mathbb{Z}$. Es gibt einen surjektiven Homomorphismus $\varphi_m : \mathbb{Z} \to \mathbb{Z}_m$, der durch $\varphi_m(n) = n \,\text{Rest}\, m$ definiert ist, und $\text{Kern}\,\varphi_m = m\mathbb{Z}$. Damit ist nach Satz 13.8 $\mathbb{Z}/m\mathbb{Z} \cong \mathbb{Z}_m$.

Sei $\varphi : \mathbb{R} \to \mathbb{C}^\times$ der Homomorphismus mit $\varphi(x) = e^{2\pi ix}$ für jedes $x \in \mathbb{R}$. Hier ist $\text{Kern}\,\varphi = \mathbb{Z}$ und $\text{Bild}\,\varphi = \{z \in \mathbb{C}^\times : |z| = 1\}$, und daraus folgt nach Satz 13.8, dass $\mathbb{R}/\mathbb{Z} \cong \{z \in \mathbb{C}^\times : |z| = 1\}$.

Für $n \geq 2$ sei S_n die Menge aller Bijektionen $\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$. Dann ist S_n eine endliche Gruppe der Ordnung $n!$. Sie heißt die *symmetrische Gruppe vom Grad n*. Das Produkt von σ und τ in S_n, das eigentlich $\sigma \circ \tau$ ist, wird aber einfach als $\sigma\tau$ geschrieben.

Die Elemente von S_n heißen *Permutationen vom Grad n*. Ein Element $\sigma \in S_n$ wird häufig angegeben in der Form

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}\ .$$

Zum Beispiel sind $\begin{pmatrix} 1&2&3&4&5\\ 3&1&5&4&2 \end{pmatrix}$ und $\begin{pmatrix} 1&2&3&4&5\\ 5&4&1&2&3 \end{pmatrix}$ Permutationen vom Grad 5.

Sind σ, $\tau \in S_n$, so gilt für das Produkt $\sigma\tau$:

$$\begin{pmatrix} 1 & \cdots & n \\ \sigma(1) & \cdots & \sigma(n) \end{pmatrix} \begin{pmatrix} 1 & \cdots & n \\ \tau(1) & \cdots & \tau(n) \end{pmatrix} = \begin{pmatrix} 1 & \cdots & n \\ \sigma(\tau(1)) & \cdots & \sigma(\tau(n)) \end{pmatrix}\ .$$

Zum Beispiel ist $\begin{pmatrix} 1&2&3&4&5\\ 3&1&5&4&2 \end{pmatrix} \begin{pmatrix} 1&2&3&4&5\\ 5&4&1&2&3 \end{pmatrix} = \begin{pmatrix} 1&2&3&4&5\\ 2&4&3&1&5 \end{pmatrix}$ in S_5.

Eine Permutation $\sigma \in S_n$ heißt *Transposition*, wenn $\{1 \leq j \leq n : \sigma(j) \neq j\}$ aus genau zwei Elementen besteht.

Satz 13.9 *Jede Permutation ist als Produkt von Transpositionen darstellbar. (Diese Darstellung als Produkt von Transpositionen ist aber nie eindeutig.)*

Beweis Das weiß jedes fünfjährige Kind. □

Definiere eine Abbildung $\text{sign} : S_n \to \{-1, 1\}$ durch $\text{sign}(\sigma) = (-1)^{\delta(\sigma)}$, wobei $\delta(\sigma)$ die Anzahl der Elemente in der Menge

$$\{(i,j) : 1 \leq i < j \leq n \text{ mit } \sigma(i) > \sigma(j)\}$$

ist; $\text{sign}(\sigma)$ heißt die *Signatur* von σ.

Lemma 13.18 *Für jedes $\sigma \in S_n$ ist*

$$\text{sign}(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(i) - \sigma(j)}{i - j} \, .$$

Beweis Sei $P = \{(i,j) : 1 \leq i < j \leq n\}$ und definiere eine Abbildung $\Delta : P \to \mathbb{Z}$ durch $\Delta((i,j)) = i - j$. Für jedes $\sigma \in S_n$ definiere $\hat{\sigma} : P \to P$ durch

$$\hat{\sigma}((i,j)) = \begin{cases} (\sigma(i), \sigma(j)) & \text{falls } \sigma(i) < \sigma(j) \, , \\ (\sigma(j), \sigma(i)) & \text{falls } \sigma(j) < \sigma(i) \, . \end{cases}$$

Man sieht leicht, dass $\hat{\sigma}$ surjektiv ist und damit ist $\hat{\sigma}$ eine Bijektion. Setze auch

$$\varepsilon(\sigma, (i,j)) = \begin{cases} 1 & \text{falls } \sigma(i) < \sigma(j) \, , \\ -1 & \text{falls } \sigma(j) < \sigma(i) \, ; \end{cases}$$

dann ist $\sigma(i) - \sigma(j) = \varepsilon(\sigma, (i,j))\Delta(\hat{\sigma}((i,j)))$ für jedes $(i,j) \in P$ und folglich ist

$$\begin{aligned} \prod_{(i,j) \in P} (\sigma(i) - \sigma(j)) &= \prod_{(i,j) \in P} \varepsilon(\sigma, (i,j)) \Delta(\hat{\sigma}((i,j))) \\ &= \prod_{(i,j) \in P} \varepsilon(\sigma, (i,j)) \prod_{(i,j) \in P} \Delta(\hat{\sigma}((i,j))) \\ &= \prod_{(i,j) \in P} \varepsilon(\sigma, (i,j)) \prod_{(i,j) \in P} \Delta((i,j)) \quad \text{(da } \hat{\sigma} \text{ eine Bijektion ist)} \\ &= \prod_{(i,j) \in P} \varepsilon(\sigma, (i,j)) \prod_{(i,j) \in P} (i - j) \, . \end{aligned}$$

Aber $\prod_{(i,j)\in P} \varepsilon(\sigma,(i,j)) = \operatorname{sign}(\sigma)$ und daraus ergibt sich, dass

$$\operatorname{sign}(\sigma) = \prod_{(i,j)\in P} \varepsilon(\sigma,(i,j)) = \prod_{(i,j)\in P} \frac{\sigma(i)-\sigma(j)}{i-j} = \prod_{1\le i<j\le n} \frac{\sigma(i)-\sigma(j)}{i-j} \; . \; \square$$

Betrachte nun $\{-1,1\}$ als Gruppe und zwar als Untergruppe von $\mathbb{R}^\times = \mathbb{R}\setminus\{0\}$.

Satz 13.10 *Die Abbildung sign* : $S_n \to \{-1,1\}$ *ist ein Homomorphismus.*

Beweis Die Schreibweise aus dem Beweis für Lemma 13.18 wird verwendet. Seien σ, $\tau \in S_n$; dann gilt $\varepsilon(\sigma\tau,(i,j)) = \varepsilon(\sigma,\hat\tau((i,j)))\varepsilon(\tau,(i,j))$ für jedes $(i,j)\in P$ und daraus folgt, dass

$$\begin{aligned}
\operatorname{sign}(\sigma\tau) &= \prod_{(i,j)\in P} \varepsilon(\sigma\tau,(i,j)) \\
&= \prod_{(i,j)\in P} \varepsilon(\sigma,\hat\tau((i,j)))\varepsilon(\tau,(i,j)) \\
&= \prod_{(i,j)\in P} \varepsilon(\sigma,\hat\tau((i,j))) \prod_{(i,j)\in P} \varepsilon(\tau,(i,j)) \\
&= \prod_{(i,j)\in P} \varepsilon(\sigma,(i,j)) \prod_{(i,j)\in P} \varepsilon(\tau,(i,j)) \quad \text{(da } \hat\tau \text{ eine Bijektion ist)} \\
&= \operatorname{sign}(\sigma)\operatorname{sign}(\tau) \; .
\end{aligned}$$

Damit ist sign ein Homomorphismus. $\square$

Eine Permutation $\sigma \in S_n$ heißt *gerade*, wenn $\operatorname{sign}(\sigma) = 1$ und *ungerade*, wenn $\operatorname{sign}(\sigma) = -1$. Jede Transposition ist offensichtlich ungerade. Sei nun $\sigma \in S_n$ und sei $\sigma = \tau_1 \cdots \tau_m$ eine Darstellung von σ als Produkt von Transpositionen. Nach Satz 13.10 ist dann $\operatorname{sign}(\sigma) = \prod_{j=1}^m \operatorname{sign}(\tau_j) = (-1)^m$. Nach Satz 13.9 ist also eine Permutation genau dann gerade bzw. ungerade, wenn sie als Produkt von einer geraden bzw. ungeraden Anzahl von Transpostionen dargestellt werden kann.

Sei A_n die Menge der geraden Permutationen in S_n; da $A_n = \operatorname{Kern}\operatorname{sign}$, ist nach Lemma 13.16 (1) A_n ein Normalteiler von S_n. A_n heißt die *alternierende Gruppe vom Rang n*. Nach Satz 13.8 ist $S_n/A_n \cong \{-1,1\}$, und insbesondere ist $[S_n : A_n] = 2$. Daraus folgt, dass $\operatorname{ord} A_n = \frac{1}{2}n!$.

Eine Gruppe G heißt *einfach*, wenn $\{e\}$ und G die einzigen Normalteiler von G sind. Zum Beispiel ist die Gruppe $\mathbb{Z}_p$ einfach, wenn p Primzahl ist, da nach Satz 13.2 $\{e\}$ und $\mathbb{Z}_p$ die einzigen Untergruppen von $\mathbb{Z}_p$ sind.

Satz 13.11 *Für* $n \ge 5$ *ist* A_n *eine einfache Gruppe.*

Beweis Übung für Streberinnen und Streber. □

Satz 13.12 (Satz von Cayley) *Sei G eine endliche Gruppe der Ordnung n mit $n \geq 2$. Dann ist G isomorph zu einer Untergruppe von S_n.*

Beweis Wähle irgendeine bijektive Abbildung $\gamma : \{1, 2, \ldots, n\} \to G$. Für jedes $a \in G$ sei $\sigma_a : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ die Abbildung, die gegeben ist durch

$$\sigma_a(j) = \gamma^{-1}(a\gamma(j))$$

für jedes $j = 1, \ldots, n$. Dann ist σ_a injektiv, da

$$\begin{aligned} \sigma_a(i) = \sigma_a(j) &\Rightarrow \gamma^{-1}(a\gamma(i)) = \gamma^{-1}(a\gamma(j)) \Rightarrow a\gamma(i) = a\gamma(j) \\ &\Rightarrow a^{-1}a\gamma(i) = a^{-1}a\gamma(j) \Rightarrow e\gamma(i) = e\gamma(j) \\ &\Rightarrow \gamma(i) = \gamma(j) \Rightarrow i = j\,. \end{aligned}$$

Damit ist σ_a bijektiv, d.h., $\sigma_a \in S_n$. Man kann also eine Abbildung $\psi : G \to S_n$ definieren durch $\psi(a) = \sigma_a$ und diese Abbildung ψ ist dann ein Homomorphismus: Für alle a, $b \in G$ und $1 \leq j \leq n$ ist

$$\begin{aligned} (\psi(a)\psi(b))(j) &= \psi(a)\,(\psi(b)(j)) = \psi(a)(\sigma_b(j)) \\ &= \psi(a)(\gamma^{-1}(b\gamma(j))) = \sigma_a(\gamma^{-1}(b\gamma(j))) = \gamma^{-1}(a\gamma(\gamma^{-1}(b\gamma(j)))) \\ &= \gamma^{-1}(ab\gamma(j)) = \sigma_{ab}(j) = \psi(ab)(j) \end{aligned}$$

und folglich ist $\psi(ab) = \psi(a)\psi(b)$ für alle a, $b \in G$. Ferner ist ψ injektiv, da

$$\begin{aligned} \psi(a) = \psi(b) &\Rightarrow \psi(a)(j) = \psi(b)(j) \text{ für alle } j \\ &\Rightarrow \gamma^{-1}(a\gamma(j)) = \gamma^{-1}(b\gamma(j)) \text{ für alle } j \Rightarrow a = b\,. \end{aligned}$$

Sei $H = \operatorname{Bild}\psi$; nach Lemma 13.16 (2) ist H eine Untergruppe von S_n und ψ kann als bijektiven Homomorphismus von G auf H angesehen werden. Damit sind G und H isomorph. □

14 Ringe

Ein 4-Tupel $(R, +, \cdot, 0)$ bestehend aus einer Menge R, einer Verknüpfung

$$\begin{aligned} + : R \times R &\to R \\ (a, b) &\mapsto a + b \end{aligned}$$

(genannt *Addition*), einer Verknüpfung

$$\begin{aligned} \cdot : R \times R &\to R \\ (a, b) &\mapsto ab \end{aligned}$$

(genannt *Multiplikation*) und einem Element $0 \in R$ heißt *Ring*, wenn gilt:

(R1) $(a + b) + c = a + (b + c)$ für alle a, b, $c \in R$.

(R2) $a + b = b + a$ für alle a, $b \in R$.

(R3) Für alle $a \in R$ gilt $0 + a = a$.

(R4) Zu jedem $a \in R$ gibt es ein Element $-a \in R$ mit $(-a) + a = 0$.

(R5) $(ab)c = a(bc)$ für alle a, b, $c \in R$.

(R6) $a(b + c) = ab + ac$ und $(b + c)a = ba + ca$ für alle a, b, $c \in R$.

Bemerkung: Nach der üblichen Konvention soll die Addition in R weniger stark binden als die Multiplikation. ($ab + ac$ bedeutet also $(ab) + (ac)$.)

Ist $(R, +, \cdot, 0)$ ein Ring, so ist $(R, +, 0)$ eine abelsche Gruppe.

Lemma 14.1 *Sei $(R, +, \cdot, 0)$ ein Ring.*

(1) Das Nullelement 0 ist eindeutig: Ist $0' \in R$ ein Element mit $0' + a = a$ für alle $a \in R$, so ist $0' = 0$.

(2) Zu jedem $a \in R$ gibt es genau ein Element $-a \in R$ mit $(-a) + a = 0$.

Beweis Übung. □

Wenn es aus dem Kontext klar ist, welche Verknüpfungen $+$ und $\cdot$ und welches Element 0 gemeint sind, dann schreibt man meistens lediglich R statt $(R, +, \cdot, 0)$.

Lemma 14.2 *Sei R ein Ring; dann gilt $a \cdot 0 = 0 \cdot a = 0$ für alle $a \in R$.*

Beweis Übung. □

Ein Ring R heißt *kommutativ*, wenn gilt: $ab = ba$ für alle a, $b \in R$.

Lemma 14.3 *Sei R ein Ring und seien $1, 1' \in R$ Elemente mit $1a = a1 = a$ und $1'a = a1' = a$ für alle $a \in R$. Dann ist $1' = 1$.*

Beweis Übung. □

Ein Ring R heißt *Ring mit 1*, wenn es ein Element $1 \in R$ mit $1 \neq 0$ gibt, so dass $1a = a1 = a$ für alle $a \in R$. Ist R ein Ring mit 1, so ist nach Lemma 14.3 das Einselement 1 eindeutig.

Beispiele von Ringen

1. Jeder Körper ist ein kommutativer Ring mit 1.

2. $\mathbb{Z}$ ist ein kommutativer Ring mit 1.

3. Für jedes $n > 1$ ist $\mathbb{Z}_n$ ein kommutativer Ring mit 1.

4. Für jeden Körper K ist $M(n \times n, K)$ ein Ring mit 1 (mit Einselement E_n). Dieser Ring ist aber nicht kommutativ, wenn $n > 1$.

5. Sei V ein Vektorraum über einem Körper K und sei $\mathrm{End}(V)$ die Menge aller Endomorphismen von V. Dann ist $\mathrm{End}(V)$ ein Ring mit 1. Dieser Ring ist aber nur kommutativ, wenn $\dim V = 1$.

6. Sei $[0,1] = \{x \in \mathbb{R} : 0 \leq x \leq 1\}$ und sei $\mathrm{C}([0,1])$ die Menge aller stetigen Abbildungen von $[0,1]$ nach $\mathbb{R}$. Dann ist $\mathrm{C}([0,1])$ ein kommutativer Ring mit 1, wobei hier die Addition durch $(f+g)(x) = f(x) + g(x)$ und die Multiplikation durch $(fg)(x) = f(x)g(x)$ für alle $x \in [0,1]$ definiert sind.

Sei R ein Ring und seien $a_1, \ldots, a_n \in R$ mit $n \geq 3$; dann ist die 'Summe' von $a_1, \ldots, a_n$ unabhängig von der Reihenfolge der einzelnen Additionen und wird mit $a_1 + \cdots + a_n$ bezeichnet. Genauso ist das 'Produkt' von $a_1, \ldots, a_n$ unabhängig von der Reihenfolge der einzelnen Multiplikationen und wird mit $a_1 \cdots a_n$ bezeichnet.

Sei R ein Ring; eine Teilmenge S von R heißt *Unterring* von R, wenn $0 \in S$ und für alle $a, b \in S$ auch $-a$, $a + b$ und ab in S liegen. (S ist also bereits selbst ein Ring bezüglich der in R gegebenen Operationen.)

Sei R ein Ring; eine Teilmenge U von R heißt *Linksideal* bzw. *Rechtsideal* von R, wenn gilt:

— $(U, +, 0)$ ist eine Untergruppe von $(R, +, 0)$. Es gilt also $0 \in U$ und für alle $a, b \in U$ sind $-a$ und $a + b$ in U.

— Für alle $a \in R$, $u \in U$ ist $au \in U$ bzw. ist $ua \in U$.

Nun heißt U *Ideal*, wenn es Links- und Rechtsideal ist. Insbesondere sind R und $\{0\}$ stets Ideale von R.

Sei R ein Ring und U ein Linksideal oder ein Rechtsideal von R; dann ist U ein Unterring von R. Ist R ein Ring mit 1 und $1 \in U$, dann ist $U = R$.

Wenn R kommutativ ist, ist jedes Linksideal ein Rechtsideal und umgekehrt.

Beispiel Sei $R = M(2 \times 2, \mathbb{R})$ der Ring der 2×2 Matrizen über $\mathbb{R}$ und sei

$$U = \left\{ \begin{pmatrix} a & 0 \\ b & 0 \end{pmatrix} : a, b \in \mathbb{R} \right\};$$

dann ist U ein Linksideal von R aber kein Rechtsideal.

Sei R ein kommutativer Ring; für jedes $a \in R$ setze $(a) = \{ab : b \in R\}$. Es ist dann klar, dass (a) ein Ideal von R ist. Ein Ideal U von R heißt *Hauptideal*, wenn $U = (a)$ für ein $a \in R$.

Satz 14.1 *Jedes Ideal von $\mathbb{Z}$ ist ein Hauptideal.*

Beweis Übung. □

Seien R, R' Ringe; eine Abbildung $\varphi : R \to R'$ heißt ein *Homomorphismus* (oder ein *Ring-Homomorphismus*), wenn gilt:

$$\varphi(a+b) = \varphi(a) + \varphi(b) \quad \text{und} \quad \varphi(ab) = \varphi(a)\varphi(b)$$

für alle a, $b \in R$.

Lemma 14.4 *Sei $\varphi : R \to R'$ ein Homomorphismus. Dann gilt $\varphi(0) = 0$ und $\varphi(-a) = -\varphi(a)$ für alle $a \in R$.*

Beweis Dies folgt unmittelbar aus Lemma 13.13. □

Für einen Homomorphismus $\varphi : R \to R'$ setze

$$\operatorname{Kern} \varphi = \{a \in R : \varphi(a) = 0\} \quad \text{und} \quad \operatorname{Bild} \varphi = \varphi(R) \, .$$

Lemma 14.5 *(1)* $\operatorname{Kern} \varphi$ *ist ein Ideal von R.*

(2) $\operatorname{Bild} \varphi$ *ist ein Unterring von R'.*

Beweis (1) Da $\varphi(0) = 0$, ist $0 \in \operatorname{Kern} \varphi$. Seien a, $b \in \operatorname{Kern}\ \varphi$; nach Lemma 14.3 ist $\varphi(-a) = -\varphi(a) = -0 = 0$, und $\varphi(a+b) = \varphi(a) + \varphi(b) = 0 + 0 = 0$. Damit sind $-a$ und $a+b$ auch in $\operatorname{Kern} \varphi$. Seien nun $a \in R$ und $u \in \operatorname{Kern} \varphi$; dann gilt nach Lemma 14.2, dass $\varphi(au) = \varphi(a)\varphi(u) = \varphi(a) \cdot 0 = 0$ und genauso gilt $\varphi(ua) = \varphi(u)\varphi(a) = 0 \cdot \varphi(a) = 0$. Folglich ist $\operatorname{Kern} \varphi$ ein Ideal von R.

(2) Da $\varphi(0) = 0$, ist $0 \in \operatorname{Bild}\varphi$. Seien a', $b' \in \operatorname{Bild}\varphi$; es gibt also a, $b \in R$ mit $a' = \varphi(a)$ und $b' = \varphi(b)$. Nach Lemma 14.4 ist $-a' = -\varphi(a) = \varphi(-a)$, und ferner gilt $a' + b' = \varphi(a) + \varphi(b) = \varphi(a + b)$ und $a'b' = \varphi(a)\varphi(b) = \varphi(ab)$. Damit liegen $-a'$, $a' + b'$ und $a'b'$ auch in $\operatorname{Bild}\varphi$. Folglich ist $\operatorname{Bild}\varphi$ ein Unterring von R'. □

Im Folgenden sei R ein kommutativer Ring mit 1.

Bezeichne nun mit $\Sigma(R)$ die Menge aller Folgen $\{a_n\}_{n\geq 0}$ aus R. Verknüpfungen $+ : \Sigma(R) \times \Sigma(R) \to \Sigma(R)$ und $\cdot : \Sigma(R) \times \Sigma(R) \to \Sigma(R)$ werden definiert durch $\{a_n\}_{n\geq 0} + \{b_n\}_{n\geq 0} = \{a_n + b_n\}_{n\geq 0}$ und $\{a_n\}_{n\geq 0}\{b_n\}_{n\geq 0} = \{c_n\}_{n\geq 0}$, wobei

$$c_n = a_0b_n + a_1b_{n-1} + \cdots + a_{n-1}b_1 + a_nb_0 \ .$$

Dann ist $\Sigma(R)$ ein kommutativer Ring mit 1. (Der Beweis dafür ist eine Übung.) Das Nullelement ist die Folge $0 = \{0_n\}_{n\geq 0}$, wobei $0_n = 0$ für alle $n \geq 0$, und das Einselement ist die Folge $1 = \{1_n\}_{n\geq 0}$, wobei

$$1_n = \begin{cases} 1 & \text{falls } n = 0 \ , \\ 0 & \text{sonst} \ . \end{cases}$$

Sei nun $\Sigma_o(R)$ die Teilmenge von $\Sigma(R)$, die definiert ist durch

$$\Sigma_o(R) = \{\{a_n\}_{n\geq 0} \in \Sigma(R) : a_n \neq 0 \text{ für nur endlich viel } n \geq 0\} \ .$$

Lemma 14.6 *Für alle α, $\beta \in \Sigma_o(R)$ sind $-\alpha$, $\alpha + \beta$ und $\alpha\beta$ auch in $\Sigma_o(R)$. Ferner sind 0, $1 \in \Sigma(R)$ Elemente von $\Sigma_o(R)$.*

Beweis Seien $\alpha = \{a_n\}_{n\geq 0}$, $\beta = \{b_n\}_{n\geq 0} \in \Sigma_o(R)$. Es gibt also M, $N \geq 0$, so dass $a_n = 0$ für alle $n > M$ und $b_n = 0$ für alle $n > N$. Sei nun $\alpha\beta = \{c_n\}_{n\geq 0}$; dann ist $c_n = a_0b_n + a_1b_{n-1} + \cdots + a_{n-1}b_1 + a_nb_0 = 0$, falls $n > M + N$, d.h., $\alpha\beta \in \Sigma_o(R)$. Die anderen Teile sind klar. □

Nach Lemma 14.6 ist $\Sigma_o(R)$, mit den Verknüpfungen aus $\Sigma(R)$, ein kommutativer Ring mit 1. Sei nun x ein Symbol, das als 'Unbestimmte' angesehen werden soll. Ist $\alpha = \{a_n\}_{n\geq 0}$ ein Element von $\Sigma_o(R)$ mit $a_n = 0$ für jedes $n > m$, so wird α durch den Ausdruck

$$a_0 + a_1x + a_2x^2 + \cdots + a_mx^m$$

dargestellt. Diese Darstellung ist nicht eindeutig, sind aber $a_0 + a_1x + \cdots + a_mx^m$ und $b_0 + b_1x + \cdots + b_\ell x^\ell$ zwei Darstellungen und ist $m \leq \ell$, so ist $a_j = b_j$ für $j = 1, \ldots, m$ und $b_j = 0$ für $m < j \leq \ell$.

Man schreibt nun $R[x]$ statt $\Sigma_o(R)$; $R[x]$ heißt der *Polynomring* über R in der Unbestimmten x, die Elemente aus $R[x]$ heißen *Polynome* in x mit Koeffizienten aus R.

Sei $i : R \to R[x]$ die Abbildung, die durch $i(a) = a$ definiert ist. Mit anderen Worten ist $i(a) = \{a_n\}_{n\geq 0}$, wobei

$$a_n = \begin{cases} a & \text{falls } n = 0 \,, \\ 0 & \text{sonst} \,. \end{cases}$$

Diese Abbildung ist injektiv und folglich kann man jedes Element a aus R mit dem Element $i(a)$ aus $R[x]$ identifizieren. Auf diese Weise wird R als Teilmenge von $R[x]$ betrachtet. Da $i(a+b) = i(a)+i(b)$ und $i(ab) = i(a)i(b)$ für alle $a, b \in R$ und $i(0) = 0$ und $i(1) = 1$, erhält man dann die Addition und Multiplikation auf R als Einschränkungen der entsprechenden Verknüpfungen auf $R[x]$.

Ist $f \in R[x]$ mit $f \neq 0$, so hat f eine eindeutige Darstellung

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m$$

mit $a_m \neq 0$. Die Zahl m heißt der *Grad* des Polynoms f und wird mit $\operatorname{Grad} f$ bezeichnet. Das Element a_m heißt der *Leitkoeffizient* von f; ist $a_m = 1$, so heißt f *normiert*.

Für jedes $a \in R$ definiere eine Abbildung $\Phi_a : R[x] \to R$ durch

$$\Phi_a(f) = a_0 + a_1 a + a_2 a^2 + \cdots + a_m a^m$$

falls $f = a_0+a_1x+a_2x^2+\cdots+a_mx^m$ mit $a_m \neq 0$, (wobei $a^2 = aa$, $a^3 = aaa$ usw.). Also ist $\Phi_a(f)$ der Wert des Polynoms f an der Stelle a und folglich schreibt man meistens $f(a)$ statt $\Phi_a(f)$.

Lemma 14.7 *Die Abbildung $\Phi_a : R[x] \to R$ ist ein Homomorphismus.*

Beweis Übung. □

Die Abbildung Φ_a wird *Einsetzhomomorphismus* genannt.

Im Folgenden sei K ein Körper.

Lemma 14.8 *Seien $f, g \in K[x]$ mit $f \neq 0$, $g \neq 0$; dann ist*

$$\operatorname{Grad} fg = \operatorname{Grad} f + \operatorname{Grad} g \,.$$

Beweis Sei $\lambda_0 + \lambda_1 x + \cdots + \lambda_m x^m$ bzw. $\mu_0 + \mu_1 x + \cdots + \mu_n x^n$ die eindeutige Darstellung von f mit $\lambda_m \neq 0$ bzw. von g mit $\mu_n \neq 0$. Dann ist

$$fg = \nu_0 + \nu_1 x + \cdots + \nu_{m+n} x^{m+n}$$

mit $\nu_{m+n} = \lambda_m \mu_n \neq 0$, und folglich ist $\operatorname{Grad} fg = m + n = \operatorname{Grad} f + \operatorname{Grad} g$. □

Lemma 14.9 *(1) Sind f, $g \in K[x]$ mit $f \neq 0$ und $g \neq 0$, so ist $fg \neq 0$.*
(2) Sind f, g_1, $g_2 \in K[x]$ mit $f \neq 0$ und $fg_1 = fg_2$, so ist $g_1 = g_2$.

Beweis (1) Dies folgt unmittelbar aus Lemma 14.8.
(2) Da $f(g_1 - g_2) = 0$ und $f \neq 0$, ist nach (1) $g_1 - g_2 = 0$, d.h., $g_1 = g_2$. □

Satz 14.2 (Division mit Rest in $K[x]$) *Seien f, $g \in K[x]$ mit $g \neq 0$. Dann gibt es eindeutige Polynome t, $r \in K[x]$ mit $f = tg + r$ und entweder $r = 0$ oder* $\operatorname{Grad} r < \operatorname{Grad} g$.

Beweis Existenz von t und r: Da $0 = 0g+0$, kann man annehmen, dass $f \neq 0$. Ist $\operatorname{Grad} f < \operatorname{Grad} g$, setze einfach $t = 0$ und $r = f$. Es kann also auch angenommen werden, dass $\operatorname{Grad} f \geq \operatorname{Grad} g$. Der Beweis erfolgt nun per Induktion nach $\operatorname{Grad} f$.

Sei $\operatorname{Grad} f = 0$. Da $\operatorname{Grad} f \geq \operatorname{Grad} g$, ist auch $\operatorname{Grad} g = 0$. Folglich ist $f = \lambda_0$ und $g = \mu_0$ mit λ_0, $\mu_0 \in K$ und $\lambda_0 \neq 0$, $\mu_0 \neq 0$, und damit ist

$$f = \lambda_0 = (\lambda_0\mu_0^{-1})\mu_0 = (\lambda_0\mu_0^{-1})g + 0 \; .$$

Sei nun $m \geq 1$ und nehme an, dass die Aussage über die Existenz von t und r für alle Polynome f mit $\operatorname{Grad} f < m$ richtig ist.. Sei $f \in K[x]$ mit $\operatorname{Grad} f = m$. Dann ist $f = \lambda_0 + \lambda_1 x + \cdots + \lambda_m x^m$, $g = \mu_0 + \mu_1 x + \cdots + \mu_n x^n$ mit $\lambda_m \neq 0$, $\mu_n \neq 0$ und $m \geq n$. Setze $f_1 = f - (\lambda_m\mu_n^{-1}x^{m-n})g$, wobei $x^0 = 1$; es gilt

$$\begin{aligned} f_1 &= (\lambda_0 + \cdots + \lambda_m x^m) - (\lambda_m\mu_n^{-1})(\mu_0 x^{m-n} + \mu_0 x^{m-n+1} + \cdots + \mu_n x^m) \\ &= (\lambda_0 + \cdots + \lambda_{m-1}x^{m-1}) - (\lambda_m\mu_n^{-1})(\mu_0 x^{m-n} + \cdots + \mu_{n-1}x^{m-1}) \end{aligned}$$

und damit ist entweder $f_1 = 0$ oder $\operatorname{Grad} f_1 \leq m - 1$. Nach Induktionsannahme gibt es also Polynome t_1, $r \in K[x]$ mit $f_1 = t_1 g + r$ und entweder $r = 0$ oder $\operatorname{Grad} r < \operatorname{Grad} g$. Setze $t = t_1 + \lambda_m\mu_n^{-1}x^{m-n}$; dann ist

$$\begin{aligned} f = f_1 + (\lambda_m\mu_n^{-1}x^{m-n})g &= t_1 g + r + (\lambda_m\mu_n^{-1}x^{m-n})g \\ &= (t_1 + \lambda_m\mu_n^{-1}x^{m-n})g + r = tg + r \; . \end{aligned}$$

Eindeutigkeit von t und r: Nehme an, dass es auch Polynome t', r' mit $f = t'g+r'$ und entweder $r' = 0$ oder $\operatorname{Grad} r' < \operatorname{Grad} g$ gibt. Dann ist entweder $r' = r$ oder $\operatorname{Grad}(r' - r) < \operatorname{Grad} g$. Ferner ist

$$0 = f - f = (tg + r) - (t'g + r') = (t - t')g + (r - r')$$

und daraus folgt, dass $(t - t')g = r' - r$. Also ist $t = t'$ (und damit auch $r = r'$), sonst wäre nach Lemma 14.8

$$\operatorname{Grad}(r' - r) = \operatorname{Grad}(t - t')g = \operatorname{Grad}(t - t') + \operatorname{Grad} g \geq \operatorname{Grad} g \; . \; \square$$

Satz 14.3 *Sei U ein Ideal von $K[x]$ mit $U \neq \{0\}$. Dann gibt es ein eindeutiges normiertes Polynom $g \in U$, so dass $U = (g)$. Insbesondere ist jedes Ideal von $K[x]$ ein Hauptideal, (da $\{0\} = (0)$).*

Beweis Wähle ein Polynom $g' \in U \setminus \{0\}$, so dass $\operatorname{Grad} g' \leq \operatorname{Grad} h$ für alle $h \in U \setminus \{0\}$ und setzen $g = \lambda^{-1}g'$, wobei λ der Leitkoeffizient von g' ist. Dann ist g normiert, $g \in U$ und $\operatorname{Grad} g = \operatorname{Grad} g' \leq \operatorname{Grad} h$ für alle $h \in U \setminus \{0\}$. Nun ist $fg \in U$ für alle $f \in K[x]$, da U ein Ideal ist, und damit ist $(g) \subset U$. Sei andererseits $f \in U$; nach Satz 14.2 gibt es Polynome t, $r \in K[x]$ mit $f = tg + r$ und entweder $r = 0$ oder $\operatorname{Grad} r < \operatorname{Grad} g$. Aber $r = f - tg \in U$, und daraus ergibt sich, dass $r = 0$ (sonst wäre $\operatorname{Grad} r < \operatorname{Grad} g \leq \operatorname{Grad} r$). Also ist $f = tg \in (g)$, d.h. $U \subset (g)$ und folglich ist $U = (g)$.

Eindeutigkeit von g: Sei h ein normiertes Polynom mit $U = (h)$. Dann gibt es u, $v \in K[x]$ mit $g = uh$ und $h = vg$, und daraus folgt, dass $\operatorname{Grad} h \leq \operatorname{Grad} g$ und $\operatorname{Grad} g \leq \operatorname{Grad} h$ und also ist $\operatorname{Grad} h = \operatorname{Grad} g$. Aber $h - g \in U$, und entweder $h = g$ oder $\operatorname{Grad}(h - g) < \operatorname{Grad} g$, (da h und g beide normierte Polynome vom gleichen Grad sind). Damit ist $h = g$. □

Ist $f \in K[x]$ ein Polynom, so heißt $\lambda \in K$ *Nullstelle* von f, wenn $f(\lambda) = 0$, (d.h., wenn $\Phi_\lambda(f) = 0$).

Satz 14.4 *Sei $f \in K[x]$ ein Polynom mit $f \neq 0$ und $\operatorname{Grad} f \geq 1$ und sei λ eine Nullstelle von f. Dann gibt es ein eindeutiges Polynom $g \in K[x]$, so dass $f = (\lambda - x)g$. Ferner ist $g \neq 0$ und $\operatorname{Grad} g = (\operatorname{Grad} f) - 1$.*

Beweis Nach Satz 14.2 gibt es Polynome g, $r \in K[x]$ mit $f = (\lambda - x)g + r$ und entweder $r = 0$ oder $\operatorname{Grad} r < \operatorname{Grad}(\lambda - x) = 1$ Dies bedeutet also, dass $r = \mu$ für ein $\mu \in K$. Da aber Φ_λ ein Homomorphismus ist, ist

$$\begin{aligned} 0 = f(\lambda) = \Phi_\lambda((\lambda - x)g + \mu) &= \Phi_\lambda(\lambda - x)\Phi_\lambda(g) + \Phi_\lambda(\mu) \\ &= (\lambda - \lambda)\Phi_\lambda(g) + \mu = 0\Phi_\lambda(g) + \mu = \mu \,, \end{aligned}$$

d.h., $\mu = 0$ und damit ist $f = (\lambda - x)g$. Es ist klar, dass $g \neq 0$ und nach Lemma 14.8 ist $\operatorname{Grad} g = (\operatorname{Grad} f) - 1$. Gilt auch $f = (\lambda - x)h$, so ist nach Lemma 14.9 (2) $h = g$, da $\lambda - x \neq 0$. □

Lemma 14.10 *Seien $\lambda_1, \ldots, \lambda_n \in K$ (mit $n \geq 1$) und seien f, $g \in K[x]$ mit $\operatorname{Grad} g \geq 1$ und $(\lambda_1 - x) \cdots (\lambda_n - x) = fg$. Dann ist λ_k eine Nullstelle von g für mindestens ein k.*

Beweis Sei $m = \operatorname{Grad} f = n - \operatorname{Grad} g \leq n-1$, und nehme an, dass $g(\lambda_k) \neq 0$ für jedes $k = 1, \ldots, m$. Da $g(\lambda_1) \neq 0$ und $f(\lambda_1)g(\lambda_1) = (fg)(\lambda_1) = 0$, ist $f(\lambda_k) = 0$ und folglich gibt es nach Satz 14.4 $f_1 \in K[x]$ mit $f = (\lambda_1 - x)f_1$. Nach Lemma 4.9 (2) ist dann $(\lambda_2 - x)\cdots(\lambda_n - x) = f_1 g$; insbesondere ist $f_2(\lambda_2) = 0$. Nach wiederholter Anwendung dieses Verfahren gibt es $f_1, \ldots, f_m \in K[x]$, so dass $(\lambda_{k+1} - x)\cdots(\lambda_n - x) = f_k g$ für $k = 1, \ldots, m$. Aber dann ist $\operatorname{Grad} f_m = 0$, folglich ist $(\lambda_{m+1} - x)\cdots(\lambda_n - x) = \lambda g$ für ein $\lambda \in K \setminus \{0\}$ und damit ist $g(\lambda_k) = 0$ für jedes $k = m+1, \ldots, n$. □

Viele der für einen Körper eingeführten Definitionen machen auch für einen Ring einen Sinn. Im folgenden sei R ein Ring.

Sei $n \in \mathbb{N}$; mit R^n wird die Menge aller n-Tupel von Elementen aus R bezeichnet. Ein Element von R^n hat also die Form $(a_1, \ldots, a_n)$ mit $a_1, \ldots, a_n$ Elementen aus R. Das Element $(0, \ldots, 0) \in R^n$ wird mit 0 bezeichnet.

Eine Addition auf R^n wird erklärt durch

$$(a_1, \ldots, a_n) + (b_1, \ldots, b_n) = (a_1 + b_1, \ldots, a_n + b_n)$$

und Multiplikation eines Elements von R^n links mit einem Element von R durch

$$a(b_1, \ldots, b_n) = (ab_1, \ldots, ab_n) \; .$$

Seien $m, n \geq 1$; eine $m \times n$ *Matrix* über R ist eine Anordnung von mn Elementen von R nach folgendem Schema

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Die Menge der $m \times n$ Matrizen über R wird mit $\mathrm{M}(m \times n, R)$ bezeichnet. Die Schreibweise $A = (a_{ij})_{1 \leq i \leq m,\, 1 \leq j \leq n}$ und $A = (a_{ij})$ werden weiter verwendet. Ferner werden Zeilen und Spalten wie im Fall eines Körpers definiert; die Zeilen von A werden als Elemente von R^n und die Spalten als Elemente von R^m betrachtet. Das Element

$$\begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \in \mathrm{M}(m \times n, R)$$

wird mit 0 bezeichnet.

Sei $A = (a_{ij}) \in \mathrm{M}(\ell \times m, R)$ eine $\ell \times m$ Matrix und $B = (a_{jk}) \in \mathrm{M}(m \times n, R)$ eine $m \times n$ Matrix über R. Für alle $i = 1, \ldots, \ell$, $k = 1, \ldots, n$ sei

$$c_{ik} = \sum_{j=1}^{m} a_{ij} b_{jk} \; ;$$

es gibt also eine $\ell \times n$ Matrix $C = (c_{ik}) \in \mathrm{M}(\ell \times n, R)$ über R. Diese Matrix C heißt das *Produkt* von A und B und wird mit AB bezeichnet.

Satz 14.5 *Seien $A \in \mathrm{M}(\ell \times m, R)$, $B \in \mathrm{M}(m \times n, R)$ und $C \in \mathrm{M}(n \times p, R)$. Dann gilt $(AB)C = A(BC)$. (Die Matrizenmultiplikation ist also assoziativ.)*

Beweis Seien $1 \leq \lambda \leq \ell$, $1 \leq \rho \leq p$. Dann ist $\sum_{\nu=1}^{n}\big(\sum_{\mu=1}^{m} a_{\lambda\mu}b_{\mu\nu}\big)c_{\nu\rho}$ das λ,ρ-te Element in $(AB)C$ und $\sum_{\mu=1}^{m} a_{\lambda\mu}\big(\sum_{\nu=1}^{n} b_{\mu\nu}c_{\nu\rho}\big)$ ist das λ,ρ-te Element in $A(BC)$. Da aber die Multiplikation in R assoziativ ist, ist

$$\sum_{\mu=1}^{m} a_{\lambda\mu}\Big(\sum_{\nu=1}^{n} b_{\mu\nu}c_{\nu\rho}\Big) = \sum_{\nu=1}^{n}\Big(\sum_{\mu=1}^{m} a_{\lambda\mu}b_{\mu\nu}\Big)\, c_{\nu\rho}\ .$$

Daraus folgt, dass $(AB)C = A(BC)$. □

Seien $A_1 \in \mathrm{M}(n_0 \times n_1, R)$, $A_2 \in \mathrm{M}(n_1 \times n_2, R)$, …, $A_m \in \mathrm{M}(n_{m-1} \times n_m, R)$. Mit Hilfe von Satz 14.5 kann man durch Induktion zeigen, dass das 'Produkt' der Matrizen $A_1, \ldots, A_m$ unabhängig von der Reihenfolge der Multiplikationen ist; dieses 'Produkt' wird mit $A_1 \cdots A_m$ bezeichnet.

Sei $n \geq 1$; nach Satz 14.5 ist $\mathrm{M}(n \times n, R)$ (mit komponentenweiser Addition und Matrizenmultiplikation) ein Ring. (Man sieht leicht, die restlichen Bedingungen in der Definition eines Rings sind erfüllt.) Ist $n \geq 2$ und ist die Multiplikation in R nicht trivial (es gibt also a, $b \in R$ mit $ab \neq 0$), so ist der Ring $\mathrm{M}(n \times n, R)$ nicht kommutativ.

Im folgenden sei R ein Ring mit 1 und sei $E_n \in \mathrm{M}(n \times n, R)$ folgende Matrix:

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}\ .$$

Die Matrix E_n heißt wieder *Einheitsmatrix*. Man sieht leicht, dass $AE_n = A$ und $E_nB = B$ für alle $A \in \mathrm{M}(m \times n, R)$ und alle $B \in \mathrm{M}(n \times p, R)$. Insbesondere ist $\mathrm{M}(n \times n, R)$ ein Ring mit Einselement E_n.

Lemma 14.11 *Sei $A \in \mathrm{M}(n \times n, R)$. Dann gibt es höchstens eine Matrix A', so dass $AA' = A'A = E_n$.*

Beweis Seien B, $C \in \mathrm{M}(n \times n, R)$ mit $AB = BA = E_n = AC = CA$. Dann gilt

$$B = BE_n = B(AC) = (BA)C = E_nC = C\ .\ \square$$

Eine Matrix $A \in \mathrm{M}(n \times n, R)$ heißt *invertierbar*, wenn es $A' \in \mathrm{M}(n \times n, R)$ gibt, so dass $AA' = A'A = E_n$. Nach Lemma 14.11 ist diese Matrix A' eindeutig und sie wird mit A^{-1} bezeichnet, d.h., A^{-1} ist die eindeutige Matrix mit

$$AA^{-1} = A^{-1}A = E_n \ .$$

Lemma 14.12 *(1) Die Einheitsmatrix E_n ist invertierbar und $E_n^{-1} = E_n$.*

(2) Ist $A \in \mathrm{M}(n \times n, R)$ invertierbar, so ist A^{-1} invertierbar und $(A^{-1})^{-1} = A$.

(3) Sind A, $B \in \mathrm{M}(n \times n, R)$ invertierbar, so ist das Produkt AB invertierbar und $(AB)^{-1} = B^{-1}A^{-1}$.

Beweis (1) Dies ist klar, da $E_nE_n = E_nE_n = E_n$.

(2) Dies ist auch klar, da $A^{-1}A = AA^{-1} = E_n$.

(3) Dies folgt aus der Assoziativität der Matrizenmultiplikation, da

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AE_nA^{-1} = AA^{-1} = E_n$$

und $(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}E_nB = B^{-1}B = E_n$. □

15 Determinante einer Matrix

Im Folgenden sei R ein kommutativer Ring. Ist $a \in R$ und $\varepsilon \in \{-1, 1\}$, dann sei εa das Element von R, das definiert ist durch

$$\varepsilon a = \begin{cases} a & \text{falls } \varepsilon = 1 \ , \\ -a & \text{falls } \varepsilon = -1 \ . \end{cases}$$

Sind $\varepsilon, \varepsilon' \in \{-1, 1\}$ und $a, b \in R$, dann ist $(\varepsilon\varepsilon')a = \varepsilon(\varepsilon' a)$, da $-(-a) = a$, und $\varepsilon(ab) = (\varepsilon a)b$, da $-(ab) = (-a)b$. Ferner ist $\varepsilon(a + b) = (\varepsilon a) + (\varepsilon b)$, da $-(a + b) = (-a) + (-b)$.

Ist $A = (a_{i,j}) \in \mathrm{M}(n \times n, R)$ eine $n \times n$ Matrix über R, so heißt das Element

$$\sum_{\sigma \in S_n} \mathrm{sign}(\sigma) a_{1\,\sigma(1)} a_{2\,\sigma(2)} \cdots a_{n\,\sigma(n)}$$

die *Determinante* von A und wird mit $\det A$ bezeichnet. Insbesondere ist

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{22}$$

und für $n = 3$ ist

$$a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31} + a_{13}a_{21}a_{32}$$

die Determinante von A.

Eine Matrix $A = (a_{ij}) \in \mathrm{M}(n \times n, R)$ heißt *obere Dreiecksmatrix*, wenn $a_{ij} = 0$ für alle $i > j$.

Lemma 15.1 *Sei $A = (a_{ij}) \in \mathrm{M}(n \times n, R)$ eine obere Dreiecksmatrix. Dann ist*

$$\det A = a_{11}a_{22} \cdots a_{nn} \ .$$

Beweis Sei $\sigma \in S_n$ eine Permutation mit $\sigma \neq e$, wobei e die Permutation ist mit $e(i) = i$ für alle i. Dann gibt es $1 \leq i \leq n$ mit $\sigma(i) < i$ (warum?) und folglich ist $a_{i\sigma(i)} = 0$. Damit ist nach Lemma 14.2 $a_{1\,\sigma(1)} a_{2\,\sigma(2)} \cdots a_{n\,\sigma(n)} = 0$ und daraus ergibt sich, dass $\det A = a_{11}a_{22} \cdots a_{nn}$. □

Seien $v_1, \ldots, v_n$ Elemente aus R^n; bezeichne dann mit $[v_1, \ldots, v_n]$ die Matrix, die $v_1, \ldots, v_n$ als Zeilen hat.

Satz 15.1 *Seien $v_1, \ldots, v_n \in R^n$ und $1 \leq k \leq n$.*

(1) Ist $v_k = bu + cw$ mit $u, w \in R^n$ und $b, c \in R$, so ist

$$\begin{aligned} &\det[v_1, \ldots, v_{k-1}, v_k, v_{k+1}, \ldots, v_n] \\ &\qquad = b \det[v_1, \ldots, v_{k-1}, u, v_{k+1}, \ldots, v_n] + c \det[v_1, \ldots, v_{k-1}, w, v_{k+1}, \ldots, v_n] \ . \end{aligned}$$

(2) Ist $v_k = v_\ell$ für ein ℓ mit $k \neq \ell$, so ist $\det[v_1, \ldots, v_n] = 0$.

Beweis Sei $[v_1, \ldots, v_n] = (a_{ij})$.

(1) Sei $[v_1, \ldots, v_{k-1}, u, v_{k+1}, \ldots, v_n] = (a^u_{ij})$, $[v_1, \ldots, v_{k-1}, w, v_{k+1}, \ldots, v_n] = (a^w_{ij})$. Dann ist

$$\begin{aligned}
&\det[v_1, \ldots, v_{k-1}, v_k, v_{k+1}, \ldots, v_n] \\
&\quad = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\,\sigma(1)} \cdots a_{(k-1)\,\sigma(k-1)} a_{k\,\sigma(k)} a_{(k+1)\,\sigma(k+1)} \cdots a_{n\,\sigma(n)} \\
&\quad = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\,\sigma(1)} \cdots a_{(k-1)\,\sigma(k-1)} (b a^u_{k\,\sigma(k)} + c a^w_{k\,\sigma(k)}) a_{(k+1)\,\sigma(k+1)} \cdots a_{n\,\sigma(n)} \\
&\quad = b \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a^u_{1\,\sigma(1)} \cdots a^u_{(k-1)\sigma(k-1)} a^u_{k\,\sigma(k)} a^u_{(k+1)\,\sigma(k+1)} \cdots a^u_{n\,\sigma(n)} \\
&\qquad\qquad + c \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a^w_{1\,\sigma(1)} \cdots a^w_{(k-1)\sigma(k-1)} a^w_{k\,\sigma(k)} a^w_{(k+1)\,\sigma(k+1)} \cdots a^w_{n\,\sigma(n)} \\
&\quad = b \det[v_1, \ldots, v_{k-1}, u, v_{k+1}, \ldots, v_n] + c \det[v_1, \ldots, v_{k-1}, w, v_{k+1}, \ldots, v_n] \,.
\end{aligned}$$

(2) Es gibt $\ell \neq k$, so dass $v_\ell = v_k$, d.h., $a_{kj} = a_{\ell j}$ für alle $j = 1, \ldots, n$. Sei $\tau \in S_n$ die Transposition, die die Elemente k und ℓ vertauscht und sei $T_\tau : S_n \to S_n$ die durch $T_\tau(\sigma) = \sigma\tau$ definierte Abbildung. Da $\sigma\tau\tau = \sigma$, ist $T_\tau(T_\tau(\sigma)) = \sigma$ für jedes $\sigma \in S$ und nach Satz 13.10 ist

$$\operatorname{sign}(T_\tau(\sigma)) = \operatorname{sign}(\sigma\tau) = \operatorname{sign}(\sigma) \operatorname{sign}(\tau) = -\operatorname{sign}(\sigma)$$

für alle $\sigma \in S_n$. Sei $A_n = \{\sigma \in S_n : \operatorname{sign}(\sigma) = 1\}$; also ist $T_\tau(A_n) \subset S_n \setminus A_n$ und $T_\tau(S_n \setminus A_n) \subset A_n$. Damit ist $T_\tau(A_n) = S_n \setminus A_n$ und jede Permutation in $S_n \setminus A_n$ hat eine eindeutige Darstellung der Form $\sigma\tau$ mit $\sigma \in A_n$. Ferner gilt

$$a_{1\,\sigma\tau(1)} \cdots a_{n\,\sigma\tau(n)} = a_{1\,\sigma(1)} \cdots a_{n\,\sigma(n)}$$

für alle $\sigma \in S_n$, da $a_{i\,\sigma\tau(i)} = a_{i\,\sigma(i)}$, falls $i \notin \{k, \ell\}$, $a_{k\,\sigma\tau(k)} = a_{k\,\sigma(\ell)} = a_{\ell\,\sigma(\ell)}$ und $a_{\ell\,\sigma\tau(\ell)} = a_{\ell\,\sigma(k)} = a_{k\,\sigma(k)}$. Daraus folgt, dass

$$\begin{aligned}
\det A &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\,\sigma(1)} \cdots a_{n\,\sigma(n)} \\
&= \sum_{\sigma \in A_n} \operatorname{sign}(\sigma) a_{1\,\sigma(1)} \cdots a_{n\,\sigma(n)} + \sum_{\sigma \in A_n} \operatorname{sign}(\sigma\tau) a_{1\,\sigma\tau(1)} \cdots a_{n\,\sigma\tau(n)} \\
&= \sum_{\sigma \in A_n} \operatorname{sign}(\sigma) a_{1\,\sigma(1)} \cdots a_{n\,\sigma(n)} + \operatorname{sign}(\tau) \sum_{\sigma \in A_n} \operatorname{sign}(\sigma) a_{1\,\sigma(1)} \cdots a_{n\,\sigma(n)} = 0 \,,
\end{aligned}$$

da $\operatorname{sign}(\tau) = -1$. □

Satz 15.2 *Seien $v_1, \ldots, v_n \in R^n$; dann gilt*

$$\det[v_{\pi(1)}, \ldots, v_{\pi(n)}] = \operatorname{sign}(\pi) \det[v_1, \ldots, v_n]$$

für alle $\pi \in S_n$.

Beweis Sei $A = [v_1, \ldots, v_n]$ mit $A = (a_{ij})$ und $B = [v_{\pi(1)}, \ldots, v_{\pi(n)}]$ mit $B = (b_{ij})$, also gilt $b_{ij} = a_{\pi(i)\,j}$ für alle $1 \leq i,\, j \leq n$. Sei $T_\pi : S_n \to S_n$ die durch $T_\pi(\sigma) = \sigma\pi$ definierte Abbildung. Dann ist T_π injektiv, (da $\sigma = (\sigma\pi)\pi^{-1} = (\sigma'\pi)\pi^{-1} = \sigma'$, wenn $\sigma\pi = \sigma'\pi$), und damit eine Bijektion. Folglich gibt es zu jedem $\sigma \in S_n$ eine eindeutige Permutation $\sigma' \in S_n$, so dass $\sigma = \sigma'\pi$. Ferner ist

$$a_{\pi(1)\,\sigma\pi(1)} \cdots a_{\pi(n)\,\sigma\pi(n)} = a_{1\,\sigma(1)} \cdots a_{n\,\sigma(n)}$$

für alle $\sigma \in S_n$, und daraus ergibt sich, dass

$$\begin{aligned}
\det B &= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) b_{1\,\sigma(1)} \cdots b_{n\,\sigma(n)} = \sum_{\sigma\in S_n} \operatorname{sign}(\sigma\pi) b_{1\,\sigma\pi(1)} \cdots b_{n\,\sigma\pi(n)} \\
&= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma\pi) a_{\pi(1)\,\sigma\pi(1)} \cdots a_{\pi(n)\,\sigma\pi(n)} = \sum_{\sigma\in S_n} \operatorname{sign}(\sigma\pi) a_{1\,\sigma(1)} \cdots a_{n\,\sigma(n)} \\
&= \operatorname{sign}(\pi) \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) a_{1\,\sigma(1)} \cdots a_{n\,\sigma(n)} = \operatorname{sign}(\pi) \det A \,. \quad \square
\end{aligned}$$

Satz 15.3 *Für alle* $A,\, B \in \mathrm{M}(n \times n, R)$ *ist*

$$\det AB = (\det A)(\det B) \,.$$

Beweis Sei $A = (a_{ij})$ und seien $v_1, \ldots, v_n$ die Zeilen von B. Nach der Definition für die Multiplikation von Matrizen ist $AB = [w_1, \ldots, w_n]$, wobei

$$w_i = a_{i1} v_1 + \cdots + a_{in} v_n$$

für $i = 1, \ldots, n$. Daraus folgt nach Satz 15.1 (1), dass

$$\begin{aligned}
\det AB = \det\, [w_1, \ldots, w_n] &= \det[a_{11} v_1 + \cdots + a_{1n} v_n, w_2, \ldots, w_n] \\
&= \sum_{j_1=1}^{n} a_{1j_1} \det[v_{j_1}, w_2, \ldots, w_n] \\
&= \sum_{j_1=1}^{n} a_{1j_1} \det[v_{j_1}, a_{21} v_1 + \cdots + a_{2n} v_n, w_3, \ldots, w_n] \\
&\quad \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} a_{1j_1} a_{2j_2} \det[v_{j_1}, v_{j_2}, w_3, \ldots, w_n] \\
&\qquad\qquad \vdots \\
&= \sum_{j_1=1}^{n} \cdots \sum_{j_n=1}^{n} a_{1j_1} \cdots a_{nj_n} \det[v_{j_1}, \ldots, v_{j_n}] \,.
\end{aligned}$$

Ferner ist nach Satz 15.1 (2) $\det[v_{j_1}, \ldots, v_{j_n}] = 0$, falls die Indizes $j_1, \ldots, j_n$ nicht alle verschieden sind. Aber die Indizes $j_1, \ldots, j_n$ sind alle verschieden genau dann,

wenn es eine Permutation $\sigma \in S_n$ gibt, so dass $j_k = \sigma(k)$ für $k = 1, \ldots, n$. Daraus ergibt sich nach Satz 15.2, dass

$$\begin{aligned}\sum_{j_1=1}^{n} \cdots \sum_{j_n=1}^{n} & a_{1j_1} \cdots a_{nj_n} \det[v_{j_1}, \ldots, v_{j_n}] \\ &= \sum_{\sigma \in S_n} a_{1\,\sigma(1)} \cdots a_{n\,\sigma(n)} \det[v_{\sigma(1)}, \ldots, v_{\sigma(n)}] \\ &= \sum_{\sigma \in S_n} a_{1\,\sigma(1)} \cdots a_{n\,\sigma(n)} \operatorname{sign}(\sigma) \det[v_1, \ldots, v_n] \\ &= \sum_{\sigma \in S_n} a_{1\,\sigma(1)} \cdots a_{n\,\sigma(n)} \operatorname{sign}(\sigma) \det B \\ &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\,\sigma(1)} \cdots a_{n\,\sigma(n)} \det B = (\det A)(\det B)\ ,\end{aligned}$$

und damit ist $\det AB = (\det A)(\det B)$. □

Für $A = (a_{ij}) \in \mathrm{M}(n \times n, R)$ sei $A^t \in \mathrm{M}(n \times n, R)$ die durch $A^t = (a^t_{ij})$ definierte Matrix, wobei $a^t_{ij} = a_{ji}$ für alle $1 \leq i, j \leq n$; A^t heißt die *zu A transponierte Matrix*.

Satz 15.4 *Für jede Matrix $A \in \mathrm{M}(n \times n, R)$ ist $\det A^t = \det A$.*

Beweis Sei $A = (a_{ij})$. Für jedes $\sigma \in S_n$ gilt

$$a_{\sigma(1)\,1} \cdots a_{\sigma(n)\,n} = a_{\sigma(\sigma^{-1}(1))\,\sigma^{-1}(1)} \cdots a_{\sigma(\sigma^{-1}(n))\,\sigma^{-1}(n)} = a_{1\,\sigma^{-1}(1)} \cdots a_{n\,\sigma^{-1}(n)}$$

und $\operatorname{sign}(\sigma^{-1}) = \operatorname{sign}(\sigma)$ (da $\operatorname{sign}(\sigma^{-1}) \operatorname{sign}(\sigma) = \operatorname{sign}(\sigma^{-1}\sigma) = \operatorname{sign}(e) = 1$). Ferner gibt es zu jedem $\sigma \in S_n$ eine eindeutige Permutation $\tau \in S_n$, so dass $\sigma = \tau^{-1}$. Daraus ergibt sich, dass

$$\begin{aligned}\det A &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\,\sigma(1)} \cdots a_{n\,\sigma(n)} = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma^{-1}) a_{1\,\sigma^{-1}(1)} \cdots a_{n\,\sigma^{-1}(n)} \\ &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{\sigma(1)\,1} \cdots a_{\sigma(n)\,n} = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a^t_{1\,\sigma(1)} \cdots a^t_{n\,\sigma(n)} = \det A^t\ .\ \square\end{aligned}$$

Seien $w_1, \ldots, w_n$ Elemente aus R^n; dann bezeichnet $[w_1, \ldots, w_n]'$ die Matrix, die $w_1, \ldots, w_n$ als Spalten hat.

Satz 15.5 *Seien $w_1, \ldots, w_n \in R^n$ und $1 \leq k \leq n$.*

(1) Für jedes $a \in R$ ist

$$\det[w_1, \ldots, w_{k-1}, aw_k, w_{k+1}, \ldots, w_n]' = a \det[w_1, \ldots, w_n]'\ .$$

(2) Sei $\hat{w}_k \in R^n$; dann ist

$$\begin{aligned}&\det[w_1,\dots,w_{k-1},w_k+\hat{w}_k,w_{k+1},\dots,w_n]'\\&\quad=\det[w_1,\dots,w_{k-1},w_k,w_{k+1},\dots,w_n]'+\det[w_1,\dots,w_{k-1},\hat{w}_k,w_{k+1},\dots,w_n]'\,.\end{aligned}$$

(3) Ist $w_k = w_\ell$ für ein ℓ mit $k \neq \ell$, so ist $\det[w_1,\dots,w_n]' = 0$.

(4) Für alle $\pi \in S_n$ gilt

$$\det[w_{\pi(1)},\dots,w_{\pi(n)}]' = \operatorname{sign}(\pi)\det[w_1,\dots,w_n]'\,.$$

Beweis Da $[w_1,\dots,w_n]'$ die zu $[w_1,\dots,w_n]$ transponierte Matrix ist, folgt dies unmittelbar aus Satz 15.1, Satz 15.2 und Satz 15.4. □

Satz 15.6 *Seien m, $\ell \geq 1$ mit $m+\ell = n$ und sei $A \in \mathrm{M}(n\times n, R)$ eine Matrix mit der Gestalt*

$$\begin{pmatrix} A_1 & B\\ 0 & A_2\end{pmatrix}\,,$$

wobei $A_1 \in \mathrm{M}(m\times m, R)$, $A_2 \in \mathrm{M}(\ell\times\ell, R)$ und $B \in \mathrm{M}(m\times\ell, R)$. Dann gilt

$$\det A = (\det A_1)(\det A_2)\,.$$

Beweis Sei $A = (a_{ij})$; also ist $a_{ij} = 0$ für alle $m < i \leq n$, $1 \leq j \leq m$. Ferner sei auch $A_1 = (a^1_{ij})$, $A_2 = (a^2_{ij})$; es gilt $a^1_{ij} = a_{ij}$ für alle $1 \leq i, j \leq m$ und $a^2_{ij} = a_{(m+i)\,(m+j)}$ für alle $1 \leq i, j \leq \ell$.

Sei $\sigma \in S_n$ mit $a_{1\,\sigma(1)}a_{2\,\sigma(2)}\cdots a_{n\,\sigma(n)} \neq 0$; dann ist $\sigma(i) > m$ für alle $i > m$ und damit auch $\sigma(i) \leq m$ für alle $i \leq m$, da σ eine Bijektion ist. Folglich ist $\sigma \in P_{m,n}$ wobei

$$\begin{aligned}P_{m,n} = \{\sigma \in S_n : \sigma(\{1,\dots,m\}) &\subset \{1,\dots,m\}\\ &\text{und}\ \ \sigma(\{m+1,\dots,n\}) \subset \{m+1,\dots,n\}\}\,.\end{aligned}$$

Sei $\psi : S_m \times S_\ell \to S_m$ die Abbildung mit

$$\psi(\tau,\varrho)(k) = \begin{cases}\tau(k) & \text{falls } k \in \{1,\dots,m\}\,,\\ \varrho(k-m)+m & \text{falls } k \in \{m+1,\dots,n\}\,.\end{cases}$$

Dann ist ψ injektiv mit $\psi(S_m \times S_\ell) = P_{m,n}$. Ferner gilt

$$\operatorname{sign}(\psi(\tau,\varrho)) = \operatorname{sign}(\tau)\operatorname{sign}(\varrho)$$

für alle $\tau \in S_m$, $\varrho \in S_\ell$. Daraus folgt, dass

$$\begin{aligned}
\det A &= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) a_{1\,\sigma(1)} a_{2\,\sigma(2)} \cdots a_{n\,\sigma(n)} \\
&= \sum_{\sigma\in P_{m,n}} \operatorname{sign}(\sigma) a_{1\,\sigma(1)} a_{2\,\sigma(2)} \cdots a_{n\,\sigma(n)} \\
&= \sum_{(\tau,\varrho)\in S_m\times S_\ell} \operatorname{sign}(\psi(\tau,\varrho)) a_{1\,\psi(\tau,\varrho)(1)} a_{2\,\psi(\tau,\varrho)(2)} \cdots a_{n\,\psi(\tau,\varrho)(n)} \\
&= \sum_{\tau\in S_m} \sum_{\varrho\in S_\ell} \operatorname{sign}(\tau)\operatorname{sign}(\varrho) a^1_{1\,\tau(1)} \cdots a^1_{m\,\tau(m)} a^2_{1\,\varrho(1)} \cdots a^2_{\ell\,\varrho(\ell)} \\
&= \sum_{\tau\in S_m} \operatorname{sign}(\tau)\operatorname{sign}(\varrho) a^1_{1\,\tau(1)} \cdots a^1_{m\,\tau(m)} \sum_{\varrho\in S_\ell} a^2_{1\,\varrho(1)} \cdots a^2_{\ell\,\varrho(\ell)} \\
&= (\det A_1)(\det A_2)\ . \ \square
\end{aligned}$$

Sei R' ein weiterer kommutativer Ring und $\varphi : R' \to R$ ein Homomorphismus. Dann kann eine Abbildung $\hat\varphi : \mathrm{M}(n\times n, R') \to \mathrm{M}(n\times n, R)$ definiert werden durch $\hat\varphi(A) = (\varphi(a_{ij}))$, wobei $A = (a_{ij})$.

Lemma 15.2 *Die Abbildung $\hat\varphi$ ist ein Homomorphismus und*

$$\det \hat\varphi(A) = \varphi(\det A)$$

für alle $A \in \mathrm{M}(n\times n, R')$.

Beweis Es ist klar, dass $\hat\varphi(A+B) = \hat\varphi(A) + \hat\varphi(B)$ für alle A, $B \in \mathrm{M}(n\times n, R)$ und dass $\hat\varphi(E_n) = E_n$. Seien A, $B \in \mathrm{M}(n\times n, R)$ mit $A = (a_{ij})$, $B = (b_{ij})$; setze $AB = (c_{ij})$. Dann gilt

$$\varphi(c_{ik}) = \varphi\Big(\sum_{j=1}^n a_{ij} b_{jk}\Big) = \sum_{j=1}^n \varphi(a_{ij})\varphi(b_{jk})$$

und daraus folgt, dass $\hat\varphi(C) = \hat\varphi(A)\hat\varphi(B)$. Dies zeigt, dass $\hat\varphi$ ein Homomorphismus ist. Sei nun $A = (a_{ij}) \in \mathrm{M}(n\times n, R')$. Dann gilt

$$\begin{aligned}
\det \hat\varphi(A) &= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) \varphi(a_{1\,\sigma(1)}) \cdots \varphi(a_{n\,\sigma(n)}) \\
&= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) \varphi(a_{1\,\sigma(1)} \cdots a_{n\,\sigma(n)}) = \sum_{\sigma\in S_n} \varphi(\operatorname{sign}(\sigma) a_{1\,\sigma(1)} \cdots\ a_{n\,\sigma(n)}) \\
&= \varphi\Big(\sum_{\sigma\in S_n} \operatorname{sign}(\sigma) a_{1\,\sigma(1)} \cdots a_{n\,\sigma(n)}\Big) = \varphi(\det A)\ . \ \square
\end{aligned}$$

Satz 15.7 *Sei $P = (p_{ij}) \in \mathrm{M}(n \times n, R[x])$ eine Matrix von Polynomen und sei $a \in R$. Für jedes $1 \leq i, j \leq n$ sei $a_{ij} = p_{ij}(a)$ der Wert des Polynoms p_{ij} an der Stelle a. Dann ist der Wert des Polynoms $\det P$ an der Stelle a gleich der Determinanten der Matrix $A = (a_{ij})$, d.h., es gilt $(\det P)(a) = \det A$.*

Beweis Dies folgt unmittelbar aus Lemma 15.2 mit $R' = R[x]$ und mit φ dem Einsetzhomomorphismus $\Phi_a : R[x] \to R$. □

Satz 15.8 *Sei R ein kommutativer Ring mit 1. Dann gilt*

$$(\det A)(\det A^{-1}) = 1$$

für jede invertierbare Matrix $A \in \mathrm{M}(n \times n, R)$. Insbesondere ist dann $\det A \neq 0$.

Beweis Es gilt $AA^{-1} = E_n$ und damit ist nach Lemma 15.1 und Satz 15.3

$$(\det A)(\det A^{-1}) = \det E_n = 1 \cdots 1 = 1 \; .$$

Daraus folgt nach Lemma 14.2, dass $\det A \neq 0$, da $0 \neq 1$. □

Ist R ein kommutativer Ring mit 1, dann heißen A, $B \in \mathrm{M}(n \times n, R)$ *ähnlich*, wenn es eine invertierbare Matrix $P \in \mathrm{M}(n \times n, R)$ gibt, so dass $B = P^{-1}AP$.

Satz 15.9 *Ähnliche Matrizen haben die gleiche Determinante: Sind die Matrizen A, $B \in \mathrm{M}(n \times n, R)$ ähnlich, so ist $\det B = \det A$.*

Beweis Es gibt eine invertierbare Matrix $P \in \mathrm{M}(n \times n, R)$, so dass $B = P^{-1}AP$, und daraus folgt nach Satz 15.3, dass

$$\begin{aligned} \det B = \det(P^{-1}AP) &= (\det P^{-1})(\det A)\ (\det P) = (\det P^{-1})(\det P)(\det A) \\ &= (\det(P^{-1} \det P))(\det A) = (\det E_n)(\det A) = \det A \; . \end{aligned}$$ □

Im Folgenden sie K ein Körper.

Satz 15.10 *Eine Matrix $A \in \mathrm{M}(n \times n, K)$ ist invertierbar genau dann, wenn $\det A \neq 0$.*

Beweis Nach Satz 15.8 braucht man hier nur zu zeigen, dass $\det A = 0$, wenn A nicht invertierbar ist. Außerdem kann man annehmen, dass $n \geq 2$, (da der Fall $n = 1$ trivial richtig ist). Sei also A nicht invertierbar und seien $v_1, \ldots, v_n$ die Zeilen von A. Nach Satz 6.3 sind $v_1, \ldots, v_n$ linear abhängig und folglich gibt es

nach Satz 3.4 (2) ein k mit $1 \leq k \leq n$, so dass $v_k \in \mathrm{L}(v_1, \ldots, v_{k-1}, v_{k+1}, \ldots, v_n)$, d.h., es gibt $\lambda_1, \ldots, \lambda_{k-1}, \lambda_{k+1}, \ldots, \lambda_n$ aus K, so dass

$$v_k = \lambda_1 v_1 + \cdots + \lambda_{k-1} v_{k-1} + \lambda_{k+1} v_{k+1} + \cdots + \lambda_n v_n \; .$$

Nach Satz 15.1 ist dann

$$\det A = \det[v_1, \ldots, v_n] = \sum_{j \neq k} \lambda_j \det[v_1, \ldots, v_{k-1}, v_j, v_{k+1}, \ldots, v_n] = 0 \; . \quad \square$$

Sei $A \in \mathrm{M}(n \times n, K)$ eine invertierbare Matrix. Nach Satz 6.3 ist das zu (A, b) gehörige lineare Gleichungssystem eindeutig lösbar für jedes $b \in K^n$. Das nächste Ergebnis liefert eine Formel für diese eindeutige Lösung.

Satz 15.11 (Cramersche Regel) *Sei $b = (b_1, \ldots, b_n) \in K^n$ und für jedes j sei $A_j^b \in \mathrm{M}(n \times n, K)$ die Matrix, die man erhält, wenn die j-te Spalte von A durch b ersetzt wird. Sei $(\lambda_1, \ldots, \lambda_n) \in K^n$ die eindeutige Lösung des zu (A, b) gehörigen linearen Gleichungssystems. Dann ist für $j = 1, \ldots, n$*

$$\lambda_j = (\det A)^{-1} (\det A_j^b)$$

(und nach Satz 15.10 ist hier $\det A \neq 0$).

Beweis Seien $w_1, \ldots, w_n$ die Spalten von A; also ist $\lambda_1 w_1 + \cdots + \lambda_n w_n = b$ und daraus ergibt sich nach Satz 15.5, dass

$$\begin{aligned} \det A_j^b &= \det[w_1, \ldots, w_{j-1}, b, w_{j+1}, \ldots, w_n]' \\ &= \det[w_1, \ldots, w_{j-1}, \lambda_1 w_1 + \cdots + \lambda_n w_n, w_{j+1}, \ldots, w_n]' \\ &= \sum_{i=1}^{n} \lambda_i \det[w_1, \ldots, w_{j-1}, w_i, w_{j+1}, \ldots, w_n]' \\ &= \lambda_j \det[w_1, \ldots, w_{j-1}, w_j, w_{j+1}, \ldots, w_n]' = \lambda_j \det A \; , \end{aligned}$$

d.h., $\lambda_j = (\det A)^{-1} (\det A_j^b)$. $\square$

Seien a, b, c, d, u, $v, \in K$ mit $ad \neq bc$. Nach Satz 15.11 ist die eindeutige Lösung (x', y') des linearen Gleichungssystems

$$\begin{aligned} ax + by &= u \\ cx + dy &= v \end{aligned}$$

gegeben durch

$$x' = \left(\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} \right)^{-1} \det \begin{pmatrix} u & b \\ v & d \end{pmatrix} = (ad - cb)^{-1} (ud - vb) \; ,$$

$$y' = \left(\det \begin{pmatrix} a & b \\ c & d \end{pmatrix}\right)^{-1} \det \begin{pmatrix} a & u \\ c & v \end{pmatrix} = (ad - cb)^{-1}(av - cu) \; .$$

Im Allgemeinen ist die Cramersche Regel eine sehr unpraktische Methode zur Berechnung der Lösung linearer Gleichungssysteme. Viel weniger aufwendig ist der Gaußsche Algorithmus.

Nun werden einige Hilfssätze zur Berechnung der Deteminanten betrachtet. Im Folgenden sei R wieder ein kommutativer Ring.

Sei $A \in \mathrm{M}(n \times n, R)$ und $1 \leq i, j \leq n$; mit A_{ij} wird die aus A durch Weglassen der i-ten Zeile und der j-ten Spalte entstehende $(n-1) \times (n-1)$ Matrix.

Satz 15.12 *Sei $A = (a_{ij}) \in \mathrm{M}(n \times n, R)$. Für jedes $k = 1, \ldots, n$ gilt dann*

$$\det A = \sum_{i=1}^{n} (-1)^{i+k} a_{ik} \det A_{ik} \; .$$

Beweis Sei k fest (mit $1 \leq k \leq n$) und für $i = 1, \ldots, n$ sei

$$S_n^i = \{\sigma \in S_n : \sigma(i) = k\} \; .$$

Die Teilmengen $S_n^1, \ldots, S_n^n$ bilden also eine Partition der Menge S_n; damit ist

$$\begin{aligned} \det A &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1\,\sigma(1)} \cdots a_{n\,\sigma(n)} = \sum_{i=1}^{n} \sum_{\sigma \in S_n^i} \operatorname{sign}(\sigma) a_{1\,\sigma(1)} \cdots a_{n\,\sigma(n)} \\ &= \sum_{i=1}^{n} a_{ik} \sum_{\sigma \in S_n^i} \operatorname{sign}(\sigma) a_{1\,\sigma(1)} \cdots a_{i-1\,\sigma(i-1)} a_{i+1\,\sigma(i+1)} \cdots a_{n\,\sigma(n)} \; . \end{aligned}$$

Für $i = 1, \ldots, n$ sei $T_i : S_n^i \to S_{n-1}$ die Abbildung, die definiert ist durch

$$T_i(\sigma)(\ell) = \begin{cases} \sigma(\ell) & \text{falls } \ell < i \text{ und } \sigma(\ell) < k \; , \\ \sigma(\ell) - 1 & \text{falls } \ell < i \text{ und } \sigma(\ell) > k \; , \\ \sigma(\ell + 1) & \text{falls } \ell \geq i \text{ und } \sigma(\ell) < k \; , \\ \sigma(\ell + 1) - 1 & \text{falls } \ell \geq i \text{ und } \sigma(\ell) > k \; . \end{cases}$$

Dann ist T_i eine Bijektion und für jedes $\sigma \in S_n^i$ ist

$$a_{1\,\sigma(1)} \cdots a_{i-1\,\sigma(i-1)} a_{i+1\,\sigma(i+1)} \cdots a_{n\,\sigma(n)} = b_{1\,\tau(1)} \cdots b_{n-1\,\tau(n-1)} \; ,$$

wobei $\tau = T_i(\sigma)$ und $A_{ik} = (b_{\ell j})_{1 \leq \ell \leq n-1,\, 1 \leq j \leq n-1}$. Folglich ist

$$\begin{aligned} &\sum_{\sigma \in S_n^i} \operatorname{sign}(\sigma) a_{1\,\sigma(1)} \cdots a_{i-1\,\sigma(i-1)} a_{i+1\,\sigma(i+1)} \cdots a_{n\,\sigma(n)} \\ &\qquad = (-1)^{i+k} \sum_{\tau \in S_{n-1}} \operatorname{sign}(\tau) b_{1\,\tau(1)} \cdots b_{n-1\,\tau(n-1)} = (-1)^{i+k} \det A_{ik} \; , \end{aligned}$$

da $\operatorname{sign}(T_i(\sigma)) = (-1)^{i+k}\operatorname{sign}(\sigma)$ (warum?). Daraus ergibt sich,

$$\begin{aligned}\det A &= \sum_{i=1}^{n} a_{ik} \sum_{\sigma\in S_n^i} \operatorname{sign}(\sigma) a_{1\,\sigma(1)}\cdots a_{i-1\,\sigma(i-1)} a_{i+1\,\sigma(i+1)}\cdots a_{n\,\sigma(n)} \\ &= \sum_{i=1}^{n} (-1)^{i+k} a_{ik} \det A_{ik}\ . \quad \square\end{aligned}$$

Die Berechnungsformel in Satz 15.12 heißt *Entwicklung der Determinante nach der k-ten Spalte*. Es gibt natürlich auch eine entsprechende Berechnungsformel für die *Entwicklung der Determinante nach der k-ten Zeile*:

Satz 15.13 *Sei $A = (a_{ij}) \in \mathrm{M}(n\times n, R)$. Für jedes $k = 1, \ldots, n$ gilt dann*

$$\det A = \sum_{j=1}^{n} (-1)^{k+j} a_{kj} \det A_{kj}\ .$$

Beweis Nach Satz 15.12 und Satz 15.4 ist

$$\begin{aligned}\det A = \det A^t &= \sum_{i=1}^{n} (-1)^{i+k} a_{ik}^t \det (A^t)_{ik} \\ &= \sum_{i=1}^{n} (-1)^{i+k} a_{ki} \det (A_{ki})^t = \sum_{j=1}^{n} (-1)^{k+j} a_{kj} \det A_{kj}\ . \quad \square\end{aligned}$$

Lemma 15.3 *Sei $A = (a_{ij}) \in \mathrm{M}(n\times n, R)$ und sei $\breve{A} = (\breve{a}_{ij})$ das Element von $\mathrm{M}(n\times n, R)$, das definiert ist durch*

$$\breve{a}_{ij} = (-1)^{i+j} \det A_{ji}$$

für alle $1 \le i,\, j \le n$. Dann gilt $A\breve{A} = \breve{A}A = (\det A) E_n$.

Beweis Ist $B \in \mathrm{M}(n\times n, R)$ und $1 \le i,\, k \le n$, so wird mit $B^{k\leftarrow i}$ die $n\times n$ Matrix bezeichnet, die man erhält, wenn die k-te Zeile von B durch die i-te Zeile von B ersetzt wird. Also ist $B^{k\leftarrow k} = B$ und $\det B^{k\leftarrow i} = 0$, falls $k \ne i$, da dann $B^{k\leftarrow i}$ zwei gleiche Zeilen hat. Man beachte, dass $B_{kj}^{k\leftarrow i} = B_{kj}$ für alle $1 \le i,\, k,\, j \le n$.

Setze $\breve{A}A = (c_{ik})$; für $1 \le i,\, k \le n$ ist dann Satz 15.13

$$\begin{aligned}c_{ik} &= \sum_{j=1}^{n} a_{ij}\breve{a}_{jk} = \sum_{j=1}^{n} a_{ij}(-1)^{j+k} \det A_{kj} = \sum_{j=1}^{n} (-1)^{k+j} a_{ij} \det A_{kj} \\ &= \sum_{j=1}^{n} (-1)^{k+j} a_{kj}^{k\leftarrow i} \det A_{kj}^{k\leftarrow i} = \det A^{k\leftarrow i} = \begin{cases} \det A & \text{falls } i = k\ , \\ 0 & \text{falls } i \ne k\ , \end{cases}\end{aligned}$$

und folglich ist $\breve{A}A = (\det A) E_n$. Das gleiche Argument (mit Spalten statt Zeilen und Satz 15.12 statt Satz 15.13) zeigt, dass $A\breve{A} = (\det A) E_n$. $\square$

Satz 15.14 *Eine Matrix $A \in \mathrm{M}(n \times n, R)$ ist genau dann invertierbar, wenn das Element* $\det A$ *von R invertierbar ist. (Ein Element $a \in R$ heißt invertierbar, wenn es $b \in R$ mit $ab = 1$ gibt. In diesem Fall ist b eindeutig und man schreibt $b = a^{-1}$.)*

Beweis Ist $\det A$ invertierbar, so ist nach Lemma 15.3 die Matrix A invertierbar mit $A^{-1} = (\det A)^{-1}\breve{A}$. Ist umgekehrt A invertierbar, so ist $AA^{-1} = E_n$ und damit ist $1 = \det AA^{-1} = (\det A)(\det A^{-1})$. Folglich ist $\det A$ invertierbar mit $(\det A)^{-1} = \det A^{-1}$. □

Im Folgenden sei K ein Körper. Für die praktische Berechnung der Determinante einer Matrix aus $\mathrm{M}(n\times n, K)$ sind im Allgemeinen die in Satz 15.12 und Satz 15.13 gegebenen Berechnungsformeln zu aufwendig. Es ist meistens besser, das folgende Ergebnis in Zusammenhang mit dem Gaußschen Algorithmus zu benutzen .

Satz 15.15 *Seien A, $B \in \mathrm{M}(n \times n, K)$.*

(1) Erhält man B aus A durch Addition eines Vielfachen einer Zeile zu einer anderen Zeile, so ist $\det B = \det A$.

(2) Erhält man B aus A durch Vertauschen zweier Zeilen, so ist $\det B = -\det A$.

Beweis (1) Seien $v_1, \ldots, v_n$ die Zeilen von A und $v'_1, \ldots, v'_n$ die Zeilen von B. Dann gibt es $\lambda \in K$ und $1 \leq j,\, k \leq n$ mit $j \neq k$, so dass

$$v'_i = \begin{cases} v_i & \text{falls } i \neq j\ , \\ v_j + \lambda v_k & \text{falls } i = j\ , \end{cases}$$

und daraus folgt nach Satz 15.1, dass

$$\begin{aligned} \det B &= \det[v'_1, \ldots, v'_n] = \det[v_1, \ldots, v_{j-1}, v_j + \lambda v_k, v_{j+1}, \ldots, v_n] \\ &= \det[v_1, \ldots, v_{j-1}, v_j, v_{j+1}, \ldots, v_n] + \lambda \det[v_1, \ldots, v_{j-1}, v_k, v_{j+1}, \ldots, v_n] \\ &= \det A + \lambda 0 = \det A\ . \end{aligned}$$

(2) Dies folgt unmittelbar aus Satz 15.2. □

Sei $A \in \mathrm{M}(n \times n, K)$; nach dem Gaußschen Algorithmus (Satz 1.5) läßt sich A durch eine Folge von elementaren Zeilenumformungen in eine Matrix B mit Zeilen-Stufen-Form überführen und nach Satz 15.3 ist dann $\det A = (-1)^m \det B$, wobei m die Anzahl der verwendeten Vertauschungen ist. Aber eine Matrix in Zeilen-Stufen-Form ist insbesondere eine obere Dreiecksmatrix und daher kann ihre Determinante direkt mit Hilfe von Lemma 15.1 berechnet werden.

16 Das charakteristische Polynom

Im Folgenden sei R ein kommutativer Ring mit 1.

Sei $A = (a_{ij}) \in \mathrm{M}(n \times n, R)$; da $R \subset R[x]$, kann A auch als Element von $\mathrm{M}(n \times n, R[x])$ betrachtet werden. Das Element

$$\chi_A = \det(A - xE_n)$$

von $R[x]$ heißt *das charakteristische Polynom von* A. Es gilt also $\chi_A = \det P$, wobei $P = (p_{ij}) \in \mathrm{M}(n \times n, R[x])$ die Matrix ist mit

$$p_{ij} = \begin{cases} a_{ii} - x & \text{falls } i = j \ , \\ a_{ij} & \text{falls } i \neq j \ . \end{cases}$$

Insbesondere ist für $n = 2$

$$\begin{aligned} \chi_A &= \det \begin{pmatrix} a_{11} - x & a_{12} \\ a_{21} & a_{22} - x \end{pmatrix} \\ &= (a_{11} - x)(a_{22} - x) - a_{12}a_{21} = (a_{11}a_{22} - a_{12}a_{21}) - (a_{11} + a_{22})x + x^2 \end{aligned}$$

(und man merke, dass $a_{11}a_{22} - a_{12}a_{21} = \det A$), und für $n = 3$ ist

$$\begin{aligned} \chi_A &= \det \begin{pmatrix} a_{11} - x & a_{12} & a_{13} \\ a_{21} & a_{22} - x & a_{23} \\ a_{31} & a_{32} & a_{33} - x \end{pmatrix} \\ &= d - (a_{11}a_{22} + a_{22}a_{33} + a_{11}a_{33} - a_{23}a_{32} - a_{12}a_{21} - a_{13}a_{31})x \\ &\quad + (a_{11} + a_{22} + a_{33})x^2 - x^3 \ , \end{aligned}$$

wobei

$$d = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31} + a_{13}a_{21}a_{32}$$

und damit ist $d = \det A$.

Satz 16.1 *Für jedes $A \in \mathrm{M}(n \times n, R)$ ist χ_A ein Polynom vom Grad n. Ist ferner $A = (a_{ij}) \in \mathrm{M}(n \times n, R)$ und*

$$\chi_A = a_0 + a_1 x + \cdots + a_n x^n \ ,$$

so ist $a_0 = \det A$, $a_{n-1} = (-1)^{n-1}(a_{11} + a_{22} + \cdots + a_{nn})$ und $a_n = (-1)^n$.

Beweis Es gilt $\chi_A = \det P$, wobei $P = (p_{ij}) \in \mathrm{M}(n \times n, R[x])$ wie oben ist. Sei

$$p_d = p_{11} \cdots p_{nn} = (a_{11} - x) \cdots (a_{nn} - x) \ ;$$

dann ist p_d ein Polynom vom Grad n und

$$\chi_A = \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) p_{1\,\sigma(1)} \cdots p_{n\,\sigma(n)} = p_d + \sum_{\sigma\in S'_n} \operatorname{sign}(\sigma) p_{1\,\sigma(1)} \cdots p_{n\,\sigma(n)} \ ,$$

wobei $S'_n = S_n \setminus \{e\}$. Sei $\sigma \in S'_n$; dann gibt es mindestens ein i mit $\sigma(i) \neq i$ (und damit $p_{i\,\sigma(i)} = a_{i\,\sigma(i)}$) und folglich ist $p_{1\,\sigma(1)} \cdots p_{n\,\sigma(n)}$ entweder 0 oder ein Polynom vom Grad n_σ mit $n_\sigma < n$. Daraus ergibt sich, dass χ_A ein Polynom vom Grad n ist. Der Beweis für den Rest ist eine Übung. □

Satz 16.2 *Ähnliche Matrizen haben das gleiche charakteristische Polynom: Sind A, $B \in \mathrm{M}(n \times n, R)$ ähnlich, so ist $\chi_B = \chi_A$.*

Beweis Es gibt eine invertierbare Matrix $P \in \mathrm{M}(n \times n, R)$, so dass $B = P^{-1}AP$. Betrachte die Matrizen A, B, P, P^{-1} und E_n als Elemente von $\mathrm{M}(n \times n\, R[x])$; dann gilt

$$B - xE_n = P^{-1}AP - x(P^{-1}E_nP) = P^{-1}AP - P^{-1}(xE_n)P = P^{-1}(A - xE_n)P \ ,$$

da $P^{-1}P = E_n$ auch in $\mathrm{M}(n \times n, R[x])$ gilt, und daraus folgt nach Satz 15.3, dass

$$\begin{aligned}
\chi_B &= \det(B - xE_n) = \det(P^{-1}(A - xE_n)P) \\
&= (\det P^{-1})(\det(A - xE_n))(\det P) = (\det P^{-1})(\det P)(\det(A - xE_n)) \\
&= (\det(P^{-1}P))(\det(A - xE_n)) = (\det E_n)(\det(A - xE_n)) \\
&= \det(A - xE_n) = \chi_A \ . \ \square
\end{aligned}$$

Im Folgenden sei K ein Körper.

Satz 16.3 *Sei $A \in \mathrm{M}(n \times n, K)$; dann sind die Eigenwerte von A genau die Nullstellen des charakteristischen Polynoms χ_A. Mit anderen Worten ist $\lambda \in K$ ein Eigenwert von A genau dann, wenn $\chi_A(\lambda) = 0$.*

Beweis Nach Satz 6.3 ist $\lambda \in K$ ein Eigenwert von A genau dann, wenn die Matrix $A - \lambda E_n \in \mathrm{M}(n \times n, K)$ nicht invertierbar ist. Folglich ist nach Satz 15.10 λ genau dann ein Eigenwert, wenn $\det(A - \lambda E_n) = 0$. Aber nach Satz 15.7 ist

$$\det(A - \lambda E_n) = (\det(A - xE_n))(\lambda) = \chi_A(\lambda) \ .$$

Damit sind die Eigenwerte von A genau die Nullstellen von χ_A. □

Beispiel Sei $A \in \mathrm{M}(2 \times 2, \mathbb{R})$ folgende Matrix:

$$\begin{pmatrix} 2 & 4 \\ 1 & -1 \end{pmatrix}$$

Dann ist $\chi_A = -6 - x + x^2 = (3 - x)(-2 - x)$ und damit sind 3 und -2 die Eigenwerte von A.

Es gibt numerische Verfahren für die Berechnung der Nullstellen von reellen und komplexen Polynomen. Diese Verfahren kann also zur Berechnung der Eigenwerte reeller und komplexer Matrizen angewendet werden.

Sei $A \in \mathrm{M}(n \times n, K)$; nun gibt es das folgende Verfahren, das prüft, ob die Matrix A diagonalisierbar ist:

(1) Man bestimme die (verschiedenen) Nullstellen $\lambda_1, \ldots, \lambda_m$ von χ_A. Nach Satz 16.3 sind dann $\lambda_1, \ldots, \lambda_m$ die Eigenwerte von A.

(2) Für jedes $j = 1, \ldots, m$ ist nach Lemma 10.4

$$\mathrm{E}(A, \lambda_j) = \mathrm{Lös}(A - \lambda_j E_n, 0)$$

und also kann mit Hilfe von Satz 16.1 und des Gaußschen Algorithmus eine Basis $(u_1^j, \ldots, u_{p_j}^j)$ von $\mathrm{E}(A, \lambda_j)$ konstruiert werden.

(3) Nach Satz 10.6 ist A diagonalisierbar genau dann, wenn $p_1 + \cdots + p_m = n$.

(4) Nehme an, dass A diagonalisierbar ist und setze

$$(w_1, \ldots, w_n) = (u_1^1, \ldots, u_{p_1}^1, u_1^2, \ldots, u_{p_2}^2, \ldots, u_1^m, \ldots, u_{p_m}^m) \; .$$

Sei $Q \in \mathrm{M}(n \times n, K)$ die Matrix, die $w_1, \ldots, w_n$ als Spalten hat. Nach Satz 10.2 ist Q invertierbar und $Q^{-1}AQ = D$, wobei D die Diagonalmatrix ist mit den Einträgen $\lambda_1, \ldots, \lambda_1, \ldots, \lambda_m, \ldots, \lambda_m$ auf der Diagonalen, wobei für jedes j der Eigenwert λ_j genau p_j-mal vorkommt.

Ein Körper K heißt *algebraisch abgeschlossen*, wenn jedes Polynom $f \in K[x]$ mit $f \neq 0$ und $\operatorname{Grad} f \geq 1$ eine Nullstelle besitzt. Ist K algebraisch abgeschlossen, so gilt nach Satz 16.3, dass jede Matrix $A \in \mathrm{M}(n \times n, K)$ einen Eigenwert besitzt.

Satz 16.4 (Fundamentalsatz der Algebra) $\mathbb{C}$ *ist algebraisch abgeschlossen.*

Beweis Trotz seines Namens ist der Fundamentalsatz der Algebra eigentlich ein Satz der Analysis und einen Beweis dafür wird man erst in Analysis II oder in einer Vorlesung über Funktionentheorie kennenlernen. □

Nach Satz 16.3 und Satz 16.4 besitzt jede komplexe Matrix einen Eigenwert.

Der Körper $\mathbb{R}$ der reellen Zahlen ist nicht algebraisch abgeschlossen. Zum Beispiel besitzt das Polynom $1 + x^2 \in \mathbb{R}[x]$ keine Nullstelle. (Dagegen sind i und $-i$ die Nullstellen des komplexen Polynoms $1 + x^2 \in \mathbb{C}[x]$.) Damit besitzt

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

(als reelle Matrix) keinen Eigenwert, da $\chi_A = 1 + x^2$.

Die Körper $\mathbb{F}_p$, p eine Primzahl, sind ebenfalls nicht algebraisch abgeschlossen.

Sei nun V ein endlichdimensionaler Vektorraum über K mit $\dim V = n \geq 1$.

Sei $f : V \to V$ ein Endomorphismus von V. Seien α und β Basen von V und sei A bzw. B die Matrix von f bezüglich der Basis α bzw. bezüglich der Basis β. Nach Satz 8.8 sind die Matrizen A und B ähnlich und daraus folgt nach Satz 16.2, dass $\chi_B = \chi_A$. Man kann also *das charakteristische Polynom χ_f von f* durch $\chi_f = \chi_A$ definieren, wobei A die Matrix von f bezüglich irgendeiner Basis von V ist.

Satz 16.5 *Sei $f : V \to V$ ein Endomorphismus. Dann sind die Eigenwerte von f genau die Nullstellen von χ_f. Mit anderen Worten ist $\lambda \in K$ ein Eigenwert von f genau dann, wenn $\chi_f(\lambda) = 0$.*

Beweis Dies folgt unmittelbar aus Satz 16.3 und Satz 10.3 (1). (Ist A die Matrix von f bezüglich einer Basis α von V und ist $\lambda \in K$, so behauptet Satz 10.3 (1), dass λ genau dann ein Eigenwert von f ist, wenn es ein Eigenwert von A ist.) □

Erinnerung: Ist $f : V \to V$ ein Endomorphismus, so heißt ein Untervektorraum U von V *f-invariant*, wenn $f(U) \subset U$.

Satz 16.6 *Sei $f : V \to V$ ein Endomorphismus und sei U ein f-invarianter Untervektorraum von V mit $\dim U \geq 1$. Dann gibt ein Polynom $p \in K[x]$, so dass $\chi_f = \chi_{f_{|U}} p$. (Das Polynom $\chi_{f_{|U}}$ ist also ein Teiler von χ_f.)*

Beweis Dies ist trivial richtig, wenn $U = V$ und also kann man annehmen, dass $m = \dim U < n$; setze $\ell = n - m$. Nach dem Basisergänzungssatz gibt es eine Basis $(v_1, \ldots, v_n)$ von V, so dass $(v_1, \ldots, v_m)$ eine Basis von U ist. Sei $A \in \mathrm{M}(n \times n, K)$ die Matrix von f bezüglich $(v_1, \ldots, v_n)$. Dann hat A die Gestalt

$$\begin{pmatrix} A' & C \\ 0 & B \end{pmatrix}$$

wobei $A' \in \mathrm{M}(m \times m, K)$, $B \in \mathrm{M}(\ell \times \ell, K)$ und $C \in \mathrm{M}(m \times \ell, K)$, und A' ist die Matrix von $f_{|U}$ bezüglich $(u_1, \ldots, u_m)$. Nun gilt

$$A - xE_n = \begin{pmatrix} A' - xE_m & C \\ 0 & B - xE_\ell \end{pmatrix}$$

und daraus ergibt sich nach Satz 15.6, dass

$$\chi_f = \det(A - xE_n) = \det(A' - xE_m) \det(B - xE_\ell) = \chi_{f_{|U}} \chi_B \ . \quad \square$$

Erinnerung: Eine Matrix $A \in \mathrm{M}(n \times n, K)$ heißt *trigonalisierbar*, wenn sie ähnlich zu einer oberen Dreiecksmatrix ist und ein Endomorphismus $f : V \to V$ heißt *trigonalisierbar*, wenn es eine f-invariante Fahne gibt. (Eine Folge $V_0, V_1, \ldots, V_n$ von Untervektorräumen von V heißt eine *Fahne*, wenn

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

und $\dim V_j = j$ für jedes $j = 0, \ldots, n$, und eine Fahne $V_0, V_1, \ldots, V_n$ heißt *f-invariant*, wenn für jedes j der Untervektorraum V_j f-invariant ist.)

Ist $f : V \to V$ ein Endomorphismus von V und ist $A \in \mathrm{M}(n \times n, K)$ die Matrix von f bezüglich irgendeiner Basis von V, so ist nach Satz 11.1 f trigonalisierbar genau dann, wenn A trigonalisierbar ist.

Sei $f \in K[x]$ ein Polynom mit $f \neq 0$ und $\operatorname{Grad} f \geq 1$; man sagt, dass f *in Linearfaktoren zerfällt*, wenn es Elemente $\lambda_1, \ldots, \lambda_n, \lambda \in K$ gibt, so dass

$$f = \lambda(\lambda_1 - x) \cdots (\lambda_n - x) \; .$$

Dann ist $\lambda \neq 0$, $n = \operatorname{Grad} f$ und $\lambda_1, \ldots, \lambda_n$ sind genau die Nullstellen von f, da

$$f(\mu) = \lambda(\lambda_1 - \mu) \cdots (\lambda_n - \mu) = 0$$

genau dann, wenn $\mu = \lambda_j$ für ein j. Zerfällt f in Linearfaktoren, so sieht man leicht, dass die Darstellung $f = \lambda(\lambda_1 - x) \cdots (\lambda_n - x)$ bis auf die Reihenfolge der Faktoren eindeutig ist.

Satz 16.7 *Ist K algebraisch abgeschlossen, so zerfällt jedes Polynom $f \in K[x]$ mit $f \neq 0$ und $\operatorname{Grad} f \geq 1$ in Linearfaktoren.*

Beweis Sei $f \in K[x]$ mit $f \neq 0$ und $\operatorname{Grad} f \geq 1$; da K algebraisch abgeschlossen ist, besitzt f eine Nullstelle λ und nach Satz 14.4 gibt es dann ein eindeutiges Polynom $g \in K[x]$, so dass $f = (\lambda - x)g$. Da $g \neq 0$ und $\operatorname{Grad} g = (\operatorname{Grad} f) - 1$, erfolgt der Beweis nun per Induktion nach $n = \operatorname{Grad} f$. □

Insbesondere folgt aus Satz 16.4 und Satz 16.7, dass jedes Polynom $f \in \mathbb{C}[x]$ mit $f \neq 0$ und $\operatorname{Grad} f \geq 1$ in Linearfaktoren zerfällt. Dagegen zerfällt zum Beispiel das Polynom $1 + x^2 \in \mathbb{R}[x]$ nicht in Linearfaktoren, da dieses Polynom keine Nullstellen besitzt.

Sei $A \in \mathrm{M}(n \times n, K)$; zerfällt das charakteristische Polynom χ_A in Linearfaktoren, dann gibt es $\lambda_1, \ldots, \lambda_n \in K$, so dass

$$\chi_f = (\lambda_1 - x) \cdots (\lambda_n - x) \; ,$$

da nach Satz 16.1 $(-1)^n$ der Leitkoeffizient von χ_A ist. Das Gleiche gilt auch für χ_f, wenn $f : V \to V$ ein Endomorphismus ist.

Satz 16.8 *(1) Eine Matrix $A \in \mathrm{M}(n \times n, K)$ ist genau dann trigonalisierbar, wenn das charakteristische Polynom χ_A in Linearfaktoren zerfällt.*

(2) Ein Endomorphismus $f : V \to V$ von V ist genau dann trigonalisierbar, wenn das charakteristische Polynom χ_f in Linearfaktoren zerfällt.

Beweis Nehme zuerst an, dass die Matrix $A \in \mathrm{M}(n \times n, K)$ trigonalisierbar ist. Dann ist A ähnlich zu einer oberen Dreiecksmatrix $B = (b_{ij}) \in \mathrm{M}(n \times n, K)$ und nach Satz 16.2 ist $\chi_A = \chi_B$. Aber die Matrix $B - xE_n \in \mathrm{M}(n \times n, K[x])$ ist auch eine obere Dreiecksmatrix mit den Einträgen $b_{11} - x$, $b_{22} - x$, $\ldots$, $b_{nn} - x$ auf der Diagonalen und daraus folgt nach Lemma 15.1, dass

$$\chi_A = \chi_B = \det(B - xE_n) = (b_{11} - x) \cdots (b_{nn} - x) \; .$$

Damit zerfällt das charakteristische Polynom χ_A in Linearfaktoren.

Ist der Endomorphismus $f : V \to V$ trigonalisierbar, dann folgt unmittelbar aus Satz 11.1, dass auch das charakteristische Polynom χ_f in Linearfaktoren zerfällt.

Nehme nun an, dass das charakteristische Polynom χ_f in Linearfaktoren zerfällt, es gibt aso $\lambda_1, \ldots, \lambda_n \in K$, so dass

$$\chi_f = (\lambda_1 - x) \cdots (\lambda_n - x) \; .$$

Sei U ein f-invarianter Untervektorraum von V mit $\dim U = m \geq 1$. Nach Satz 16.6 gibt es ein Polynom $p \in K[x]$, so dass $\chi_f = \chi_{f_{|U}} p$ und daraus folgt nach Lemma 14.10, dass $\chi_{f_{|U}}(\lambda_k) = 0$ für mindestens ein k. Dann ist aber λ_k ein Eigenwert von $f_{|U}$ und damit enthält U einen Eigenvektor von f. Daraus ergibt sich nach Satz 11.2, dass f trigonalisierbar ist.

Hier ist ein anderer Beweis, der Satz 11.2 nicht benutzt: Es wird durch Induktion nach $n = \dim V$ gezeigt, dass f trigonalisierbar ist. Die Aussage ist trivial richtig, wenn $n = 1$, da in diesem Fall jeder Endomorphismus trigonalisierbar ist. (Es gibt nur eine Fahne und sie ist stets invariant.) Sei also $n > 1$ und nehme an, dass g immer dann trigonalisierbar ist, wenn $1 \leq \dim W < n$ und $g : W \to W$ ein Endomorphismus von W ist, für den χ_g in Linearfaktoren zerfällt.

Da χ_f in Linearfaktoren zerfällt, gibt es $\lambda_1, \ldots, \lambda_n \in K$, so dass

$$\chi_f = (\lambda_1 - x) \cdots (\lambda_n - x) \; .$$

Dann ist insbesondere $\chi_f(\lambda_1) = 0$ und daraus folgt nach Satz 16.5, dass λ_1 Eigenwert von f ist. Sei v_1 ein Eigenvektor von f zum Eigenwert λ_1; nach dem Basisergänzungssatz gibt es dann $v_2, \ldots, v_n \in V$, so dass $(v_1, \ldots, v_n)$ eine Basis von V ist. Sei $A = (a_{ij}) \in \mathrm{M}(n \times n, K)$ die Matrix von f bezüglich der Basis

$(v_1, \ldots, v_n)$. Da $f(v_1) = \lambda_1 v_1$, ist $a_{11} = \lambda_1$ und $a_{i1} = 0$ für $i > 1$; A hat also folgende Gestalt:

$$\begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix},$$

wobei $B \in \mathrm{M}((n-1) \times (n-1), K)$. Folglich ist nach Satz 15.6

$$\chi_A = \det(A - xE_n) = (\lambda_1 - x)(\det(B - xE_{n-1}))$$

und nach Lemma 14.9 (2) ist dann $\chi_B = \det(B - xE_{n-1}) = (\lambda_2 - x) \cdots (\lambda_n - x)$, d.h., das charakteristische Polynom χ_B zerfällt in Linearfaktoren.

Sei nun $W = \mathrm{L}(v_2, \ldots, v_n)$; da $(v_2, \ldots, v_n)$ eine Basis von W ist, gibt es nach Lemma 5.11 einen eindeutigen Endomorphismus $g : W \to W$ von W, so dass

$$g(v_j) = a_{2j}v_2 + \cdots + a_{nj}v_n$$

für $j = 2, \ldots, n$, und natürlich ist B die Matrix von g bezüglich $(v_2, \ldots, v_n)$. Insbesondere zerfällt das charakteristische Polynom $\chi_g = \chi_B$ in Linearfaktoren und damit ist nach der Induktionsannahme g trigonalisierbar. Es gibt also eine g-invariante Fahne

$$\{0\} = W_0 \subset W_1 \subset \cdots \subset W_{n-1} = W$$

und nach wiederholter Anwendung des Basisergänzungssatzes gibt es eine Basis $(w_1, \ldots, w_{n-1})$ von W, so dass für jedes $j = 1, \ldots, n-1$ $(w_1, \ldots, w_j)$ eine Basis von W_j ist. Dann ist nach Lemma 3.12 und Satz 4.6 (2) $(v_1, w_1, \ldots, w_{n-1})$ eine Basis von V, da $v_1 \notin \mathrm{L}(v_2, \ldots, v_n) = \mathrm{L}(w_1, \ldots, w_{n-1})$. Setze $V_1 = \mathrm{L}(v_1)$ und für $j = 2, \ldots, n$ sei $V_j = \mathrm{L}(v_1, w_1, \ldots, w_{j-1})$; also ist $V_0 = \{0\}, V_1, \ldots, V_n$ eine Fahne. Für jedes $j = 2, \ldots, n$ ist nun

$$f(v_j) = a_{1j}v_1 + \cdots + a_{nj}v_n = a_{1j}v_1 + g(v_j)$$

und damit ist $(f - g)(v_j) \in \mathrm{L}(v_1)$ für $j = 2, \ldots, n$. Daraus ergibt sich nach Satz 3.2, dass $(f - g)(w) \in \mathrm{L}(v_1)$ für alle $w \in \mathrm{L}(v_2, \ldots, v_n) = W$. Folglich ist

$$f(w_k) \in \mathrm{L}(v_1, g(w_k)) \subset \mathrm{L}(v_1, w_1, \ldots, w_k)$$

für $k = 1, \ldots, n-1$, und daher ist $f(V_j) \subset V_j$ für $j = 1, \ldots, n$, d.h., die Fahne $V_0, V_1, \ldots, V_n$ ist f-invariant. Dies zeigt, dass f trigonalisierbar ist.

Nehme schließlich an, dass das charakteristische Polynom χ_A in Linearfaktoren zerfällt. Da A die Matrix von φ_A bezüglich der kanonischen Basis von K^n ist, ist $\chi_{\varphi_A} = \chi_A$. Damit zerfällt χ_{φ_A} in Linearfaktoren und daraus ergibt sich, dass φ_A trigonalisierbar ist. Folglich gilt nach Satz 11.1, dass A trigonalisierbar ist. □

Satz 16.8 zeigt, dass jede komplexe Matrix trigonalisierbar ist. Genauso ist jeder Endomorphismus eines endlichdimensionalen $\mathbb{C}$-Vektorraumes trigonalisierbar.

Im Beweis für Satz 16.8 wird die entsprechende Aussage für Endomorphismen benutzt, um zu zeigen, dass A trigonalisierbar ist, wenn χ_A in Linearfaktoren zerfällt. Es ist aber möglich, dies direkt zu beweisen: Sei $A_n \in \mathrm{M}(n \times n, K)$ eine Matrix, für die das charakteristische Polynom χ_{A_n} in Linearfaktoren zerfällt. Sei λ_1 ein Eigenwert von A_n und wähle eine Basis $(v_1, \ldots, v_n)$ von K^n, so dass v_1 ein Eigenvektor von A_n zum Eigenwert λ_1 ist. Sei $P_n \in \mathrm{M}(n \times n, K)$ die Matrix mit Spalten $v_1, \ldots, v_n$ (also ist P_n invertierbar). Setze $A'_n = P_n^{-1} A_n P_n$; die Matrix $A'_n = (a'_{ij})$ hat die Gestalt:

$$A'_n = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & & & \\ \vdots & & A_{n-1} & \\ 0 & & & \end{pmatrix}$$

wobei $A_{n-1} \in \mathrm{M}((n-1) \times (n-1), K)$. Dann gilt nach Satz 15.6, dass

$$\chi_{A_n} = \det(A_n - xE_n) = (\lambda_1 - x)\det(A_{n-1} - xE_{n-1}) = (\lambda_1 - x)\chi_{A_{n-1}}$$

und daraus ergibt sich, dass $\chi_{A_{n-1}}$ in Linearfaktoren zerfällt. Nehme an, dass es eine invertierbare Matrix $R_{n-1} = (r_{ij}) \in \mathrm{M}((n-1) \times (n-1), K)$ gibt, so dass $B_{n-1} = R_{n-1}^{-1} A_{n-1} R_{n-1}$ eine obere Dreiecksmatrix ist. Setze

$$Q_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & R_{n-1} & \\ 0 & & & \end{pmatrix} \quad \text{und} \quad B_n = \begin{pmatrix} \lambda_1 & \alpha_1 & \cdots & \alpha_{n-1} \\ 0 & & & \\ \vdots & & B_{n-1} & \\ 0 & & & \end{pmatrix},$$

wobei für jedes $j = 1, \ldots, n-1$

$$\alpha_j = a'_{12} r_{1j} + a'_{13} r_{2j} + \cdots + a'_{1n} r_{n-1\,j}\ .$$

Dann ist Q_n invertierbar und B_n ist eine obere Dreiecksmatrix, und man prüft leicht nach, dass $A'_n = Q_n B_n Q_n^{-1}$. Damit ist $A_n = R_n B_n R_n^{-1}$, wobei $R_n = P_n Q_n$. Dies zeigt, dass die $n \times n$ Matrix A_n trigonalisierbar ist, wenn die $(n-1) \times (n-1)$ Matrix A_{n-1} trigonalisierbar ist. Durch Induktion nach n kann man auf diese Weise beweisen, dass A_n trigonalisierbar ist.

17 Das minimale Polynom

In Folgenden sei R ein kommutativer Ring mit 1.

Lemma 17.1 *Zu jeder Matrix $B \in \mathrm{M}(n \times n, R)$ gibt es einen eindeutigen Ring-Homomorphismus $\Phi_B : R[x] \to \mathrm{M}(n \times n, R)$ mit $\Phi_B(1) = E_n$ und $\Phi_B(x) = B$.*

Beweis Definiere $\Phi_B : R[x] \to \mathrm{M}(n \times n, R)$ durch $\Phi_B(0) = 0$ und

$$\Phi_B(a_0 + a_1 x + \cdots + a_m x^m) = a_0 E_n + a_1 B + \cdots + a_m B^m \ ,$$

falls $m \geq 0$ und $a_m \neq 0$. Dann gilt $\Phi_B(1) = E_n$ und $\Phi_B(B) = B$ und man sieht leicht, dass Φ_B ein Homomorphismus ist, da $A^{p+q} = A^p A^q$ und $x^{p+q} = x^p x^q$ für alle p, $q \geq 0$. Ist $\varphi : R[x] \to \mathrm{M}(n \times n, R)$ ein Homomorphismus mit $\varphi(1) = E_n$ und $\varphi(B) = B$, dann gilt $\varphi(x^m) = B^m$ für alle $m \geq 1$ und damit ist

$$\varphi(a_0 + \cdots + a_m x^m) = a_0 E_n + \cdots + a_m B^m = \Phi_B(a_0 + \cdots + a_m x^m)$$

für alle $a_0, \ldots, a_m \in R$, $m \geq 0$, und folglich ist $\varphi = \Phi_B$. □

Satz 17.1 (Cayley-Hamilton) *Für jedes $A \in \mathrm{M}(n \times n, R)$ gilt $\Phi_A(\chi_A) = 0$.*

Beweis Sei $A_\chi = (A - xE_n)^t = A^t - xE_n \in \mathrm{M}(n \times n, R[x])$. Dann ist

$$\chi_A = \det(A - xE_n) = \det(A - xE_n)^t = \det A_\chi \ ,$$

und daraus ergibt sich nach Lemma 15.3, dass

$$\breve{A}_\chi A_\chi = (\det A_\chi) E_n = \chi_A E_n$$

in $\mathrm{M}(n \times n, R[x])$, d.h., die Matrix $\breve{A}_\chi A_\chi$ ist das χ_A-fache der Einheitsmatrix in $\mathrm{M}(n \times n, R[x])$.

Nach Lemma 15.2 induziert nun der Homomorphismus $\Phi_A : R[x] \to \mathrm{M}(n \times n, R)$ einen Homomorphismus $\hat{\Phi}_A : \mathrm{M}(n \times n, R[x]) \to \mathrm{M}(n \times n, \mathrm{M}(n \times n, R))$. Schreibe $\chi_A(A) = \Phi_A(\chi_A)$. Es gilt dann

$$\hat{\Phi}_A(\chi_A E_n) = \hat{\Phi}_A(\breve{A}_\chi A_\chi) = \hat{\Phi}_A(\breve{A}_\chi)\hat{\Phi}_A(A_\chi)$$

in $\mathrm{M}(n \times n, \mathrm{M}(n \times n, R))$ und $\hat{\Phi}_A(\chi_A E_n) = \chi_A(A) E_n$, d.h., $\hat{\Phi}_A(\chi_A E_n)$ ist das $\chi_A(A)$-fache der Einheitsmatrix in $\mathrm{M}(n \times n, \mathrm{M}(n \times n, R))$.

Definiere $\otimes : \mathrm{M}(n \times n, \mathrm{M}(n \times n, R)) \times (R^n)^n \to (R^n)^n$ durch

$$B \otimes \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \ ,$$

wobei $w_i = b_{i1}v_1 + \cdots + b_{in}v_n$ mit $B = (b_{ij})$ und wobei die Elemente von R^n und $(R^n)^n$ als Spaltenvektoren betrachtet werden. Man prüft leicht nach, dass

$$(CB) \otimes \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = C \otimes \left(B \otimes \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right)$$

für alle B, $C \in \mathrm{M}(n \times n, \mathrm{M}(n \times n, R))$, $v_1, \ldots, v_n \in R^n$. Seien $e_1, \ldots, e_n \in R^n$ die üblichen Einheitsvektoren. Für $B = \hat{\Phi}_A(A^t - xE_n)$ gilt $b_{ij} = a_{ji}E_n - \delta_{ij}A$ und folglich ist

$$\hat{\Phi}_A(A^t - xE_n) \otimes \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} ,$$

da für jedes $i = 1, \ldots, n$

$$\begin{aligned} a_{1i}E_ne_1 + \cdots + a_{ni}E_ne_n &= a_{1i}e_1 + \cdots + a_{ni}e_n \\ &= s_i = Ae_i = 0 + \cdots + Ae_i + \cdots + 0 \end{aligned}$$

mit $s_1, \ldots, s_n \in R^n$ die Spalten von A. Für $B = \hat{\Phi}_A(\chi_A E_n)$ gilt andererseits $b_{ij} = \delta_{ij}\chi_A(A)$ und folglich ist

$$\hat{\Phi}_A(\chi_A E_n) \otimes \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} \chi_A(A)e_1 \\ \vdots \\ \chi_A(A)e_n \end{pmatrix} .$$

Daraus ergibt sich, dass

$$\begin{aligned} \begin{pmatrix} \chi_A(A)e_1 \\ \vdots \\ \chi_A(A)e_n \end{pmatrix} &= \hat{\Phi}_A(\chi_A E_n) \otimes \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \hat{\Phi}_A(\breve{A}_\chi A_\chi) \otimes \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \\ &= \left(\hat{\Phi}_A(\breve{A}_\chi)\hat{\Phi}_A(A^t - xE_n)\right) \otimes \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \\ &= \hat{\Phi}_A(\breve{A}_\chi) \otimes \left(\hat{\Phi}_A(A^t - xE_n) \otimes \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} \right) \\ &= \hat{\Phi}_A(\breve{A}_\chi) \otimes \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} ; \end{aligned}$$

d.h., $\chi_A(A)e_k = 0$ für $k = 1, \ldots, n$. Aber $\chi_A(A)e_k$ ist die k-te Spalte von $\chi_A(A)$, und also ist $\chi_A(A) = 0$. □

Im Folgenden sei nun K ein Körper. Für jede Matrix $A \in \mathrm{M}(n \times n, K)$ sei $\Phi_A : K[x] \to \mathrm{M}(n \times n, K)$ immer noch der eindeutige Ring-Homomorphismus mit $\Phi_A(1) = E_n$ und $\Phi_A(x) = A$.

Satz 17.2 *Sei $A \in \mathrm{M}(n\times n, K)$; dann gibt es ein eindeutiges normiertes Polynom $\mathrm{m}_A \in K[x]$ mit der folgenden Eigenschaft: Für $p \in K[x]$ gilt $\Phi_A(p) = 0$ genau dann, wenn m_A ein Teiler von p ist. (m_A heißt das minimale Polynom von A).*

Beweis Nach Satz 17.1 ist $\Phi_A(\chi_A) = 0$, d.h., $\chi_A \in \operatorname{Kern} \Phi_A$. Aber $\chi_A \neq 0$ und also ist $\operatorname{Kern} \Phi_A \neq \{0\}$. Nach Satz 14.3 gibt es also ein eindeutiges normiertes Polynom $\mathrm{m}_A \in K[x]$, so dass $\operatorname{Kern} \Phi_A = (\mathrm{m}_A)$. Für ein Polynom $p \in K[x]$ gilt daher $\Phi_A(p) = 0$ genau dann, wenn m_A ein Teiler von p ist. □

Nach dem Satz von Cayley-Hamilton teilt m_A das charakteristische Polynom χ_A.

Lemma 17.2 *Seien A, $B \in \mathrm{M}(n \times n, K)$ ähnlich. Dann gilt $\mathrm{m}_A = \mathrm{m}_B$.*

Beweis Es gibt eine invertierbare Matrix $P \in \mathrm{M}(n \times n, K)$, so dass $B = P^{-1}AP$. Dann ist $B^m = P^{-1}A^mP$ für jedes $m \geq 0$ und damit $\Phi_B(p) = P^{-1}\Phi_A(p)P$ für jedes $p \in K[x]$. Insbesondere gilt

$$\Phi_B(\mathrm{m}_A) = P^{-1}\Phi_A(\mathrm{m}_A)P = P^{-1}0P = 0$$

und folglich ist m_B ein Teiler von m_A. Genauso ist m_A ein Teiler von m_B und daher ist $\mathrm{m}_A = \mathrm{m}_B$, da die Polynome m_A und m_B beide normiert sind. □

Satz 17.3 *Sei $A \in \mathrm{M}(n \times n, K)$. Dann sind die Eigenwerte von A genau die Nullstellen von m_A und damit haben m_A und χ_A die gleichen Nullstellen.*

Beweis Da m_A das charakteristische Polynom χ_A teilt, ist jede Nullstelle von m_A auch eine Nullstelle von χ_A. Sei umgekehrt λ eine Nullstelle von χ_A, d.h., λ ist ein Eigenwert von A. Sei $v \in K^n$ ein Eigenvektor von A zum Eigenwert λ und sei $\mathrm{m}_A = a_0 + a_1x + \cdots + a_mx^m$ (und also ist $a_m = 1$). Da $Av = \lambda v$, gilt auch $A^kv = \lambda^k v$ für jedes $k \geq 0$ und damit ist

$$\begin{aligned}(a_0 + a_1\lambda + \cdots + a_m\lambda^m)v& \\ &= a_0v + a_1\lambda v + \cdots + a_m\lambda^m v = a_0E_nv + a_1Av + \cdots + a_mA^mv \\ &= (a_0E_n + a_1A + \cdots + a_mA^m)v = \Phi_A(\mathrm{m}_A)v = 0v = 0\ .\end{aligned}$$

Folglich ist $a_0 + a_1\lambda + \cdots + a_m\lambda^m = 0$, da $v \neq 0$, d.h., λ ist eine Nullstelle von m_A. □

18 Euklidische und unitäre Vektorräume

Sei V ein reeller Vektorraum (d.h. ein Vektorraum über dem Körper $\mathbb{R}$). Eine Abbildung $s : V \times V \to \mathbb{R}$ heißt *Bilinearform*, wenn

$$s(\lambda_1 u_1 + \lambda_2 u_2, v) = \lambda_1 s(u_1, v) + \lambda_2 s(u_2, v) \quad \text{und}$$
$$s(v, \lambda_1 u_1 + \lambda_2 u_2) = \lambda_1 s(v, u_1) + \lambda_2 s(v, u_2)$$

für alle u_1, u_2, $v \in V$ und alle λ_1, $\lambda_2 \in \mathbb{R}$. Eine Bilinearform $s : V \times V \to \mathbb{R}$ heißt *symmetrisch*, wenn $s(u, v) = s(v, u)$ für alle u, $v \in V$. Eine symmetrische Bilinearform $s : V \times V \to \mathbb{R}$ heißt *positiv definit*, wenn $s(v, v) > 0$ für alle $v \in V$ mit $v \neq 0$. Ist s positiv definit, so ist insbesondere $s(v, v) \geq 0$ für alle $v \in V$, da $s(0, 0) = 0$ für jede Bilinearform s. Eine positiv definite symmetrische Bilinearform nennt man *Skalarprodukt*.

Ein Paar $(V, \langle \cdot, \cdot \rangle)$ bestehend aus einem reellen Vektorraum V und einem Skalarprodukt $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ heißt *euklidischer Vektorraum*.

Ist $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum, so wird die durch

$$\|v\| = \sqrt{\langle v, v \rangle}$$

definierte Abbildung $\| \cdot \| : V \to \mathbb{R}^+$ die *Norm* genannt.

Beispiele: (1) Definiere $\cdot : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ durch

$$(x_1, \ldots, x_n) \cdot (y_1, \ldots, y_n) = \sum_{k=1}^{n} x_k y_k \ .$$

Dann ist $(\mathbb{R}^n, \cdot)$ ein euklidischer Vektorraum, der $\mathbb{R}^n$ *mit dem üblichen Skalarprodukt* genannt wird. Es gilt

$$\|x\| = \sqrt{\sum_{j=1}^{n} x_j^2}$$

für alle $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$.

(2) Seien a, $b \in \mathbb{R}$ mit $a < b$ und sei $\mathrm{C}([a, b], \mathbb{R})$ der reelle Vektorraum aller stetigen Abbildungen von $[a, b]$ nach $\mathbb{R}$. Sei $\langle \cdot, \cdot \rangle : \mathrm{C}([a, b], \mathbb{R}) \times \mathrm{C}([a, b], \mathbb{R}) \to \mathbb{R}$ die Abbildung, die definiert ist durch

$$\langle f, g \rangle = \int_a^b f(t) g(t)\, dt \ .$$

Dann ist $(\mathrm{C}([a, b], \mathbb{R}), \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum. (Der Beweis dafür ist eine Übung.)

(3) Sei ℓ_2 der reelle Vektorraum aller quadratisch summierbaren Folgen reeller Zahlen. (Eine Folge $\{x_n\}_{n\geq 0}$ heißt quadratisch summierbar, wenn es ein $N \geq 0$ gibt, so dass $\sum_{n=0}^{m} x_n^2 \leq N$ für alle $m \geq 0$, d.h., wenn die unendliche Reihe $\sum_{n\geq 0} x_n^2$ konvergiert.) Seien $\{x_n\}_{n\geq 0}$, $\{y_n\}_{n\geq 0} \in \ell_2$; da $|x_n y_n| \leq \frac{1}{2}(x_n^2 + y_n^2)$ für jedes $n \geq 0$, konvergiert die unendliche Reihe $\sum_{n\geq 0} x_n y_n$. Folglich kann eine Abbildung $\langle \cdot, \cdot \rangle : \ell_2 \times \ell_2 \to \mathbb{R}$ durch

$$\langle \{x_n\}_{n\geq 0}, \{y_n\}_{n\geq 0} \rangle = \sum_{n\geq 0} x_n y_n$$

definiert werden und dann ist $(\ell_2, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum. (Der Beweis dafür ist eine Übung.)

In den folgenden zwei Sätzen sowie in Lemma 18.1 sei $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum.

Satz 18.1 (Cauchy-Schwarzsche Ungleichung) *Für alle u, $v \in V$ gilt*

$$|\langle u, v \rangle| \leq \sqrt{\langle u, u \rangle \langle v, v \rangle} \, .$$

Beweis Für jedes $\lambda \in \mathbb{R}$ gilt

$$\begin{aligned} 0 &\leq \langle u - \lambda v, u - \lambda v \rangle = \langle u, u - \lambda v \rangle - \lambda \langle v, u - \lambda v \rangle \\ &= \langle u, u \rangle - \lambda \langle u, v \rangle - \lambda \langle v, u \rangle + \lambda^2 \langle v, v \rangle = \langle u, u \rangle - 2\lambda \langle u, v \rangle + \lambda^2 \langle v, v \rangle \, , \end{aligned}$$

d.h. $\langle v, v \rangle \lambda^2 - 2\langle u, v \rangle \lambda + \langle u, u \rangle \geq 0$ für alle $\lambda \in \mathbb{R}$. Daraus ergibt sich, dass

$$\langle u, v \rangle^2 \leq \langle u, u \rangle \langle v, v \rangle$$

und damit $|\langle u, v \rangle| \leq \sqrt{\langle u, u \rangle \langle v, v \rangle}$. (Sind b, $c \in \mathbb{R}$, $a \geq 0$ und $a\lambda^2 - 2b\lambda + c \geq 0$ für alle $\lambda \in \mathbb{R}$, so ist $b^2 \leq ac$.) □

Satz 18.2 *Die Norm $\| \cdot \| : V \to \mathbb{R}^+$ hat folgende Eigenschaften:*

(1) $\|v\| = 0$ genau dann, wenn $v = 0$.

(2) $\|\lambda v\| = |\lambda| \|v\|$ für alle $v \in V$, $\lambda \in \mathbb{R}$.

(3) Dreiecksungleichung $\|u + v\| \leq \|u\| + \|v\|$ für alle u, $v \in V$.

Beweis (1) und (2) sind klar. (3): Nach Satz 18.1 ist

$$\begin{aligned} \|u + v\|^2 &= \langle u + v, u + v \rangle = \langle u, u \rangle + 2\langle u, v \rangle + \langle v, v \rangle \\ &= \|u\|^2 + 2\langle u, v \rangle + \|v\|^2 \leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 = (\|u\| + \|v\|)^2 \end{aligned}$$

und daraus folgt, dass $\|u + v\| \leq \|u\| + \|v\|$. □

Lemma 18.1 (Polarisierungsidentitäten) *Für alle u, $v \in V$ gilt*

$$\langle u, v\rangle = \tfrac{1}{2}(\|u\|^2 + \|v\|^2 - \|u - v\|^2) = \tfrac{1}{4}(\|u + v\|^2 - \|u - v\|^2) \; .$$

Beweis Übung. □

Der Begriff des Skalarproduktes wird nun auch für komplexe Vektorräume eingeführt. Ein komplexer Vektorraum zusammen mit einem Skalarprodukt heißt unitärer Vektorraum.

Sei V ein komplexer Vektorraum (d.h. ein Vektorraum über dem Körper $\mathbb{C}$). Eine Abbildung $s : V \times V \to \mathbb{C}$ heißt *Sesquilinearform*, wenn

$$\begin{aligned} s(\lambda_1 u_1 + \lambda_2 u_2, v) &= \lambda_1 s(u_1, v) + \lambda_2 s(u_2, v) \quad \text{und} \\ s(v, \lambda_1 u_1 + \lambda_2 u_2) &= \overline{\lambda}_1 s(v, u_1) + \overline{\lambda}_2 s(v, u_2) \end{aligned}$$

für alle u_1, u_2, $v \in V$ und alle λ_1, $\lambda_2 \in \mathbb{C}$. Eine Sesquilinearform $s : V \times V \to \mathbb{C}$ heißt *Hermitesche Form*, wenn $s(u, v) = \overline{s(v, u)}$ für alle u, $v \in V$. Ist s eine Hermitesche Form, dann gilt $s(v, v) = \overline{s(v, v)}$ für alle $v \in V$, d.h. $s(v, v) \in \mathbb{R}$ für alle $v \in V$. Eine Hermitesche Form $s : V \times V \to \mathbb{C}$ heißt *positiv definit*, wenn $s(v, v) > 0$ für alle $v \in V$ mit $v \neq 0$. Ist s positiv definit, so ist insbesondere $s(v, v) \geq 0$ für alle $v \in V$, da $s(0, 0) = 0$ für jede Sesquilinearform s. Eine positiv definite Hermitesche Form nennt man *Skalarprodukt*.

Ein Paar $(V, \langle \cdot, \cdot \rangle)$ bestehend aus einem komplexen Vektorraum V und einem Skalarprodukt $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ heißt *unitärer Vektorraum*.

Ist $(V, \langle \cdot, \cdot \rangle)$ ein unitärer Vektorraum, so wird die durch

$$\|v\| = \sqrt{\langle v, v \rangle}$$

definierte Abbildung $\| \cdot \| : V \to \mathbb{R}^+$ die *Norm* genannt.

Beispiele: (1) Definiere $\cdot : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ durch

$$(z_1, \ldots, z_n) \cdot (w_1, \ldots, w_n) = \sum_{j=1}^{n} z_j \overline{w}_j \; .$$

Dann ist $(\mathbb{C}^n, \cdot)$ ein unitärer Vektorraum , der $\mathbb{C}^n$ *mit dem üblichen Skalarprodukt* genannt wird. Es gilt $\|z\| = \sqrt{\sum_{j=1}^{n} |z_j|^2}$ für alle $z = (z_1, \ldots, z_n) \in \mathbb{C}^n$.

(2) Seien a, $b \in \mathbb{R}$ mit $a < b$ und sei $\mathrm{C}([a, b], \mathbb{C})$ der komplexe Vektorraum aller stetigen Abbildungen von $[a, b]$ nach $\mathbb{C}$. Dann ist $(\mathrm{C}([a, b], \mathbb{C}), \langle \cdot, \cdot \rangle)$ ein unitärer Vektorraum, wobei $\langle \cdot, \cdot \rangle : \mathrm{C}([a, b], \mathbb{C}) \times \mathrm{C}([a, b], \mathbb{C}) \to \mathbb{C}$ durch

$$\langle f, g \rangle = \int_a^b f(t) \overline{g(t)} \, dt$$

definiert ist. (Der Beweis dafür ist eine Übung.)

(3) Sei $\ell_2^{\mathbb{C}}$ der komplexe Vektorraum aller Folgen quadratisch summierbarer komplexer Zahlen. (Eine Folge $\{z_n\}_{n\geq 0}$ heißt quadratisch summierbar, wenn es ein $N \geq 0$ gibt, so dass $\sum_{n=0}^{m} |z_n|^2 \leq N$ für alle $m \geq 0$, d.h., wenn die unendliche Reihe $\sum_{n\geq 0} |z_n|^2$ konvergiert.) Seien $\{z_n\}_{n\geq 0}$, $\{w_n\}_{n\geq 0} \in \ell_2^{\mathbb{C}}$; dann konvergiert die unendliche Reihe $\sum_{n\geq 0} z_n\overline{w}_n$, da $|z_n\overline{w}_n| \leq \frac{1}{2}(|z_n|^2 + |w_n|^2)$. Folglich kann eine Abbildung $\langle\cdot,\cdot\rangle : \ell_2^{\mathbb{C}} \times \ell_2^{\mathbb{C}} \to \mathbb{C}$ durch

$$\langle\{z_n\}_{n\geq 0}, \{w_n\}_{n\geq 0}\rangle = \sum_{n\geq 0} z_n\overline{w}_n$$

definiert werden und dann ist $(\ell_2^{\mathbb{C}}, \langle\cdot,\cdot\rangle)$ ein unitärer Vektorraum. (Der Beweis dafür ist eine Übung.)

In den folgenden zwei Sätzen sowie in Lemma 18.2 sei $(V, \langle\cdot,\cdot\rangle)$ ein unitärer Vektorraum.

Satz 18.3 (Cauchy-Schwarzsche Ungleichung) *Für alle u, $v \in V$ gilt*

$$|\langle u, v\rangle| \leq \|u\|\|v\| \, .$$

Beweis Wenn $v = 0$, dann ist $\|v\| = 0$ und $\langle u, v\rangle = 0$ und in diesem Fall ist $|\langle u, v\rangle| = \|u\|\|v\| = 0$. Sei also $v \neq 0$. Für jedes $\lambda \in \mathbb{C}$ gilt

$$\begin{aligned} 0 &\leq \langle u - \lambda v, u - \lambda v\rangle = \langle u, u - \lambda v\rangle - \lambda\langle v, u - \lambda v\rangle \\ &= \langle u, u\rangle - \overline{\lambda}\langle u, v\rangle - \lambda\langle v, u\rangle + \lambda\overline{\lambda}\langle v, v\rangle = \langle u, u\rangle - \overline{\lambda}\langle u, v\rangle - \lambda\overline{\langle u, v\rangle} + \lambda\overline{\lambda}\langle v, v\rangle \, , \end{aligned}$$

und insbesondere gilt mit $\lambda = \langle v, v\rangle^{-1}\langle u, v\rangle$, dass

$$\begin{aligned} 0 \leq \langle u, u\rangle &- \big(\langle v, v\rangle^{-1}\overline{\langle u, v\rangle}\big)\langle u, v\rangle \\ &- \big(\langle v, v\rangle^{-1}\langle u, v\rangle\big)\overline{\langle u, v\rangle} + \big(\langle v, v\rangle^{-1}\langle u, v\rangle\big)\big(\langle v, v\rangle^{-1}\overline{\langle u, v\rangle}\big)\langle v, v\rangle \\ = \langle u, u\rangle &- \langle v, v\rangle^{-1}\langle u, v\rangle\overline{\langle u, v\rangle} = \langle u, u\rangle - \langle v, v\rangle^{-1}|\langle u, v\rangle|^2 \, , \end{aligned}$$

d.h. $|\langle u, v\rangle|^2 \leq \langle u, u\rangle\langle v, v\rangle$ und damit ist $|\langle u, v\rangle| \leq \|u\|\|v\|$. □

Satz 18.4 *Die Norm $\|\cdot\| : V \to \mathbb{R}^+$ hat folgende Eigenschaften:*

(1) $\|v\| = 0$ genau dann, wenn $v = 0$.

(2) $\|\lambda v\| = |\lambda|\|v\|$ für alle $v \in V$, $\lambda \in \mathbb{C}$.

(3) Dreiecksungleichung) $\|u + v\| \leq \|u\| + \|v\|$ für alle u, $v \in V$.

Beweis (1) und (2) sind klar.

(3): Sei $z = x + iy \in \mathbb{C}$ mit $x, y \in \mathbb{R}$; dann ist $z + \overline{z} = 2x \in \mathbb{R}$ und also ist $z + \overline{z} \leq 2\sqrt{x^2 + y^2} = 2|z|$. Daraus folgt nach Satz 6.3, dass

$$\begin{aligned} \|u+v\|^2 &= \langle u+v, u+v\rangle = \langle u,u\rangle + \langle u,v\rangle + \langle v,u\rangle + \langle v,v\rangle \\ &= \|u\|^2 + \langle u,v\rangle + \overline{\langle u,v\rangle} + \|v\|^2 \leq \|u\|^2 + 2|\langle u,v\rangle| + \|v\|^2 \\ &\leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 = (\|u\| + \|v\|)^2 \end{aligned}$$

und damit ist $\|u + v\| \leq \|u\| + \|v\|$. □

Lemma 18.2 (Polarisierungsidentität) *Für alle $u, v \in V$ gilt*

$$\langle u, v\rangle = \tfrac{1}{4}(\|u+v\|^2 - \|u-v\|^2 + i\|u+iv\|^2 - i\|u-iv\|^2) \; .$$

Beweis Übung. □

Im Folgenden sei $(V, \langle\cdot,\cdot\rangle)$ entweder ein euklidischer Vektorraum oder ein unitärer Vektorraum. Der zugrundeliegende Körper wird mit $\mathbb{K}$ bezeichnet. Dies bedeutet, dass $\mathbb{K} = \mathbb{R}$, falls $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer Vektorraum ist, und $\mathbb{K} = \mathbb{C}$, falls $(V, \langle\cdot,\cdot\rangle)$ ein unitärer Vektorraum ist.

Vektoren $u, v \in V$ heißen *orthogonal* oder *senkrecht zueinander* (geschrieben $u \perp v$), wenn $\langle u, v\rangle = 0$. Vektoren $v_1, \ldots, v_m \in V$ heißen *orthonormal*, wenn

$$\langle v_i, v_j\rangle = \begin{cases} 1 & \text{falls } i = j \; , \\ 0 & \text{falls } i \neq j \; , \end{cases}$$

d.h., wenn $\|v_j\| = 1$ für $j = 1, \ldots, m$ und $v_i \perp v_j$ für alle $i \neq j$.

Lemma 18.3 *Orthonormale Vektoren sind linear unabhängig.*

Beweis Seien $v_1, \ldots, v_m$ orthonormale Vektoren und seien $\lambda_1, \ldots, \lambda_m \in \mathbb{K}$ mit $\lambda_1 v_1 + \cdots + \lambda_m v_m = 0$. Für jedes $j = 1, \ldots, m$ ist dann

$$\begin{aligned} 0 = \langle 0, v_j\rangle &= \langle \lambda_1 v_1 + \cdots + \lambda_m v_m, v_j\rangle \\ &= \lambda_1\langle v_1, v_j\rangle + \cdots + \lambda_m\langle v_m, v_j\rangle = \lambda_j\langle v_j, v_j\rangle = \lambda_j \; , \end{aligned}$$

d.h., $\lambda_j = 0$ für $j = 1, \ldots, m$ und damit sind $v_1, \ldots, v_m$ linear unabhängig. □

Sei U ein endlichdimensionaler Untervektorraum von V. Dann heißt eine Basis $(u_1, \ldots, u_m)$ von U *orthonormal*, wenn die Vektoren $u_1, \ldots, u_m$ orthonormal sind. Seien $u_1, \ldots, u_m \in U$ orthonormal; nach Lemma 18.3 ist $(u_1, \ldots, u_m)$ immer eine orthonormale Basis von $\mathrm{L}(u_1, \ldots, u_m)$. Daraus ergibt sich nach Satz 4.6, dass $(u_1, \ldots, u_m)$ genau dann eine orthonormale Basis von U ist, wenn $\dim U = m$.

Bemerkung: Die kanonische Basis von $\mathbb{R}^n$ ist eine orthonormale Basis von $\mathbb{R}^n$ mit dem üblichen Skalarprodukt. Genauso ist die kanonische Basis von $\mathbb{C}^n$ eine orthonormale Basis von $\mathbb{C}^n$ mit dem üblichen Skalarprodukt.

Satz 18.5 *Sei $(u_1, \ldots, u_m)$ eine orthonormale Basis eines Untervektorraumes U von V. Für alle $u \in U$ gilt dann*

$$u = \langle u, u_1\rangle u_1 + \cdots + \langle u, u_m\rangle u_m \, .$$

(Mit anderen Worten ist $\langle u, u_1\rangle u_1 + \cdots + \langle u, u_m\rangle u_m$ die eindeutige Darstellung von u als Linearkombination der Vektoren $u_1, \ldots, u_m$.)

Beweis Sei $u \in U$; da $(u_1, \ldots, u_m)$ eine Basis von U ist, gibt es nach Satz 4.1 eindeutige Elemente $\lambda_1, \ldots, \lambda_m \in \mathbb{K}$, so dass $u = \lambda_1 u_1 + \cdots + \lambda_m u_m$. Dann gilt

$$\langle u, u_j\rangle = \langle \lambda_1 u_1 + \cdots + \lambda_m u_m, u_j\rangle = \lambda_1\langle u_1, u_j\rangle + \cdots + \lambda_m\langle u_m, u_j\rangle = \lambda_j\langle u_j, u_j\rangle = \lambda_j$$

für jedes $j = 1, \ldots, m$. □

Für jedes $v \in V$ setze

$$v^* = \begin{cases} \|v\|^{-1} v & \text{falls } v \neq 0 \, , \\ 0 & \text{falls } v = 0 \, , \end{cases}$$

also gilt $\|v^*\| = 1$ für alle $v \neq 0$.

Satz 18.6 (Gram-Schmidtsches Orthonormalisierungsverfahren) *Es sei U ein endlichdimensionaler Untervektorraum von V und sei $(u_1, \ldots, u_m)$ eine Basis von U. Definiere $v_1, \ldots, v_m \in V$*

$$\begin{aligned} v_1 &= u_1^* \\ v_2 &= \big(u_2 - \langle u_2, v_1\rangle v_1\big)^* \\ &\vdots \\ v_k &= \big(u_k - \langle u_k, v_1\rangle v_1 - \langle u_k, v_2\rangle v_2 - \cdots - \langle u_k, v_{k-1}\rangle v_{k-1}\big)^* \, , \\ &\vdots \\ v_m &= \big(u_m - \langle u_m, v_1\rangle v_1 - \langle u_m, v_2\rangle \, v_2 - \cdots - \langle u_m, v_{m-1}\rangle v_{m-1}\big)^* \, . \end{aligned}$$

Dann ist $(v_1, \ldots, v_m)$ eine orthonormale Basis von U mit

$$\mathrm{L}(v_1, \ldots, v_k) = \mathrm{L}(u_1, \ldots, u_k)$$

für jedes k. Insbesondere besitzt jeder endlichdimensionale Untervektorraum von V eine orthonormale Basis.

Beweis Es wird durch Induktion nach k bewiesen, dass $v_1, \ldots, v_k$ orthonormal sind und $\mathrm{L}(v_1, \ldots, v_k) = \mathrm{L}(u_1, \ldots, u_k)$ für jedes $k = 1, \ldots, m$.

Induktionsanfang: Da $u_1 \neq 0$, ist $\|v_1\| = 1$ und $\mathrm{L}(v_1) = \mathrm{L}(u_1)$.

Induktionsschritt: Sei $1 \leq k < m$ und nehme an, dass $v_1, \ldots, v_k$ orthonormal sind und $\mathrm{L}(v_1, \ldots, v_k) = \mathrm{L}(u_1, \ldots, u_k)$. Setze

$$w_{k+1} = u_{k+1} - \langle u_{k+1}, v_1\rangle v_1 - \langle u_{k+1}, v_2\rangle v_2 - \cdots - \langle u_{k+1}, v_k\rangle v_k \; .$$

Dann ist $w_{k+1} \in \mathrm{L}(v_1, \ldots, v_k, u_{k+1})$ und $u_{k+1} \in \mathrm{L}(v_1, \ldots, v_k, w_{k+1})$, und daraus ergibt sich nach Satz 3.2, dass $\mathrm{L}(v_1, \ldots, v_k, w_{k+1}) = \mathrm{L}(v_1, \ldots, v_k, u_{k+1})$. Aber $v_{k+1} = w^*_{k+1}$ und $\mathrm{L}(v_1, \ldots, v_k) = \mathrm{L}(u_1, \ldots, u_k)$ und damit ist

$$\begin{aligned}\mathrm{L}(v_1, \ldots, v_{k+1}) &= \mathrm{L}(v_1, \ldots, v_k, v_{k+1}) = \mathrm{L}(v_1, \ldots, v_k, w_{k+1}) \\ &= \mathrm{L}(v_1, \ldots, v_k, u_{k+1}) = \mathrm{L}(u_1, \ldots, u_k, u_{k+1}) = \mathrm{L}(u_1, \ldots, u_{k+1}) \; .\end{aligned}$$

Da $u_1, \ldots, u_{k+1}$ linear unabhängig sind, gilt auch

$$u_{k+1} \notin \mathrm{L}(u_1, \ldots, u_k) = \mathrm{L}(v_1, \ldots, v_k) \; ,$$

und daher ist $w_{k+1} \neq 0$, d.h. $\|v_{k+1}\| = 1$. Für jedes $j = 1, \ldots, k$ ist nun

$$\begin{aligned}\langle v_{k+1}, v_j\rangle &= \|w_{k+1}\|^{-1}\langle w_{k+1}, v_j\rangle \\ &= \|w_{k+1}\|^{-1}\langle\, u_{k+1} - \langle u_{k+1}, v_1\rangle v_1 - \cdots - \langle u_{k+1}, v_k\rangle v_k, v_j\rangle \\ &= \|w_{k+1}\|^{-1}\big(\langle u_{k+1}, v_j\rangle - \langle u_{k+1}, v_1\rangle\langle v_1, v_j\rangle - \cdots - \langle u_{k+1}, v_k\rangle\langle v_k, v_j\rangle\big) \\ &= \|w_{k+1}\|^{-1}\big(\langle u_{k+1}, v_j\rangle - \langle u_{k+1}, v_j\rangle\langle v_j, v_j\rangle\big) \\ &= \|w_{k+1}\|^{-1}\big(\langle u_{k+1}, v_j\rangle - \langle u_{k+1}, v_j\rangle\big) = 0 \; ,\end{aligned}$$

und damit sind $v_1, \ldots, v_{k+1}$ orthonormal, da nach Induktionsannahme $v_1, \ldots, v_k$ schon orthonormal sind. Dies zeigt, dass die Vektoren $v_1, \ldots, v_{k+1}$ orthonormal sind und $\mathrm{L}(v_1, \ldots, v_{k+1}) = \mathrm{L}(u_1, \ldots, u_{k+1})$. □

Es gibt den folgenden Basisergänzungssatz für orthonormale Basen:

Satz 18.7 *Seien U, W Untervektorräume von V mit W endlichdimensional und $\{0\} \neq U \subset W$; sei $(u_1, \ldots, u_m)$ eine orthonormale Basis von U. Dann gibt es $k \geq 0$ und Vektoren $v_1, \ldots, v_k \in W$, so dass $(u_1, \ldots, u_m, v_1, \ldots, v_k)$ eine orthonormale Basis von W ist.*

Beweis Nach dem Basisergänzungssatz (Satz 4.5) gibt es $w_1, \ldots, w_k \in W$ (mit $k \geq 0$), so dass $(u_1, \ldots, u_m, w_1, \ldots, w_k)$ eine Basis von W ist. Wende nun das Gram-Schmidtsche Orthonormalisierungsverfahren auf diese Basis an und erhalte eine orthonormale Basis $(u'_1, \ldots, u'_m, v_1, \ldots, v_k)$ von W. Man sieht aber leicht, dass $u'_j = u_j$ für $j = 1, \ldots, m$, da die Vektoren $u_1, \ldots, u_m$ schon orthonormal sind. Also ist $(u_1, \ldots, u_m, v_1, \ldots, v_k)$ eine orthonormale Basis von W. □

Sei W ein endlichdimensionaler Untervektorraum von V mit $\dim W = n$ und U ein Untervektorraum von V mit $U \subset W$ und $\dim U = m \geq 1$. Nach Satz 18.7

gibt es dann eine orthonormale Basis $(v_1, \ldots, v_n)$ von W, so dass $(v_1, \ldots, v_m)$ eine orthonormale Basis von U ist.

Für jede nichtleere Teilmenge M von V setzen wir

$$M^\perp = \{v \in V : v \perp u \text{ für alle } u \in M\} .$$

$M^\perp$ heißt das *orthogonale Komplement* von M. Offensichtlich ist $\{0\}^\perp = V$ und $V^\perp = \{0\}$.

Lemma 18.4 *$M^\perp$ ist ein Untervektorraum von V.*

Beweis Da $\langle 0, u\rangle = 0$ für alle $u \in M$, ist $0 \in M^\perp$. Seien nun v_1, $v_2 \in M^\perp$ und λ_1, $\lambda_2 \in \mathbb{K}$; für jedes $u \in M^\perp$ ist dann $\langle v_1, u\rangle = \langle v_2, u\rangle = 0$ und damit auch

$$\langle \lambda_1 v_1 + \lambda_2 v_2, u\rangle = \lambda_1\langle v_1, u\rangle + \lambda_2\langle v_2, u\rangle = 0 ,$$

d.h., $\lambda_1 v_1 + \lambda_2 v_2 \in M^\perp$. Folglich ist $M^\perp$ ein Untervektorraum von V. □

Lemma 18.5 *Sei V endlichdimensional und sei U ein Untervektorraum von V mit $\{0\} \neq U \neq V$. Sei $(v_1, \ldots, v_n)$ eine orthonormale Basis von V, so dass $(v_1, \ldots, v_m)$ eine orthonormale Basis von U ist, wobei $m = \dim U$. Dann ist $(v_{m+1}, \ldots, v_n)$ eine orthonormale Basis von $U^\perp$.*

Beweis Sei $v \in U^\perp$; dann gilt $\langle v, v_j\rangle = 0$ für $j = 1, \ldots, m$, da $v_1, \ldots, v_m \in U$, und daraus folgt nach Satz 18.5, dass

$$v = \langle v, v_1\rangle v_1 + \cdots + \langle v, v_n\rangle v_n = \langle v, v_{m+1}\rangle v_{m+1} + \cdots + \langle v, v_n\rangle v_n \in \mathrm{L}(v_{m+1}, \ldots, v_n) ,$$

d.h., $U^\perp \subset \mathrm{L}(v_{m+1}, \ldots, v_n)$. Andererseits ist

$$\langle \lambda_{m+1} v_{m+1} + \cdots + \lambda_n v_n, \lambda_1 v_1 + \cdots + \lambda_m v_m\rangle = 0$$

für alle $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$, und damit ist $\langle \lambda_{m+1} v_{m+1} + \cdots + \lambda_n v_n, u\rangle = 0$ für alle $\lambda_{m+1}, \ldots, \lambda_n \in \mathbb{K}$, $u \in U$, da $(v_1, \ldots, v_m)$ eine Basis von U ist. Folglich gilt auch $\mathrm{L}(v_{m+1}, \ldots, v_n) \subset U^\perp$, d.h., $U^\perp = \mathrm{L}(v_{m+1}, \ldots, v_n)$. Aber $v_{m+1}, \ldots, v_n$ sind orthonormal, und also ist $(v_{m+1}, \ldots, v_n)$ eine orthonormale Basis von $U^\perp$. □

Satz 18.8 *Sei V endlichdimensional und sei U ein Untervektorraum von V. Dann ist $(U^\perp)^\perp = U$ und es gilt $\dim U + \dim U^\perp = \dim V$.*

Beweis Dies ist trivial richtig, wenn $U = \{0\}$ oder $U = V$; nehme also an, dass $\{0\} \neq U \neq V$. Nach Satz 18.7 gibt es eine orthonormale Basis $(v_1, \ldots, v_n)$ von V, so dass $(v_1, \ldots, v_m)$ eine orthonormale Basis von U ist, wobei $m = \dim U$, und nach Lemma 18.5 ist dann $(v_{m+1}, \ldots, v_n)$ eine orthonormale Basis von $U^\perp$. Nun ist $(v_1, \ldots, v_n)$ eine orthonormale Basis von V, so dass $(v_{m+1}, \ldots, v_n)$ eine orthonormale Basis von $U^\perp$ ist. Daraus folgt nach Lemma 18.5, dass $(v_1, \ldots, v_m)$ eine orthonormale Basis von $(U^\perp)^\perp$ ist, und damit ist $(U^\perp)^\perp = U$. Ferner ist $\dim U + \dim U^\perp = m + (n - m) = n = \dim V$. □

Bemerkung Ist V nicht endlichdimensional, so gilt immer noch $U \subset (U^\perp)^\perp$ für jeden Untervektorraum U von V. (Setze $W = U^\perp$; für alle $w \in W$, $u \in U$ ist $w \perp u$ und damit auch $u \perp w$ und folglich ist $U \subset W^\perp = (U^\perp)^\perp$.) Aber im Allgemeinen gilt $U = (U^\perp)^\perp$ nicht: Betrachte ℓ_2 mit dem Skalarprodukt

$$\langle \{x_n\}_{n\geq 0}, \{y_n\}_{n\geq 0} \rangle = \sum_{n\geq 0} x_n y_n$$

und sei $\ell_0 = \{\{x_n\}_{n\geq 0} \in \ell_2 : x_n \neq 0 \text{ für nur endlich viele } n \geq 0\}$. Dann ist ℓ_0 ein Untervektorraum von ℓ_2 mit $\ell_0^\perp = \{0\}$. (Der Beweis dafür ist eine Übung.) Damit ist $(\ell_0^\perp)^\perp = \{0\}^\perp = \ell_2 \neq \ell_0$.

Untervektorräume U_1 und U_2 von V sind *orthogonal*, und man schreibt $U_1 \perp U_2$, wenn $u_1 \perp u_2$ für alle $u_1 \in U_1$, $u_2 \in U_2$.

Sind nun $U, U_1, \ldots, U_m$ Untervektorräume von V, so heißt U die *orthogonale Summe* von $U_1, \ldots, U_m$, wenn $U = U_1 + \cdots + U_m$ und $U_j \perp U_k$ für alle $j \neq k$.

Lemma 18.6 *Eine orthogonale Summe ist direkt: Ist U die orthogonale Summe von $U_1, \ldots, U_m$, so ist auch $U = U_1 \oplus \cdots \oplus U_m$.*

Beweis Seien $u_1, \ldots, u_m \in V$ mit $u_j \in U_j$ für jedes j und $u_1 + \cdots + u_m = 0$. Da $\langle u_j, u_k \rangle = 0$ für alle $j \neq k$, haben wir für jedes $j = 1, \ldots, m$, dass

$$0 = \langle u_1 + \cdots + u_m, u_j \rangle = \langle u_1, u_j \rangle + \cdots + \langle u_m, u_j \rangle = \langle u_j, u_j \rangle$$

und damit $u_j = 0$. Daraus folgt nach Satz 3.7, dass $U = U_1 \oplus \cdots \oplus U_m$. □

Bemerkung Sei V endlichdimensional und seien U, W Untervektorräume von V. Dann ist V die orthogonale Summe von U und W genau, wenn $W = U^\perp$. (Der Beweis dafür ist eine Übung.) Dies bedeutet, dass es zu jedem Untervektorraum U von V einen eindeutigen Untervektorraum W gibt, so dass V die orthogonale Summe von U und W ist. Zum Vergleich: Ist U ein Untervektorraum von V mit $\{0\} \neq U \neq V$, dann gibt es unendlich viele Untervektorräume W, so dass $V = U \oplus W$.

Lemma 18.7 *Seien $U_1, \ldots, U_m$ endlichdimensionale Untervektorräume von V und für $j = 1, \ldots, m$ sei $(u_1^j, \ldots, u_{k_j}^j)$ eine orthonormale Basis von U_j. Dann ist die Summe $U = U_1 + \cdots + U_m$ die orthogonale Summe von $U_1, \ldots, U_m$ genau, wenn $(u_1^1, \ldots, u_{k_1}^1, \ldots, u_1^m, \ldots, u_{k_m}^m)$ eine orthonormale Basis von U ist.*

Beweis Übung. □

19 Orthogonale und unitäre Endomorphismen

Im Folgenden sei $(V, \langle\cdot,\cdot\rangle)$ entweder ein euklidischer Vektorraum oder ein unitärer Vektorraum. Der zugrundeliegende Körper wird mit $\mathbb{K}$ bezeichnet. Dies bedeutet, dass $\mathbb{K} = \mathbb{R}$, falls $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer Vektorraum ist, und $\mathbb{K} = \mathbb{C}$, falls $(V, \langle\cdot,\cdot\rangle)$ ein unitärer Vektorraum ist. Ist $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer Vektorraum und $\lambda \in \mathbb{R}$, so ist $\overline{\lambda}$ einfach als λ zu interpretieren.

Lemma 19.1 *Sei $\alpha : V \to V$ eine Abbildung, für die gilt:*

$$\langle \alpha(u), \alpha(v)\rangle = \langle u, v\rangle$$

für alle u, $v \in V$. Dann ist α linear; d.h. α ist ein Endomorphismus von V.

Beweis Für alle u, $v \in V$ gilt

$$\begin{aligned}
&\langle \alpha(u+v) - \alpha(u) - \alpha(v), \alpha(u+v) - \alpha(u) - \alpha(v)\,\rangle \\
&\qquad = \langle \alpha(u+v), \alpha(u+v)\rangle - \langle \alpha(u+v), \alpha(u)\rangle - \langle \alpha(u+v), \alpha(v)\rangle \\
&\qquad\qquad - \langle \alpha(u), \alpha(u+v)\rangle + \langle \alpha(u), \alpha(u)\rangle + \langle \alpha(u), \alpha(v)\rangle \\
&\qquad\qquad - \langle \alpha(v), \alpha(u+v)\rangle + \langle \alpha(v), \alpha(u)\rangle + \langle \alpha(v), \alpha(v)\rangle \\
&\qquad = \langle u+v, u+v\rangle - \langle u+v, u\rangle - \langle u+v, v\rangle - \langle u, u+v\rangle \\
&\qquad\qquad + \langle u, u\rangle + \langle u, v\rangle - \langle v, u+v\rangle + \langle v, u\rangle + \langle v, v\rangle \\
&\qquad = \langle (u+v) - u - v, (u+v) - u - v\rangle = \langle 0, 0\rangle = 0
\end{aligned}$$

und daraus folgt, dass $\alpha(u+v) - \alpha(u) - \alpha(v) = 0$, d.h., $\alpha(u+v) = \alpha(u) + \alpha(v)$ für alle u, $v \in V$. Ferner gilt für alle $v \in V$, $\lambda \in \mathbb{K}$, dass

$$\begin{aligned}
&\langle \alpha(\lambda v) - \lambda\alpha(v), \alpha(\lambda v) - \lambda\alpha(v)\rangle \\
&\qquad = \langle \alpha(\lambda v), \alpha(\lambda v)\rangle - \overline{\lambda}\langle \alpha(\lambda v), \alpha(v)\rangle - \lambda\langle \alpha(v), \alpha(\lambda v)\rangle + \lambda\overline{\lambda}\langle \alpha(v), \alpha(v)\rangle \\
&\qquad = \langle \lambda v, \lambda v\rangle - \overline{\lambda}\langle \lambda v, v\rangle - \lambda\langle v, \lambda v\rangle + \lambda\overline{\lambda}\langle v, v\rangle \\
&\qquad = \langle \lambda v - \lambda v, \lambda v - \lambda v\rangle = \langle 0, 0\rangle = 0\ ,
\end{aligned}$$

und daraus folgt, dass $\alpha(\lambda v) - \lambda\alpha(v) = 0$, d.h. $\alpha(\lambda v) = \lambda\alpha(v)$ für alle $v \in V$, $\lambda \in \mathbb{K}$. Damit ist α eine lineare Abbildung. □

Ein Endomorphismus $f : V \to V$ heißt *orthogonal*, falls $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer Vektorraum ist, bzw. *unitär*, falls $(V, \langle\cdot,\cdot\rangle)$ ein unitärer Vektorraum ist, wenn

$$\langle f(u), f(v)\rangle = \langle u, v\rangle$$

für alle u, $v \in V$.

Satz 19.1 *Ein Endomorphismus $f : V \to V$ ist orthogonal bzw. unitär genau dann, wenn $\|f(v)\| = \|v\|$ für alle $v \in V$.*

Beweis Ist f orthogonal bzw. unitär, dann ist

$$\|f(v)\|^2 = \langle f(v), f(v)\rangle = \langle v, v\rangle = \|v\|^2$$

und damit $\|f(v)\| = \|v\|$ für alle $v \in V$. Für die Umkehrung nehme zunächst an, dass $(V, \langle\cdot,\cdot\rangle)$ ein unitärer Vektorraum ist. Ist $\|f(w)\| = \|w\|$ für alle $w \in V$, so gilt nach Lemma 18.2, dass

$$\begin{aligned}
&4\langle f(u), f(v)\rangle \\
&\quad = \|f(u) + f(v)\|^2 - \|f(u) - f(v)\|^2 + i\|f(u) + if(v)\|^2 - i\|f(u) - if(v)\|^2 \\
&\quad = \|f(u+v)\|^2 - \|f(u-v)\|^2 + i\|f(u+iv)\|^2 - i\|f(u-iv)\|^2 \\
&\quad = \|u+v\|^2 - \|u-v\|^2 + i\|u+iv\|^2 - i\|u-iv\|^2 = 4\langle u, v\rangle
\end{aligned}$$

für alle u, $v \in V$, d.h., f ist unitär. Wenn $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer Vektorraum ist, dann benutzt man eine der Identitäten in Lemma 18.1 statt Lemma 18.2. □

Ist $f : V \to V$ orthogonal bzw. unitär, dann gilt $f(0) = 0$ und für alle u, $v \in V$ gilt $\|f(u) - f(v)\| = \|u - v\|$, da

$$\|f(u) - f(v)\|^2 = \|f(u-v)\|^2 = \langle f(u-v), f(u-v)\rangle = \langle u-v, u-v\rangle = \|u-v\|^2 .$$

Für euklidische Vektorräume ist die Umkehrung auch richtig:

Satz 19.2 *Sei $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer Vektorraum und sei $\alpha : V \to V$ eine (beliebige) Abbildung mit $\alpha(0) = 0$ und*

$$\|\alpha(u) - \alpha(v)\| = \|u - v\|$$

für alle u, $v \in V$. Dann ist α ein orthogonaler Endomorphismus. (Hier bedeutet 'beliebig': Es wird nicht vorausgesetzt, dass α eine lineare Abbildung ist.)

Beweis Seien u, $v \in V$; nach der Polarisierungsidentität in Lemma 18.1 ist

$$\begin{aligned}
2\langle\alpha(u), \alpha(v)\rangle &= \|\alpha(u)\|^2 + \|\alpha(v)\|^2 - \|\alpha(u) - \alpha(v)\|^2 \\
&= \|\alpha(u) - \alpha(0)\|^2 + \|\alpha(v) - \alpha(0)\|^2 - \|\alpha(u) - \alpha(v)\|^2 \\
&= \|u - 0\|^2 + \|v - 0\|^2 - \|u - v\|^2 \\
&= \|u\|^2 + \|v\|^2 - \|u - v\|^2 = 2\langle u, v\rangle
\end{aligned}$$

und damit ist $\langle\alpha(u), \alpha(v)\rangle = \langle u, v\rangle$ für alle u, $v \in V$. Daraus ergibt sich nach Lemma 19.1, dass α ein orthogonaler Endomorphismus ist. □

Betrachte den unitären Vektorraum $\mathbb{C} = \mathbb{C}^1$ mit dem üblichen Skalarprodukt, also ist hier die Norm einfach $|\cdot|$, und sei $\alpha : \mathbb{C} \to \mathbb{C}$ die durch $\alpha(z) = \overline{z}$ definierte Abbildung. Dann gilt $\alpha(0) = 0$ und $|\alpha(z) - \alpha(z')| = |z - z'|$ für alle z, $z' \in \mathbb{C}$, aber die Abbildung α ist nicht linear. Dies zeigt, dass für unitäre Vektorräume Satz 19.2 im Allgemeinen falsch ist.

Seien f, $g : V \to V$ orthogonale bzw. unitäre Endomorphismen von V. Dann ist der Endomorphismus $gf : V \to V$ auch orthogonal bzw. unitär, da

$$\langle (gf)(u), (gf)(v)\rangle = \langle g(f(u)), g(f(v))\rangle = \langle f(u), f(v)\rangle = \langle u, v\rangle$$

für alle u, $v \in V$.

Lemma 19.2 *Sei V endlichdimensional und sei $f : V \to V$ ein orthogonaler bzw. ein unitärer Endomorphismus von V. Dann ist f ein Automorphismus und f^{-1} ist auch orthogonal bzw. unitär.*

Beweis Sei $v \in \operatorname{Kern} f$; dann ist $0 = \langle 0, 0\rangle = \langle f(v), f(v)\rangle = \langle v, v\rangle$ und damit ist $v = 0$, d.h., $\operatorname{Kern}\ f = \{0\}$. Da aber V endlichdimensional ist, ist nach Satz 5.3 f ein Automorphismus. Ferner ist

$$\langle f^{-1}(u), f^{-1}(v)\rangle = \langle f(f^{-1}(u)), f(f^{-1}(v))\rangle = \langle u, v\rangle$$

für alle u, $v \in V$, d.h., f^{-1} ist orthogonal bzw. unitär. □

Satz 19.3 *Sei V endlichdimensional und $f : V \to V$ ein Endomorphismus von V. Dann sind äquivalent:*

(1) f ist orthogonal bzw. unitär.

(2) f ist ein Automorphismus und $\langle f^{-1}(u), v\rangle = \langle u, f(v)\rangle$ für alle u, $v \in V$. (Später stellt sich heraus, dass f^{-1} der zu f adjungierte Endomorphismus ist.)

Beweis (1) ⇒ (2): Nach Lemma 19.2 ist f ein Automorphismus und

$$\langle f^{-1}(u), v\rangle = \langle f(f^{-1}(u)), f(v)\rangle = \langle u, f(v)\rangle$$

für alle u, $v \in V$.

(2) ⇒ (1): Für alle u, $v \in V$ gilt $\langle u, v\rangle = \langle f^{-1}(f(u)), v\rangle = \langle f(u), f(v)\rangle$ und damit ist f orthogonal bzw. unitär. □

Lemma 19.3 *Sei V endlichdimensional, $(v_1, \ldots, v_n)$ eine orthonormale Basis von V und sei $f : V \to V$ ein Endomorphismus. Dann ist f orthogonal bzw. unitär genau, wenn $(f(v_1), \ldots, f(v_n))$ eine orthonormale Basis von V ist.*

Beweis Ist f orthogonal bzw. unitär, so ist $\langle f(v_j), f(v_k)\rangle = \langle v_j, v_k\rangle$ für alle j, k. Folglich sind die Vektoren $f(v_1), \ldots, f(v_n)$ orthonormal, d.h., $(f(v_1), \ldots, f(v_n))$ ist eine orthonormale Basis von V. Nehme umgekehrt an, dass $(f(v_1), \ldots, f(v_n))$ eine orthonormale Basis von V ist. Seien $u, v \in V$; dann gibt es eindeutige Darstellungen $u = \lambda_1 v_1 + \cdots + \lambda_n v_n$ und $v = \mu_1 v_1 + \cdots + \mu_n v_n$ und dann ist

$$\begin{aligned}\langle f(u), f(v)\rangle &= \langle f(\lambda_1 v_1 + \cdots + \lambda_n v_n), f(\mu_1 v_1 + \cdots + \mu_n v_n)\rangle \\ &= \langle \lambda_1 f(v_1) + \cdots + \lambda_n f(v_n), \mu_1 f(v_1) + \cdots + \mu_n f(v_n)\rangle \\ &= \lambda_1\overline{\mu}_1 + \cdots + \lambda_n\overline{\mu}_n = \langle \lambda_1 v_1 + \cdots + \lambda_n v_n, \mu_1 v_1 + \cdots + \mu_n v_n\rangle \\ &= \langle u, v\rangle \ ,\end{aligned}$$

und daraus ergibt sich, dass f orthogonal bzw. unitär ist. □

Lemma 19.4 *Sei V endlichdimensional und sei $f : V \to V$ ein orthogonaler bzw. ein unitärer Endomorphismus. Ferner sei U ein Untervektorraum von V mit $f(U) \subset U$. Dann gilt auch $f(U^\perp) \subset U^\perp$.*

Beweis Sei $v \in U^\perp$; für jedes $u \in U$ gilt dann $0 = \langle v, u\rangle = \langle f(v), f(u)\rangle$. Also ist $f(v) \perp w$ für alle $w \in f(U)$, d.h., $f(v) \in f(U)^\perp$. Daraus ergibt sich, dass $f(U^\perp) \subset f(U)^\perp$. (Dies ist auch richtig, wenn V nicht endlichdimensional ist.) Aber V ist endlichdimensional und nach Lemma 19.2 ist f ein Automorphismus, und daraus folgt nach Lemma 5.10, dass $\dim f(U) = \dim U$. Damit ist $f(U) = U$, da $f(U) \subset U$, d.h., $f(U^\perp) \subset U^\perp$. □

Satz 19.4 *Sei $f : V \to V$ ein orthogonaler bzw. ein unitärer Endomorphismus von V und sei λ ein Eigenwert von f. Dann gilt $|\lambda| = 1$. Ist also f orthogonal, so ist λ entweder -1 oder 1; ist dagegen f unitär, so gibt es $\theta \in \mathbb{R}$, so dass $\lambda = e^{i\theta}$. Sind ferner λ_1 und λ_2 verschiedene Eigenwerte von f, so sind die Eigenräume $\mathrm{E}(f, \lambda_1)$ und $\mathrm{E}(f, \lambda_2)$ orthogonal.*

Beweis Sei $v \in V$ ein Eigenvektor zum Eigenwert λ. Dann ist

$$\langle v, v\rangle = \langle f(v), f(v)\rangle = \langle \lambda v, \lambda v\rangle = \lambda\overline{\lambda}\langle v, v\rangle = |\lambda|^2\langle v, v\rangle \ ,$$

und folglich ist $|\lambda|^2 = 1$ und damit $|\lambda| = 1$, da $\langle v, v\rangle \neq 0$. Seien nun λ_1 und λ_2 verschiedene Eigenwerte von f und $v_1 \in \mathrm{E}(f, \lambda_1)$, $v_2 \in \mathrm{E}(f, \lambda_2)$. Dann ist

$$\langle v_1, v_2\rangle = \langle f(v_1), f(v_2)\rangle = \langle \lambda_1 v_1, \lambda_2 v_2\rangle = \lambda_1\overline{\lambda}_2 \langle v_1, v_2\rangle \ .$$

Aber $\lambda_1\overline{\lambda}_2 \neq 1$ (Warum?) und damit ist $\langle v_1, v_2\rangle = 0$. Folglich sind die Eigenräume $\mathrm{E}(f, \lambda_1)$ und $\mathrm{E}(f, \lambda_2)$ orthogonal. □

Das folgende Ergebnis bezieht sich nur auf unitäre Vektorräume. (Das analoge Behauptung für euklidische Vektorräume ist im Allgemeinen falsch.)

Satz 19.5 *Sei $(V, \langle\cdot,\cdot\rangle)$ ein endlichdimensionaler unitärer Vektorraum (also ist $(V, \langle\cdot,\cdot\rangle)$ ein unitärer Vektorraum mit V einem endlichdimensionalen komplexen Vektorraum), sei $f : V \to V$ ein unitärer Endomorphismus und seien $\lambda_1, \ldots, \lambda_m$ die verschiedenen Eigenwerte von f. Dann ist V die orthogonale Summe der Eigenräume $\mathrm{E}(f, \lambda_1), \ldots, \mathrm{E}(f, \lambda_m)$.*

Beweis Sei $U = \mathrm{E}(f, \lambda_1) + \cdots + \mathrm{E}(f, \lambda_m)$; nach Satz 19.4 ist U die orthogonale Summe von $\mathrm{E}(f, \lambda_1), \ldots, \mathrm{E}(f, \lambda_m)$ und also muss gezeigt werden, dass $U = V$. Nehme an, dass $U \neq V$ und setze $W = U^\perp$. Da $f(\mathrm{E}(f, \lambda_j)) \subset \mathrm{E}(f, \lambda_j)$ für jedes j, ist auch $f(U) \subset U$, und daraus folgt nach Lemma 19.4, dass $f(W) \subset W$. Sei $g : W \to W$ die Einschränkung von f auf W. Da $\dim W \geq 1$, ist χ_g ein nicht konstantes Polynom und folglich gilt nach Satz 16.4 und Satz 16.5, dass es einen Eigenwert $\lambda \in \mathbb{C}$ von g gibt. Sei $v \in V$ ein Eigenvektor von g zum Eigenwert λ. Dann ist $f(v) = g(v) = \lambda v$, d.h., λ ist auch ein Eigenwert von f, und damit $\lambda = \lambda_j$ für ein j, und v ist ein Eigenvektor von f zum Eigenwert λ. Also ist $v \in U \cap W$ und dies ist ein Widerspruch, da $U \perp W$. Daraus ergibt sich, dass $U = V$, d.h., V ist die orthogonale Summe der Eigenräume $\mathrm{E}(f, \lambda_1), \ldots, \mathrm{E}(f, \lambda_m)$. □

Satz 19.6 *Sei $(V, \langle\cdot,\cdot\rangle)$ ein endlichdimensionaler unitärer Vektorraum und sei f ein unitärer Endomorphismus. Dann gibt es eine aus Eigenvektoren von f bestehende orthonormale Basis von V. Ferner ist $|\det f| = 1$.*

Beweis Seien $\lambda_1, \ldots, \lambda_m$ die verschiedenen Eigenwerte von f und für jedes j sei $(u_1^j, \ldots, u_{k_j}^j)$ eine orthonormale Basis von $\mathrm{E}(f, \lambda_j)$. Dann ist nach Satz 19.5 und Lemma 18.7 $(v_1, \ldots, v_n) = (u_1^1, \ldots, u_{k_1}^1, \ldots, u_1^m, \ldots, u_{k_m}^m)$ eine orthonormale Basis von V, d.h., $(v_1, \ldots, v_n)$ ist eine aus Eigenvektoren bestehende orthonormale Basis von V. Sei $D = (d_{ij})$ die Matrix von f bezüglich $(v_1, \ldots, v_n)$. Dann ist D eine Diagonalmatrix und die Einträge $d_{11}, \ldots, d_{nn}$ auf der Diagonale sind alle Eigenwerte von f. Daraus folgt nach Satz 19.4 und Lemma 16.1, dass

$$|\det f| = |\det D| = |d_{11} \cdots d_{nn}| = |d_{11}| \cdots |d_{nn}| = 1 \; . \quad \square$$

Das dem Satz 19.5 entsprechende Ergebnis für orthogonale Endomorphismen ist im Allgemeinen falsch. Jetzt wird untersucht, wie orthogonale Endomorphismen dargestellt werden kann und es wird mit dem Fall begonnen, in dem $(V, \langle\cdot,\cdot\rangle)$ ein zweidimensionaler euklidischer Vektorraum ist (das heißt: $(V, \langle\,\cdot\,\cdot\rangle)$ ist ein euklidischer Vektorraum mit $\dim V = 2$).

Bis auf weiteres sei also $(V, \langle\cdot,\cdot\rangle)$ ein zweidimensionaler euklidischer Vektorraum. Sei $f : V \to V$ ein orthogonaler Endomorphismus und sei θ mit $0 \leq \theta \leq \pi$. Dann heißt f eine *Drehung um den Winkel θ*, wenn

$$\langle f(v), v\rangle = \|v\|^2 \cos\theta$$

für alle $v \in V$.

Lemma 19.5 *(1) Die Identitätsabbildung* $\mathrm{id}_V : V \to V$ *ist die einzige Drehung um den Winkel* 0*.*

(2) Die Abbildung $-\mathrm{id}_V$ *ist die einzige Drehung um den Winkel* π*.*

(3) Ist $f : V \to V$ *eine Drehung um einen Winkel* θ*, wobei* $0 < \theta < \pi$*, so besitzt* f *keinen Eigenwert.*

Beweis (1) Es ist klar, dass id_V eine Drehung um den Winkel 0 ist. Sei umgekehrt $f : V \to V$ eine Drehung um den Winkel 0, es gilt also $\langle f(v), v\rangle = \|v\|^2$ für alle $v \in V$ und daraus folgt, dass

$$\begin{aligned}\langle f(v) - v, f(v) - v\rangle &= \langle f(v), f(v)\rangle - 2\langle f(v), v\rangle + \langle v, v\rangle \\ &= \langle v, v\rangle - 2\langle f(v), v\rangle + \langle v, v\rangle = \langle v, v\rangle - 2\langle v, v\rangle + \langle v, v\rangle = 0\end{aligned}$$

und damit $f(v) = v$ für alle $v \in V$.

(2) Es ist klar, dass $-\mathrm{id}_V$ eine Drehung um den Winkel π ist. Sei umgekehrt $f : V \to V$ eine Drehung um den Winkel π, es gilt also $\langle f(v), v\rangle = -\|v\|^2$ für alle $v \in V$ und daraus folgt, dass

$$\begin{aligned}\langle f(v) + v, f(v) + v\rangle &= \langle f(v), f(v)\rangle + 2\,\langle f(v), v\rangle + \langle v, v\rangle \\ &= \langle v, v\rangle + 2\langle f(v), v\rangle + \langle v, v\rangle = \langle v, v\rangle - 2\langle v, v\rangle + \langle v, v\rangle = 0\end{aligned}$$

und damit $f(v) = -v$ für alle $v \in V$.

(3) Sei λ ein Eigenwert von f und sei $v \in V$ ein Eigenvektor zum Eigenwert λ. Dann ist $\langle f(v), v\rangle = \langle \lambda v, v\rangle = \lambda\langle v, v\rangle = \lambda\|v\|^2 \neq \|v\|^2 \cos\theta$, da nach Satz 19.4 λ nur -1 oder 1 sein kann und $-1 < \cos\theta < 1$, wenn $0 < \theta < \pi$. □

Ein orthogonaler Endomorphismus $f : V \to V$ heißt *Spiegelung*, wenn

$$\dim \mathrm{E}(f, 1) = \dim \mathrm{E}(f, -1) = 1\ .$$

Sei f eine Spiegelung; dann gilt nach Satz 19.4, dass V die orthogonale Summe der Eigenräume $\mathrm{E}(f, 1)$ und $\mathrm{E}(f, -1)$ ist. Sei v_1 bzw. v_2 ein Eigenvektor von f zum Eigenwert 1 bzw. zum Eigenwert -1 mit $\|v_1\| = \|v_2\| = 1$. Dann ist (v_1, v_2) eine orthonormale Basis von V und die Matrix von f bezüglich (v_1, v_2) ist

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

Insbesondere ist $\det f = -1$, da $\det S = -1$.

Für jedes $\theta \in \mathbb{R}$ seien D_θ, $S_\theta \in \mathrm{M}(2 \times 2, \mathbb{R})$ folgende Matrizen:

$$D_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\ , \quad S_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}\ .$$

Insbesondere ist $D_0 = E_2$, $D_\pi = -E_2$ und $S_0 = S$. Es gilt $\chi_{S_\theta} = -1 + x^2$ für jedes $\theta \in \mathbb{R}$ (und damit sind -1 und 1 beide Eigenwerte von S_θ), da

$$\det \begin{pmatrix} \cos\theta - x & \sin\theta \\ \sin\theta & -\cos\theta - x \end{pmatrix}$$
$$= -(\cos\theta - x)(\cos\theta + x) - \sin^2\theta = -\cos^2\theta - \sin^2\theta + x^2 = -1 + x^2 \ .$$

Ferner ist $\det D_\theta = 1$ und $\det S_\theta = -1$ für jedes $\theta \in \mathbb{R}$.

Satz 19.7 *(1) Sei $f : V \to V$ ein orthogonaler Endomorphismus. Dann ist f entweder eine Drehung oder eine Spiegelung.*

(2) Sei $f : V \to V$ eine Drehung um den Winkel θ, wobei $0 \leq \theta \leq \pi$, und sei (v_1, v_2) eine orthonormale Basis von V. Dann ist die Matrix von f bezüglich (v_1, v_2) entweder D_θ oder $D_{2\pi-\theta}$. Insbesondere ist $\det f = 1$.

(3) Sei $f : V \to V$ eine Spiegelung und sei θ mit $0 \leq \theta < 2\pi$. Dann gibt es eine orthonormale Basis (v_1, v_2) von V, so dass S_θ die Matrix von f bezüglich (v_1, v_2) ist.

Beweis (1) Sei (v_1, v_2) eine orthonormale Basis von V und sei

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

die Matrix von f bezüglich (v_1, v_2). Es gilt also

$$\begin{aligned} 1 &= \langle v_1, v_1\rangle = \langle f(v_1), f(v_1)\rangle = \langle av_1 + cv_2, av_1 + cv_2\rangle = a^2 + c^2 \ , \\ 1 &= \langle v_2, v_2\rangle = \langle f(v_2), f(v_2)\rangle = \langle bv_1 + dv_2, bv_1 + dv_2\rangle = b^2 + d^2 \ , \\ 0 &= \langle v_1, v_2\rangle = \langle f(v_1), f(v_2)\rangle = \langle av_1 + cv_2, bv_1 + dv_2\rangle = ab + cd \ , \end{aligned}$$

und damit gibt es ein eindeutiges θ mit $0 \leq \theta < 2\pi$, so dass $A = D_\theta$ oder $A = S_\theta$. (Warum?) Nehme zunächst an, dass $A = D_\theta$. Sei $v \in V$; dann gibt es eindeutige Elemente λ_1, $\lambda_2 \in \mathbb{R}$, so dass $v = \lambda v_1 + \lambda_2 v_2$ und

$$\begin{aligned} \langle f(v), v\rangle &= \langle f(\lambda_1 v_1 + \lambda_2 v_2), \lambda_1 v_1 + \lambda_2 v_2\rangle = \langle \lambda_1 f(v_1) + \lambda_2 f(v_2), \lambda_1 v_1 + \lambda_2 v_2\rangle \\ &= \langle \lambda_1(\cos\theta\, v_1 + \sin\theta\, v_2) + \lambda_2(-\sin\theta\, v_1 + \cos\theta\, v_2), \lambda_1 v_1 + \lambda_2 v_2\rangle \\ &= \langle (\lambda_1\cos\theta - \lambda_2\sin\theta)v_1 + (\lambda_1\sin\theta + \lambda_2\cos\theta)v_2, \lambda_1 v_1 + \lambda_2 v_2\rangle \\ &= (\lambda_1\cos\theta - \lambda_2\sin\theta)\lambda_1 + (\lambda_1\sin\theta + \lambda_2\cos\theta)\lambda_2 \\ &= (\lambda_1^2 + \lambda_2^2)\cos\theta = \|v\|^2\cos\theta \ . \end{aligned}$$

Daraus ergibt sich, dass f eine Drehung um den Winkel θ' ist, wobei

$$\theta' = \begin{cases} \theta & \text{falls } 0 \leq \theta \leq \pi \ , \\ 2\pi - \theta & \text{falls } \pi < \theta < 2\pi \ , \end{cases}$$

da $0 \leq \theta' \leq \pi$ und $\cos\theta' = \cos\theta$. Nehme nun an, dass $A = S_\theta$. Dann sind -1 und 1 Eigenwerte von f, da -1 und 1 Eigenwerte von S_θ sind, und folglich ist $\dim \mathrm{E}(f,1) = \dim \mathrm{E}(f,-1) = 1$, d.h., f ist eine Spiegelung.

(2) Sei $f : V \to V$ eine Drehung um den Winkel θ, wobei $0 \leq \theta < \pi$, sei (v_1, v_2) eine orthonormale Basis von V und sei A die Matrix von f bezüglich (v_1, v_2). Genauso wie in (1) gibt es dann ein eindeutiges θ' mit $0 \leq \theta' < 2\pi$, so dass $A = D_{\theta'}$ oder $A = S_{\theta'}$. Nach Lemma 19.5 können aber -1 und 1 nicht beide Eigenwerte von f sein und damit ist $A = S_{\theta'}$ nicht möglich, d.h., $A = D_{\theta'}$. Daraus folgt (genauso wie in (1)), dass $\langle f(v), v\rangle = \|v\|^2 \cos\theta'$ für alle $v \in V$. Insbesondere ist $\cos\theta' = \cos\theta$, d.h., θ' ist entweder θ oder $2\pi - \theta$.

(3) Sei $f : V \to V$ eine Spiegelung und sei u_1 bzw. u_2 ein Eigenvektor von f zum Eigenwert 1 bzw. zum Eigenwert -1 mit $\|u_1\| = \|u_2\| = 1$. Also ist (u_1, u_2) eine orthonormale Basis von V und $S = S_0$ ist die Matrix von f bezüglich (u_1, u_2). Sei θ mit $0 \leq \theta < 2\pi$, setze $t = \theta/2$ und definiere Vektoren v_1, $v_2 \in V$ durch $v_1 = \cos t\, u_1 - \sin t\, u_2$ und $v_2 = \sin t\, u_1 + \cos t\, u_2$. Dann ist (v_1, v_2) eine orthonormale Basis von V, da

$$\begin{aligned}
\langle v_1, v_1\rangle &= \langle \cos t\, u_1 - \sin t\, u_2, \cos t\, u_1 - \sin t\, u_2\rangle = \cos^2 t + \sin^2 t = 1\ ,\\
\langle v_2, v_2\rangle &= \langle \sin t\, u_1 + \cos t\, u_2, \sin t\, u_1 + \cos t\, u_2\rangle = \sin^2 t + \cos^2 t = 1\ ,\\
\langle v_1, v_2\rangle &= \langle \cos t\, u_1 - \sin t\, u_2, \sin t\, u_1 + \cos t\, u_2\rangle = \cos t \sin t - \sin t \cos t = 0\ ,
\end{aligned}$$

und die Matrix für den Wechsel von (u_1, u_2) nach (v_1, v_2) ist

$$D_{-t} = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}\ .$$

Daraus folgt nach Satz 8.8, dass $(D_{-t})^{-1} S_0 D_{-t}$ die Matrix von f bezüglich (v_1, v_2) ist. Aber man sieht leicht, dass $(D_{-t})^{-1} S_0 D_{-t} = D_t S_0 D_{-t} = S_\theta$. Damit ist S_θ die Matrix von f bezüglich der orthonormalen Basis (v_1, v_2). □

Sei θ mit $0 < \theta < \pi$; dann gibt es genau zwei Endomorphismen, die Drehungen um den Winkel θ sind: Ist (v_1, v_2) eine orthonormale Basis von V, so hat eine dieser Drehungen die Matrix D_θ und die andere die Matrix $D_{2\pi-\theta}$ bezüglich (v_1, v_2).

Man sagt, dass $f : V \to V$ eine *nichttriviale Drehung* ist, wenn f eine Drehung um einen Winkel θ mit $0 < \theta < \pi$ ist. Nichttriviale Drehungen sind also genau die orthogonalen Endomorphismen $f : V \to V$, die keine Eigenwerte besitzen.

Die Untersuchung von orthogonalen Endomorphismen für den Fall $\dim V = 2$ ist jetzt abgeschlossen. Im Folgenden sei nun $(V, \langle\cdot,\cdot\rangle)$ ein endlichdimensionaler euklidischer Vektorraum.

Sei U ein Untervektorraum von V. Dann ist die Einschränkung $\langle\cdot,\cdot\rangle_U$ von $\langle\cdot,\cdot\rangle$ auf $U \times U$ ein Skalarprodukt. Ist $f : V \to V$ ein orthogonaler Endomorphismus und ist U f-invariant, so ist die Einschränkung von f auf U auch orthogonal bezüglich des euklidischen Vektorraumes $(U, \langle\cdot,\cdot\rangle_U)$.

Satz 19.8 *Sei $f : V \to V$ ein orthogonaler Endomorphismus. Dann gibt es zweidimensionale f-invariante Untervektorräume $D_1, \ldots, D_m$ (mit $m \geq 0$), so dass für jedes j die Einschränkung von f auf D_j eine nichttriviale Drehung ist und V die orthogonale Summe der Untervektorräume $\mathrm{E}(f,1)$, $\mathrm{E}(f,-1)$, $D_1, \ldots, D_m$ ist.*

Beweis Der folgende Hilfssatz wird benötigt:

Satz 19.9 *Sei W ein endlichdimensionaler reeller Vektorraum mit $\dim W \geq 1$ und sei $g : W \to W$ ein Endomorphismus. Dann gibt es einen g-invarianten Untervektorraum U von W mit $\dim U$ entweder 1 oder 2.*

Beweis Sei $(w_1, \ldots, w_n)$ eine Basis von W und sei $A \in \mathrm{M}(n \times n, \mathbb{R})$ die Matrix von g bezüglich dieser Basis. Sei $A_{\mathbb{C}} \in \mathrm{M}(n \times n, \mathbb{C})$ die Matrix A als komplexe Matrix angesehen. Es gibt dann den reellen Endomorphismus $\varphi_A : \mathbb{R}^n \to \mathbb{R}^n$ und den komplexen Endomorphismus $\varphi_{A_{\mathbb{C}}} : \mathbb{C}^n \to \mathbb{C}^n$, und die Abbildung φ_A ist die Einschränkung der Abbildung $\varphi_{A_{\mathbb{C}}}$ auf der Teilmenge $\mathbb{R}^n$ von $\mathbb{C}^n$. Nach Satz 16.4 und Satz 16.5 gibt es einen Eigenwert $\lambda = \alpha + i\beta \in \mathbb{C}$ (mit α, $\beta \in \mathbb{R}$) von $\varphi_{A_{\mathbb{C}}}$; sei $z = x + iy \in \mathbb{C}^n$ (mit x, $y \in \mathbb{R}^n$) ein Eigenvektor zum Eigenwert λ. Dann gilt

$$\begin{aligned} 0 = \varphi_{A_{\mathbb{C}}}(z) - \lambda z &= \varphi_{A_{\mathbb{C}}}(x + iy) - (\alpha + i\beta)(x + iy) \\ &= \varphi_{A_{\mathbb{C}}}(x) + i\varphi_{A_{\mathbb{C}}}(y) - (\alpha x - \beta y) - i(\beta x + \alpha y) \\ &= \varphi_A(x) + i\,\varphi_A(y) - (\alpha x - \beta y) - i(\beta x + \alpha y) \end{aligned}$$

und daraus folgt, dass $\varphi_A(x) = \alpha x - \beta y$ und $\varphi_A(y) = \beta x + \alpha y$. Sei $V = \mathrm{L}(x, y)$; also ist V ein Untervektorraum von $\mathbb{R}^n$ mit $\dim V$ entweder 1 oder 2. (Da $z \neq 0$, ist $x \neq 0$ oder $y \neq 0$.) Ferner ist nach Lemma 5.8 und Satz 3.2

$$\varphi_A(V) = \varphi_A(\mathrm{L}(x, y)) = \mathrm{L}(\varphi_A(x), \varphi_A(y)) \subset \mathrm{L}(x, y) = V\ .$$

Sei $\psi : \mathbb{R}^n \to W$ die Abbildung, die gegeben ist durch

$$\psi((\lambda_1, \ldots, \lambda_n)) = \lambda_1 w_1 + \cdots + \lambda_n w_n$$

und setze $U = \psi(V)$; nach Lemma 6.7 ist ψ ein Isomorphismus und damit ergibt sich nach Lemma 5.10, dass U ein Untervektorraum von W mit $\dim U = \dim V$ ist, d.h., $\dim U$ ist entweder 1 oder 2. Aber nach Satz 8.2 gilt $g \circ \psi = \psi \circ \varphi_A$ und folglich ist $g(U) = g(\psi(V)) = \psi(\varphi_A(V)) \subset \psi(V) = U$, d.h., U ist g-invariant. □

Bemerkung Für komplexe Vektorräume ist die Situation viel einfacher: Sei W ein endlichdimensionaler komplexer Vektorraum mit $\dim W \geq 1$ und sei $g : W \to W$ ein Endomorphismus. Dann gibt es immer einen g-invarianten Untervektorraum U von W mit $\dim U = 1$. (Nach Satz 16.4 und Satz 16.5 gibt es einen Eigenwert

$\lambda \in \mathbb{C}$ von g. Sei w ein Eigenvektor zum Eigenwert λ und setze $U = \mathrm{L}(w)$; dann ist $\dim U = 1$ und $g(U) \subset U$.)

Ist $(W, \langle \cdot, \cdot \rangle)$ ein endlichdimensionaler euklidischer Vektorraum und $g : W \to W$ ein orthogonaler Endomorphismus, dann sei

$$\Delta(g) = \dim W - \dim \mathrm{E}(g, 1) - \dim \mathrm{E}(g, -1) \; .$$

Der Beweis für Satz 19.8 erfolgt durch Induktion nach $n = \Delta(f)$.

Induktionsanfang: Sei f ein orthogonaler Endomorphismus mit $\Delta(f) = 0$. Dann gilt nach Satz 19.4, dass V die orthogonale Summe der Untervektorräume $\mathrm{E}(f, 1)$ und $\mathrm{E}(f, -1)$ ist.

Induktionsschritt: Sei $n > 0$ und nehme an, dass die Aussage in Satz 19.9 für jeden orthogonalen Endomorphismus $g : W \to W$ mit $\Delta(g) < n$ richtig ist. Sei $f : V \to V$ ein orthogonaler Endomorphismus mit $\Delta(f) = n$ und setze $E = \mathrm{E}(f, 1) + \mathrm{E}(f, -1)$; nach Satz 18.8 und Satz 19.4 ist

$$\dim E^\perp = \dim V - \dim E = \dim V - \dim \mathrm{E}(f, 1) - \dim \mathrm{E}(f, -1) = n$$

und nach Lemma 19.4 ist $f(E^\perp) \subset E^\perp$, da $f(E) \subset E$. Sei f' die Einschränkung von f auf $E^\perp$; nach Satz 19.9 gibt es dann einen f'-invarianten Untervektorraum U von $E^\perp$ mit $\dim U$ entweder 1 oder 2, und hier ist tatsächlich $\dim U = 2$. (Ist $\dim U = 1$, so ist jeder Vektor in $U \setminus \{0\}$ ein Eigenvektor von f' und damit auch von f. Aber alle Eigenvektoren von f liegen in E und $E \cap U \subset E \cap E^\perp = \{0\}$.) Nun ist der Untervektorraum U auch f-invariant und die Einschränkung von f auf U ist eine nichttriviale Drehung, da U keinen Eigenvektor von f enthält. Setze $W = U^\perp$; nach Lemma 19.4 gilt $f(W) \subset W$, da $f(U) \subset U$, sei also g die Einschränkung von f auf W. Dann ist g ein orthogonaler Endomorphismus bezüglich des euklidischen Vektorraumes $(W, \langle \cdot, \cdot \rangle_W)$. Ferner ist $\mathrm{E}(g, 1) = \mathrm{E}(f, 1)$ und $\mathrm{E}(g, -1) = \mathrm{E}(f, -1)$, da $E \subset W$, und damit ist $\Delta(g) = \Delta(f) - 2 < n$. Nach der Induktionsannahme gibt es also zweidimensionale g-invariante Untervektorräume $D_1, \ldots, D_m$ (mit $m \geq 0$), so dass für jedes j die Einschränkung von g auf D_j eine nichttriviale Drehung ist und W die orthogonale Summe von $\mathrm{E}(g, 1)$, $\mathrm{E}(g, -1)$, $D_1, \ldots, D_m$ ist. Daher ist die Einschränkung von f auf D_j, die nichts anderes ist als die Einschränkung von g auf D_j, auch eine nichttriviale Drehung. Da $\mathrm{E}(g, 1) = \mathrm{E}(f, 1)$ und $\mathrm{E}(g, -1) = \mathrm{E}(f, -1)$, ist V die orthogonale Summe der Untervektorräume $\mathrm{E}(f, 1)$, $\mathrm{E}(f, -1)$, $D_1, \ldots, D_m$ und U. □

Satz 19.10 *Sei $f : V \to V$ ein orthogonaler Endomorphismus. Dann ist*

$$\det f = (-1)^k \; ,$$

wobei $k = \dim \mathrm{E}(f, -1)$. Insbesondere ist $\det f$ *entweder* 1 *oder* -1.

Beweis Nach Satz 19.8 ist V die orthogonale Summe der Untervektorräume

$$\mathrm{E}(f,1),\ \mathrm{E}(f,-1),\ D_1,\ \ldots,\ D_m\ ,$$

wobei für jedes j die Einschränkung f_j von f auf D_j eine nichttriviale Drehung ist. Sei f_0 die Einschränkung von f auf $\mathrm{E}(f,1)$ und f_0' die Einschränkung von f auf $\mathrm{E}(f,-1)$. Da eine orthogonale Summe eine direkte Summe ist, ist f die direkte Summe der Endomorphismen $f_0,\ f_0',\ f_1,\ \ldots,\ f_m$ und folglich ist

$$\det f = (\det f_0)(\det f_0')(\det f_1)\cdots(\det f_m)\ .$$

Aber $\det f_0 = 1$, $\det f_0' = (-1)^k$ und nach Satz 19.7 (2) ist $\det f_j = 1$ für jedes $j = 1,\ \ldots,\ m$. Damit ist $\det f = (-1)^k$. □

Der spezielle Fall von $\mathbb{R}^3$ mit dem üblichen Skalarprodukt wird nun untersucht.

Satz 19.11 *Sei $f : \mathbb{R}^3 \to \mathbb{R}^3$ ein orthogonaler Endomorphismus mit* $\det f = 1$. *Dann ist $\mathbb{R}^3$ die orthogonale Summe von f-invarianten Untervektorräumen L und D mit* $\dim L = 1$, $\dim D = 2$, *so dass $L \subset \mathrm{E}(f,1)$ und die Einschränkung von f auf D eine Drehung ist. Mit anderen Worten: Jeder orthogonale Endomorphismus von $\mathbb{R}^3$ mit Determinante 1 ist eine Drehung um eine Achse.*

Beweis Nach Satz 19.10 muss $\dim \mathrm{E}(f,1)$ ungerade und $\dim \mathrm{E}(f,-1)$ gerade sein. Es gibt also drei Möglichkeiten:

(1) $\dim \mathrm{E}(f,1) = 3$ und $\dim \mathrm{E}(f,-1) = 0$: Hier ist $f = \mathrm{id}_{\mathbb{R}^3}$. Sei $D = L^\perp$, wobei L ein beliebiger eindimensionaler Untervektorraum von $\mathbb{R}^3$ ist. Dann ist die Einschränkung von f auf D eine Drehung um den Winkel 0.

(2) $\dim \mathrm{E}(f,1) = 1$ und $\dim \mathrm{E}(f,-1) = 2$: Sei $L = \mathrm{E}(f,1)$ und $D = \mathrm{E}(f,-1)$. Dann ist die Einschränkung von f auf D eine Drehung um den Winkel π.

(3) $\dim \mathrm{E}(f,1) = 1$ und $\dim \mathrm{E}(f,-1) = 0$: Sei $L = \mathrm{E}(f,1)$ und $D = L^\perp$. Der Beweis für Satz 19.8 zeigt dann, dass die Einschränkung von f auf D eine nichttriviale Drehung ist. □

20 Orthogonale und unitäre Matrizen

Eine reelle Matrix $A \in \mathrm{M}(n \times n, \mathbb{R})$ heißt *orthogonal*, wenn die Spalten von A eine orthonormale Basis von $\mathbb{R}^n$ (mit dem üblichen Skalarprodukt) bilden. Also ist die Matrix $A = (a_{ij}) \in \mathrm{M}(n \times n, \mathbb{R})$ genau dann orthogonal, wenn gilt:

$$\sum_{i=1}^{n} a_{ij}a_{ij} = 1 \quad \text{für } j = 1, \ldots, n \,,$$
$$\sum_{i=1}^{n} a_{ij}a_{ik} = 0 \quad \text{für alle } j \neq k \,.$$

Ist $A = (a_{ij}) \in \mathrm{M}(n \times n, \mathbb{R})$, so bezeichnet A^t wieder die zu A transponierte Matrix, d.h., $A^t = (a^t_{ij})$, wobei $a^t_{ij} = a_{ji}$.

Satz 20.1 *Sei $A \in \mathrm{M}(n \times n, \mathbb{R})$; dann sind äquivalent:*

(1) A ist orthogonal.

(2) $A^t A = E_n$.

(3) A ist invertierbar und $A^{-1} = A^t$.

(4) $AA^t = E_n$.

(5) Die Zeilen von A bilden eine orthonormale Basis von $\mathbb{R}^n$ (mit dem üblichen Skalarprodukt).

(6) Der Endomorphismus $\varphi_A : \mathbb{R}^n \to \mathbb{R}^n$ ist orthogonal bezüglich $\mathbb{R}^n$ mit dem üblichen Skalarprodukt.

Beweis (1) $\Leftrightarrow$ (2): Sei $A = (a_{ij})$ und setze $\delta_{ij} = \begin{cases} 1 & \text{falls } i = j \,, \\ 0 & \text{falls } i \neq j \,. \end{cases}$ Dann gilt

$$\begin{aligned} A \text{ ist orthogonal} \quad &\Leftrightarrow \quad \sum_{i=1}^{n} a_{ij}a_{ik} = \delta_{jk} \text{ für alle } j,\, k \\ &\Leftrightarrow \quad \sum_{i=1}^{n} a^t_{ji}a_{ik} = \delta_{jk} \text{ für alle } j,\, k \quad \Leftrightarrow \quad A^t A = E_n \,. \end{aligned}$$

(4) $\Leftrightarrow$ (5): Genauso.

(2) $\Leftrightarrow$ (3) $\Leftrightarrow$ (4): Dies folgt unmittelbar aus Satz 6.2.

(1) $\Leftrightarrow$ (6): Seien $v_1, \ldots, v_n \in \mathbb{R}^n$ die Spalten von A; also gilt $\varphi_A(e_j) = v_j$ für jedes j, wobei $(e_1, \ldots, e_n)$ die kanonische Basis von $\mathbb{R}^n$ ist. Aber $(e_1, \ldots, e_n)$ ist eine orthonormale Basis, und daraus ergibt sich nach Lemma 19.3, dass φ_A genau dann orthogonal ist, wenn $(v_1, \ldots, v_n)$ eine orthonormale Basis von $\mathbb{R}^n$ ist. Damit ist φ_A orthogonal genau dann, wenn A orthogonal ist. $\square$

Lemma 20.1 *(1) Es gilt* $(AB)^t = B^t A^t$ *für alle* $A, B \in \mathrm{M}(n \times n, \mathbb{R})$.

(2) Ist $A \in \mathrm{M}(n \times n, \mathbb{R})$ *invertierbar, so ist* A^t *auch invertierbar und es gilt* $(A^t)^{-1} = (A^{-1})^t$.

Beweis (1) Übung.

(2) Dies folgt unmittelbar aus (1) und Satz 6.2, da

$$A^t(A^{-1})^t = (A^{-1}A)^t = E_n^t = E_n \; . \; \square$$

Die Menge der $n \times n$ orthogonalen Matrizen wird mit $\mathrm{O}(n)$ bezeichnet.

Satz 20.2 $\mathrm{O}(n)$ *ist eine Untergruppe der Gruppe* $\mathrm{GL}(n, \mathbb{R})$ *aller invertierbaren* $n \times n$ *reellen Matrizen. Mit anderen Worten:*

(1) Die Einheitsmatrix E_n *ist orthogonal.*

(2) Ist A *orthogonal, so ist* A^{-1} *orthogonal.*

(3) Sind A, B *orthogonal, so ist* AB *orthogonal.*

Beweis (1) Dies ist klar.

(2) Nach Satz 20.1, Lemma 20.1 (2) und Lemma 6.3 (3) ist

$$(A^{-1})^t A^{-1} = (A^t)^{-1} A^{-1} = (AA^t)^{-1} = E_n^{-1} = E_n$$

und damit ist nach Satz 20.1 A^{-1} auch orthogonal.

(3) Nach Satz 20.1 und Lemma 20.1 (1) ist

$$(AB)^t AB = B^t A^t AB = B^t E_n B = B^t B = E_n$$

und damit ist nach Satz 20.1 AB auch orthogonal. $\square$

Satz 20.3 *Sei* $A \in \mathrm{M}(n \times n, \mathbb{R})$ *orthogonal; dann ist* $\det A$ *entweder* 1 *oder* -1.

Beweis Nach Satz 20.1, Satz 15.3 und Satz 15.4 ist

$$(\det A)^2 = (\det A)(\det A) = (\det A^t)(\det A) = \det(A^t A) = \det E_n = 1$$

und damit ist $\det A$ entweder 1 oder -1. $\square$

Im Folgenden sei $(V, \langle \cdot, \cdot \rangle)$ ein endlichdimensionaler euklidischer Vektorraum.

Lemma 20.2 *Sei $f : V \to V$ ein Endomorphismus von V, sei $(v_1, \ldots, v_n)$ eine orthonormale Basis und sei A die Matrix von f bezüglich $(v_1, \ldots, v_n)$. Definiere $B = (b_{ij}) \in \mathrm{M}(n \times n, \mathbb{R})$ durch $b_{ij} = \langle f(v_i), f(v_j)\rangle$. Dann ist $B = A^t A$.*

Beweis Für alle i, j ist

$$b_{ij} = \langle f(v_i), f(v_j)\rangle = \Big\langle \sum_{k=1}^{n} a_{ki} v_k, \sum_{\ell=1}^{n} a_{\ell j} v_\ell \Big\rangle$$
$$= \sum_{k=1}^{n}\sum_{\ell=1}^{n} a_{ki} a_{\ell j} \langle v_k, v_\ell\rangle = \sum_{k=1}^{n} a_{ki} a_{kj} = \sum_{k=1}^{n} a^t_{ik} a_{kj} \ ,$$

d.h., $B = A^t A$. □

Satz 20.4 *Sei $f : V \to V$ ein Endomorphismus von V und sei $(v_1, \ldots, v_n)$ eine orthonormale Basis. Dann ist die Matrix A von f bezüglich $(v_1, \ldots, v_n)$ orthogonal genau, wenn der Endomorphismus f orthogonal ist.*

Beweis Definiere $B = (b_{ij}) \in \mathrm{M}(n \times n, \mathbb{R})$ durch $b_{ij} = \langle f(v_i), f(v_j)\rangle$ und also ist nach Lemma 20.2 $B = A^t A$. Nach Lemma 19.3 ist aber f orthogonal genau dann, wenn $(f(v_1), \ldots, f(v_n))$ eine orthonormale Basis von V ist und daher genau dann, wenn $B = E_n$. Folglich ist f orthogonal genau dann, wenn A orthogonal ist. □

Satz 20.5 *Sei $(v_1, \ldots, v_n)$ eine orthonormale Basis von V und sei $(u_1, \ldots, u_n)$ eine (beliebige) Basis von V; sei P die Matrix für den Wechsel von $(v_1, \ldots, v_n)$ nach $(u_1, \ldots, u_n)$. Dann ist $(u_1, \ldots, u_n)$ eine orthonormale Basis genau, wenn P orthogonal ist.*

Beweis Für alle $1 \le k, \ell \le n$ ist

$$\langle u_k, u_\ell\rangle = \Big\langle \sum_{i=1}^{n} p_{ik} v_i, \sum_{j=1}^{n} p_{j\ell} v_j \Big\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} p_{ik} p_{j\ell} \langle v_i, v_j\rangle = \sum_{j=1}^{n} p_{jk} p_{j\ell} = \sum_{j=1}^{n} p^t_{kj} p_{j\ell}$$

und folglich gilt $\langle u_k, u_\ell\rangle = \delta_{k\ell}$ für alle j, k genau dann, wenn $P^t P = E_n$. Daraus ergibt sich nach Satz 20.1, dass $(u_1, \ldots, u_n)$ genau dann eine orthonormale Basis ist, wenn die Matrix P orthogonal ist. □

Für jedes $\theta \in \mathbb{R}$ seien wieder D_θ, $S_\theta \in \mathrm{M}(2 \times 2, \mathbb{R})$ folgende Matrizen:

$$D_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \qquad S_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \ .$$

Da $\sin^2\theta + \cos^2\theta = 1$, sind D_θ und S_θ orthogonale Matrizen.

Satz 20.6 *Sei $A \in \mathrm{M}(2 \times 2, \mathbb{R})$ orthogonal; dann gibt es ein eindeutiges θ mit $0 \leq \theta < 2\pi$, so dass A entweder D_θ oder S_θ ist. Mit anderen Worten ist*

$$\{D_\theta : 0 \leq \theta < 2\pi\} \cup \{S_\theta : 0 \leq \theta < 2\pi\}$$

genau die Menge aller 2×2 orthogonalen Matrizen.

Beweis Übung. □

Sei $0 \leq \theta < 2\pi$; dann ist

$$D_\theta \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{pmatrix} .$$

Sei nun $v \in \mathbb{R}^2$; dann gibt es $r \geq 0$ und $0 \leq \alpha < 2\pi$, so dass $v = \begin{pmatrix} r\cos\alpha \\ r\sin\alpha \end{pmatrix}$, und

$$\varphi_{D_\theta}(v) = D_\theta \begin{pmatrix} r\cos\alpha \\ r\sin\alpha \end{pmatrix} = \begin{pmatrix} r\cos\alpha\cos\theta - r\sin\alpha\sin\theta \\ r\cos\alpha\sin\theta + r\sin\alpha\cos\theta \end{pmatrix} = \begin{pmatrix} r\cos(\alpha+\theta) \\ r\sin(\alpha+\theta) \end{pmatrix} .$$

Also ist φ_{D_θ} wirklich eine Drehung um den Winkel θ. Für die Matrix S_θ gilt

$$S_\theta \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\theta + y\sin\theta \\ x\sin\theta - y\cos\theta \end{pmatrix} .$$

Sei $v \in \mathbb{R}^2$ mit $v = \begin{pmatrix} r\cos\alpha \\ r\sin\alpha \end{pmatrix}$; dann ist

$$\varphi_{S_\theta}(v) = S_\theta \begin{pmatrix} r\cos\alpha \\ r\sin\alpha \end{pmatrix} = \begin{pmatrix} r\cos\alpha\cos\theta + r\sin\alpha\sin\theta \\ r\cos\alpha\sin\theta - r\sin\alpha\cos\theta \end{pmatrix} = \begin{pmatrix} r\cos(\theta-\alpha) \\ r\sin(\theta+\alpha) \end{pmatrix} .$$

Also ist φ_{S_θ} eine Spiegelung an der Geraden $y = (\theta/2)x$.

Sei $\mathrm{M}_{\mathrm{O}}(n \times n, \mathbb{R})$ die Menge aller Matrizen in $\mathrm{M}(n \times n, \mathbb{R})$ mit folgender Block-Gestalt:

$$\begin{pmatrix} E_p & & & & \\ & -E_q & & & \\ & & D_{\theta_1} & & \\ & & & \ddots & \\ & & & & D_{\theta_m} \end{pmatrix} ,$$

wobei $p \geq 0$, $q \geq 0$, $m \geq 0$ und $0 < \theta_j < \pi$ für $j = 1, \ldots, m$ (und natürlich ist dann $p + q + 2m = n$).

Satz 20.7 *Sei $A \in \mathrm{M}(n \times n, \mathbb{R})$ orthogonal; dann gibt es eine orthogonale Matrix $P \in \mathrm{M}(n \times n, \mathbb{R})$, so dass $P^{-1}AP \in \mathrm{M}_{\mathrm{O}}(n \times n, \mathbb{R})$.*

Beweis Setze $f = \varphi_A$; nach Satz 20.1 ist $f : \mathbb{R}^n \to \mathbb{R}^n$ orthogonal und folglich gibt es nach Satz 19.8 zweidimensionale f-invariante Untervektorräume $D_1, \ldots, D_m$ von $\mathbb{R}^n$ (mit $m \geq 0$), so dass für jedes j die Einschränkung f_j von f auf D_j eine nichttriviale Drehung ist und $\mathbb{R}^n$ die orthogonale Summe der Untervektorräume $E(f,1)$, $E(f,-1)$, $D_1, \ldots, D_m$ ist. Sei $1 \leq k \leq m$; nach Satz 19.7 (2) gibt es eine orthonormale Basis (u_1^k, u_2^k) von D_2, so dass D_{θ_k} die Matrix von f_k bezüglich (u_1^k, u_2^k) ist. (Ist $D_{2\pi-\theta_k}$ die Matrix von f_k bezüglich einer orthonormalen Basis (w_1, w_2), so ist D_{θ_k} die Matrix von f_k bezüglich (w_2, w_1).) Sei $(u_1, \ldots, u_p)$ eine orthonormale Basis von $E(f,1)$ und $(u'_1, \ldots, u'_q)$ eine orthonormale Basis von $E(f,-1)$. Nach Lemma 18.7 ist dann

$$(v_1, \ldots, v_n) = (u_1, \ldots, u_p, u'_1, \ldots, u'_q, u_1^1, u_2^1, \ldots, u_1^m, u_2^m)$$

eine orthonormale Basis von $\mathbb{R}^n$ und

$$B = \begin{pmatrix} E_p & & & & \\ & -E_q & & & \\ & & D_{\theta_1} & & \\ & & & \ddots & \\ & & & & D_{\theta_m} \end{pmatrix}$$

ist die Matrix von $f = \varphi_A$ bezüglich $(v_1, \ldots, v_n)$. Sei P die Matrix für den Wechsel von der kanonischen Basis $(e_1, \ldots, e_n)$ von $\mathbb{R}^n$ nach $(v_1, \ldots, v_n)$. Nach Satz 20.5 ist P orthogonal, da $(e_1, \ldots, e_n)$ eine orthonormale Basis von $\mathbb{R}^n$ ist, und nach Satz 8.8 ist $B = P^{-1}AP$, da A die Matrix von φ_A bezüglich $(e_1, \ldots, e_n)$ ist. Insbesondere ist $P^{-1}AP \in \mathrm{M_O}(n \times n, \mathbb{R})$. □

Eine komplexe Matrix $A \in \mathrm{M}(n \times n, \mathbb{C})$ heißt *unitär*, wenn die Spalten von A eine orthonormale Basis von $\mathbb{C}^n$ (mit dem üblichen Skalarprodukt) bilden. Also ist die Matrix $A = (a_{ij}) \in \mathrm{M}(n \times n, \mathbb{C})$ genau dann unitär, wenn gilt:

$$\sum_{i=1}^{n} a_{ij}\overline{a}_{ij} = 1 \text{ für } j = 1, \ldots, n \quad \text{und} \quad \sum_{i=1}^{n} a_{ij}\overline{a}_{ik} = 0 \text{ für alle } j \neq k \,,$$

(wobei $\overline{a}_{ij}$ statt $\overline{a_{ij}}$ geschrieben wird). Ist $A = (a_{ij}) \in \mathrm{M}(n \times n, \mathbb{C})$, so bezeichnet A^t die zu A transponierte Matrix, d.h., $A^t = (a_{ij}^t)$, wobei $a_{ij}^t = a_{ji}$, und $\overline{A}$ bezeichnet die zu A konjugierte Matrix, d.h., $\overline{A} = (\overline{a}_{ij})$.

Satz 20.8 *Sei $A \in \mathrm{M}(n \times n, \mathbb{C})$; dann sind äquivalent:*

(1) A ist unitär.

(2) $A^t\overline{A} = E_n$.

(3) A ist invertierbar und $A^{-1} = \overline{A}^t$.

(4) $A\overline{A}^t = E_n$.

(5) Die Zeilen von A bilden eine orthonormale Basis von $\mathbb{C}^n$ *(mit dem üblichen Skalarprodukt).*

(6) Der Endomorphismus $\varphi_A : \mathbb{C}^n \to \mathbb{C}^n$ *ist unitär bezüglich* $\mathbb{C}^n$ *mit dem üblichen Skalarprodukt.*

Beweis (1) $\Leftrightarrow$ (2): Sei $A = (a_{ij})$ und setze $\delta_{ij} = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j \end{cases}$. Dann gilt

$$\begin{aligned} A \text{ ist orthogonal} &\Leftrightarrow \sum_{i=1}^{n} a_{ij}\overline{a}_{ik} = \delta_{jk} \text{ für alle } j,\, k \\ &\Leftrightarrow \sum_{i=1}^{n} a_{ji}^t\overline{a}_{ik} = \delta_{jk} \text{ für alle } j,\, k \quad \Leftrightarrow \quad A^t\overline{A} = E_n \,. \end{aligned}$$

(4) $\Leftrightarrow$ (5): Genauso.

(2) $\Leftrightarrow$ (3) $\Leftrightarrow$ (4): Dies folgt unmittelbar aus Satz 6.2.

(1) $\Leftrightarrow$ (6): Seien $v_1, \ldots, v_n \in \mathbb{C}^n$ die Spalten von A; also gilt $\varphi_A(e_j) = v_j$ für jedes j, wobei $(e_1, \ldots, e_n)$ die kanonische Basis von $\mathbb{C}^n$ ist. Aber $(e_1, \ldots, e_n)$ ist eine orthonormale Basis, und daraus ergibt sich nach Lemma 19.3, dass φ_A genau dann unitär ist, wenn $(v_1, \ldots, v_n)$ eine orthonormale Basis von $\mathbb{C}^n$ ist. Damit ist φ_A unitär genau dann, wenn A unitär ist. □

Lemma 20.3 *(1) Es gilt* $\overline{AB} = \overline{A}\,\overline{B}$ *für alle* $A, B \in \mathrm{M}(n \times n, \mathbb{C})$.

(2) Ist $A \in \mathrm{M}(n \times n, \mathbb{C})$ *invertierbar, so ist* $\overline{A}$ *auch invertierbar und es gilt* $(\overline{A})^{-1} = \overline{A^{-1}}$.

(3) Es gilt $\overline{A}^t = \overline{A^t}$ *für alle* $A \in \mathrm{M}(n \times n, \mathbb{C})$.

Beweis (1) und (3) sind klar.

(2) Dies folgt unmittelbar aus (1) und Satz 6.2, da

$$\overline{A^{-1}}\,\overline{A} = \overline{A^{-1}A} = \overline{E}_n = E_n \,. \quad \square$$

Die Menge der $n \times n$ unitären Matrizen wird mit $\mathrm{U}(n)$ bezeichnet.

Satz 20.9 $\mathrm{U}(n)$ *ist eine Untergruppe der Gruppe* $\mathrm{GL}(n, \mathbb{C})$ *aller invertierbaren* $n \times n$ *komplexen Matrizen. Mit anderen Worten:*

(1) Die Einheitsmatrix E_n *ist unitär.*

(2) Ist A unitär, so ist A^{-1} *unitär.*

(3) Sind A, B unitär, so ist AB unitär.

Beweis (1) Dies ist klar.

(2) Nach Satz 20.8, Lemma 20.1 (2), Lemma 20.3 (2) und Lemma 6.3 (3) ist

$$(A^{-1})^t\overline{A^{-1}} = (A^t)^{-1}(\overline{A})^{-1} = (\overline{AA^t})^{-1} = E_n^{-1} = E_n$$

und damit ist nach Satz 20.8 A^{-1} auch unitär.

(3) Nach Satz 20.8, Lemma 20.1 (1) und Lemma 20.3 (1) ist

$$(AB)^t\overline{AB} = B^tA^t\overline{A}\,\overline{B} = B^tE_n\overline{B} = B^t\overline{B} = E_n$$

und damit ist nach Satz 20.8 AB auch unitär. □

Satz 20.10 *Sei $A \in \mathrm{M}(n \times n, \mathbb{C})$ unitär; dann ist $|\det A| = 1$.*

Beweis Es ist klar, dass $\det \overline{A} = \overline{\det A}$, und daraus folgt nach Satz 20.8, Satz 15.3 und Satz 15.4, dass

$$|\det A|^2 = (\det A)(\overline{\det A}) = (\det A^t)(\det \overline{A}) = \det(A^t\overline{A}) = \det E_n = 1\ .$$

Damit ist $|\det A| = 1$. □

Im Folgenden sei $(V, \langle\cdot,\cdot\rangle)$ ein endlichdimensionaler unitärer Vektorraum.

Lemma 20.4 *Sei $f : V \to V$ ein Endomorphismus von V, sei $(v_1, \ldots, v_n)$ eine orthonormale Basis und sei A die Matrix von f bezüglich $(v_1, \ldots, v_n)$. Definiere $B = (b_{ij}) \in \mathrm{M}(n \times n, \mathbb{C})$ durch $b_{ij} = \langle f(v_i), f(v_j)\rangle$. Dann ist $B = A^t\overline{A}$.*

Beweis Für alle i, j ist

$$\begin{aligned} b_{ij} = \langle f(v_i), f(v_j)\rangle &= \Big\langle \sum_{k=1}^{n} a_{ki}v_k, \sum_{\ell=1}^{n} a_{\ell j}v_\ell \Big\rangle \\ &= \sum_{k=1}^{n}\sum_{\ell=1}^{n} a_{ki}\overline{a}_{\ell j}\langle v_k, v_\ell\rangle = \sum_{k=1}^{n} a_{ki}\overline{a}_{kj} = \sum_{k=1}^{n} a_{ik}^t\overline{a}_{kj}\ , \end{aligned}$$

d.h., $B = A^t\overline{A}$. □

Satz 20.11 *Sei $f : V \to V$ ein Endomorphismus von V und sei $(v_1, \ldots, v_n)$ eine orthonormale Basis. Dann ist die Matrix A von f bezüglich $(v_1, \ldots, v_n)$ unitär genau, wenn der Endomorphismus f unitär ist.*

Beweis Definiere $B = (b_{ij}) \in \mathrm{M}(n \times n, \mathbb{C})$ durch $b_{ij} = \langle f(v_i), f(v_j)\rangle$ und also ist nach Lemma 20.4 $B = A^t\overline{A}$. Nach Lemma 19.3 ist aber f unitär genau dann, wenn $(f(v_1), \ldots, f(v_n))$ eine orthonormale Basis von V ist und daher genau dann, wenn $B = E_n$. Folglich ist f unitär genau dann, wenn A unitär ist. □

Satz 20.12 *Sei $(v_1, \ldots, v_n)$ eine orthonormale Basis von V und sei $(u_1, \ldots, u_n)$ eine (beliebige) Basis von V; sei P die Matrix für den Wechsel von $(v_1, \ldots, v_n)$ nach $(u_1, \ldots, u_n)$. Dann ist $(u_1, \ldots, u_n)$ eine orthonormale Basis genau, wenn P unitär ist.*

Beweis Für alle $1 \leq k, \ell \leq n$ ist

$$\langle u_k, u_\ell\rangle = \Big\langle \sum_{i=1}^{n} p_{ik}v_i, \sum_{j=1}^{n} p_{j\ell}v_j \Big\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} p_{ik}\overline{p}_{j\ell}\langle v_i, v_j\rangle = \sum_{j=1}^{n} p_{jk}\overline{p}_{j\ell} = \sum_{j=1}^{n} p^t_{kj}\overline{p}_{j\ell}$$

und folglich gilt $\langle u_k, u_\ell\rangle = \delta_{k\ell}$ für alle j, k genau dann, wenn $P^t\overline{P} = E_n$. Daraus ergibt sich nach Satz 20.8, dass $(u_1, \ldots, u_n)$ genau dann eine orthonormale Basis ist, wenn die Matrix P unitär ist. □

Satz 20.13 *Sei $A \in \mathrm{M}(n \times n, \mathbb{C})$ unitär; dann gibt es eine unitäre Matrix P, so dass $P^{-1}AP$ eine Diagonalmatrix ist. Insbesondere ist jede unitäre Matrix diagonalisierbar.*

Beweis Nach Satz 20.8 ist $\varphi_A : \mathbb{C}^n \to \mathbb{C}^n$ unitär (bezüglich $\mathbb{C}^n$ mit dem üblichen Skalarprodukt) und folglich gibt es nach Satz 19.6 eine aus Eigenvektoren von φ_A bestehende orthonormale Basis $(v_1, \ldots, v_n)$ von $\mathbb{C}^n$. Die Matrix D von φ_A bezüglich $(v_1, \ldots, v_n)$ ist also eine Diagonalmatrix. Sei P die Matrix für den Wechsel von der kanonischen Basis $(e_1, \ldots, e_n)$ von $\mathbb{C}^n$ nach $(v_1, \ldots, v_n)$. Nach Satz 20.12 ist P unitär, da $(e_1, \ldots, e_n)$ eine orthonormale Basis von $\mathbb{C}^n$ ist, und da A die Matrix von φ_A bezüglich $(e_1, \ldots, e_n)$ ist, ist nach Satz 8.8 $D = P^{-1}AP$, d.h. $P^{-1}AP$ ist eine Diagonalmatrix. □

21 Selbstadjungierte Endomorphismen

Im Folgenden sei $(V, \langle \cdot, \cdot \rangle)$ entweder ein euklidischer Vektorraum oder ein unitärer Vektorraum. Der zugrundeliegende Körper wird mit $\mathbb{K}$ bezeichnet. Es gilt also $\mathbb{K} = \mathbb{R}$, falls $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum ist, und $\mathbb{K} = \mathbb{C}$, falls $(V, \langle \cdot, \cdot \rangle)$ ein unitärer Vektorraum ist. Ist $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum und $\lambda \in \mathbb{R}$, so ist $\overline{\lambda}$ einfach als λ zu interpretieren.

Ein Endomorphismus $f : V \to V$ heißt *selbstadjungiert*, wenn

$$\langle f(u), v \rangle = \langle u, f(v) \rangle$$

für alle u, $v \in V$.

Satz 21.1 *Sei $(V, \langle \cdot, \cdot \rangle)$ ein unitärer Vektorraum. Dann ist ein Endomorphismus $f : V \to V$ selbstadjungiert genau, wenn $\langle f(v), v \rangle$ reell ist für alle $v \in V$.*

Beweis Ist f selbstadjungiert, so ist

$$\langle f(v), v \rangle = \langle v, f(v) \rangle = \overline{\langle f(v), v \rangle}$$

und damit ist $\langle f(v), v \rangle$ reell für alle $v \in V$. Die Umkehrung ist eine Übung. □

Für euklidische Vektorräume gibt es kein Ergebnis, das Satz 21.1 entspricht.

Satz 21.2 *Sei $(V, \langle \cdot, \cdot \rangle)$ ein unitärer Vektorraum und sei $f : V \to V$ ein selbstadjungierter Endomorphismus. Dann ist jeder Eigenwert von f reell.*

Beweis Sei $\lambda \in \mathbb{C}$ ein Eigenwert von f und sei $v \in V$ ein Eigenvektor von f zum Eigenwert λ. Dann ist

$$\lambda \langle v, v \rangle = \langle \lambda v, v \rangle = \langle f(v), v \rangle = \langle v, f(v) \rangle = \langle v, \lambda v \rangle = \overline{\lambda} \langle v, v \rangle$$

und daraus folgt, dass $\lambda = \overline{\lambda}$, da $\langle v, v \rangle \neq 0$. Damit ist $\lambda \in \mathbb{R}$. □

Sei $f : V \to V$ ein selbstadjungierter Endomorphismus; dann ist jeder Eigenwert von f reell: Für einen euklidischen Vektorraum ist dies trivial der Fall und für einen unitären Vektorraum ist dies die Behauptung von Satz 21.2.

Satz 21.3 *Sei $f : V \to V$ ein selbstadjungierter Endomorphismus und seien λ_1, λ_2 verschiedene Eigenwerte von f. Dann sind die Eigenräume $\mathrm{E}(f, \lambda_1)$ und $\mathrm{E}(f, \lambda_2)$ orthogonal.*

Beweis Seien $v_1 \in \mathrm{E}(f, \lambda_1)$, $v_2 \in \mathrm{E}(f, \lambda_2)$; dann ist

$$\begin{aligned}\lambda_1\langle v_1, v_2\rangle = \langle \lambda_1 v_1, v_2\rangle &= \langle f(v_1), v_2\rangle \\ &= \langle v_1, f(v_2)\rangle = \langle v_1, \lambda_2 v_2\rangle = \overline{\lambda}_2\langle v_1, v_2\rangle = \lambda_2\langle v_1, v_2\rangle\ ,\end{aligned}$$

da $\lambda_2 \in \mathbb{R}$, und daraus folgt, dass $\langle v_1, v_2\rangle = 0$, da $\lambda_1 \neq \lambda_2$. □

Lemma 21.1 *Sei $f : V \to V$ ein selbstadjungierter Endomorphismus und sei U ein Untervektorraum von V mit $f(U) \subset U$. Dann gilt auch $f(U^\perp) \subset U^\perp$.*

Beweis Seien $v \in U^\perp$, $u \in U$; dann gilt $\langle f(v), u\rangle = \langle v, f(u)\rangle = 0$, da $f(u) \in U$. Mit anderen Worten ist $f(v) \perp u$ für alle $v \in U^\perp$, $u \in U$; d.h., $f(U^\perp) \subset U^\perp$. □

Bemerkung Im Gegensatz zu dem entsprechenden Lemma für orthogonale bzw. unitäre Vektorräume (Lemma 19.4) braucht man hier in Lemma 21.1 nicht anzunehmen, dass V endlichdimensional ist.

Sei $f : V \to V$ ein Endomorphismus von V; ein Endomorphismus $g : V \to V$ heißt *zu f adjungiert*, wenn für alle u, $v \in V$

$$\langle g(u), v\rangle = \langle u, f(v)\rangle\ .$$

Also ist f selbstadjungiert genau dann, wenn f zu sich selbst adjungiert ist.

Lemma 21.2 *Sei $f : V \to V$ ein Endomorphismus. Dann gibt es höchstens einen Endomorphismus, der zu f adjungiert ist.*

Beweis Seien g, $h : V \to V$ Endomorphismen, die zu f adjungiert sind. Für jedes $v \in V$ gilt dann

$$\begin{aligned}\langle g(v) - h(v), g(v) - h(v)\rangle &= \langle g(v), g(v) - h(v)\rangle - \langle h(v), g(v) - h(v)\rangle \\ &= \langle v, f(g(v) - h(v))\rangle - \langle v, f(g(v) - h(v))\rangle = 0\end{aligned}$$

und damit ist $g(v) = h(v)$ für alle $v \in V$, d.h., $g = h$. □

Wenn ein (und damit der) zu f adjungierte Endomorphismus existiert, so wird dieser Endomorphismus meistens mit f^{ad} bezeichnet.

Im Folgenden sei V stets endlichdimensional ist mit $\dim V \geq 1$.

Lemma 21.3 *Seien f, $g : V \to V$ Endomorphismen und sei $(v_1, \ldots, v_n)$ eine orthonormale Basis von V. Dann ist g der zu f adjungierte Endomorphismus genau, wenn $\langle g(v_i), v_j\rangle = \langle v_i, f(v_j)\rangle$ für alle $1 \leq i, j \leq n$.*

Beweis Nehme an, dass $\langle g(v_i), v_j\rangle = \langle v_i, f(v_j)\rangle$ für alle $1 \leq i, j \leq n$. Seien nun $u, v \in V$; es gibt also eindeutige Elemente $\lambda_1, \ldots, \lambda_n, \mu_1, \ldots, \mu_m \in \mathbb{K}$, so dass $u = \lambda_1 v_1 + \cdots + \lambda_n v_n$ und $v = \mu_1 v_1 + \cdots + \mu_n v_n$, und dann ist

$$\begin{aligned}\langle u, f(v)\rangle &= \Big\langle \sum_{i=1}^{n} \lambda_i v_i, \sum_{j=1}^{n} \mu_j f(v_j)\Big\rangle \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \overline{\mu}_j \langle v_i, f(v_j)\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \overline{\mu}_j \langle g(v_i), v_j\rangle \\ &= \Big\langle \sum_{i=1}^{n} \lambda_i g(v_i), \sum_{j=1}^{n} \mu_j v_j\Big\rangle = \langle g(u), v\rangle\end{aligned}$$

und daraus folgt, dass g der zu f adjungierte Endomorphismus ist. Natürlich ist die Umkehrung trivial richtig. □

Satz 21.4 *Sei $f : V \to V$ ein Endomorphismus von V. Dann existiert der zu f adjungierte Endomorphismus $f^{\mathsf{ad}} : V \to V$. Ist ferner α eine orthonormale Basis von V und ist A die Matrix von f bezüglich α, so ist A^t, falls $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer Vektorraum ist, bzw. $\overline{A}^t$, falls $(V, \langle\cdot,\cdot\rangle)$ ein unitärer Vektorraum ist, die Matrix von f^{ad} bezüglich der Basis α.*

Beweis Sei $(v_1, \ldots, v_n)$ eine orthonormale Basis von V und $A = (a_{ij})$ die Matrix von f bezüglich $(v_1, \ldots, v_n)$. Nach Lemma 8.2 gibt es dann einen eindeutigen Endomorphismus $g : V \to V$, so dass

$$g(v_j) = \sum_{i=1}^{n} \overline{a}_{ji} v_i = \sum_{i=1}^{n} \overline{a}^t_{ij} v_i$$

für $j = 1, \ldots, n$, und folglich ist A^t bzw. $\overline{A}^t$ die Matrix von g bezüglich der Basis $(v_1, \ldots, v_n)$. Seien $1 \leq j, k \leq n$; dann ist

$$\begin{aligned}\langle g(v_j), v_k\rangle = \Big\langle \sum_{i=1}^{n} \overline{a}^t_{ij} v_i, v_k\Big\rangle &= \sum_{i=1}^{n} \overline{a}^t_{ij} \langle v_i, v_k\rangle = \overline{a}^t_{kj} = \overline{a}_{jk} \\ &= \sum_{i=1}^{n} \overline{a}_{ik} \langle v_j, v_i\rangle = \Big\langle v_j, \sum_{i=1}^{n} a_{ik} v_i\Big\rangle = \langle v_j, f(v_k)\rangle\end{aligned}$$

und damit ist nach Lemma 21.3 g der zu f adjungierte Endomorphismus. Ferner ist A^t bzw. $\overline{A}^t$ die Matrix von g bezüglich $(v_1, \ldots, v_n)$. □

Satz 21.5 *Sei $f : V \to V$ ein Endomorphismus. Dann ist* $\det f^{\mathsf{ad}} = \det f$*, falls $(V, \langle\cdot,\cdot\rangle)$ ein euklidischer Vektorraum ist, und* $\det f^{\mathsf{ad}} = \overline{\det f}$*, falls $(V, \langle\cdot,\cdot\rangle)$ ein unitärer Vektorraum ist.*

Beweis Dies folgt unmittelbar aus Satz 21.4 und Satz 15.4. □

Satz 21.6 *Sei $f : V \to V$ ein Endomorphismus. Dann sind äquivalent:*

(1) f ist orthogonal bzw. unitär.

(2) f ist ein Automorphismus und $f^{-1} = f^{\mathsf{ad}}$.

Beweis Dies ist lediglich eine Umformulierung von Satz 19.3. □

Eine Matrix $A = (a_{ij}) \in \mathrm{M}(n \times n, \mathbb{R})$ heißt *symmetrisch*, wenn $a_{ji} = a_{ij}$ für alle i, j, d.h., wenn $A = A^t$.

Eine Matrix $A = (a_{ij}) \in \mathrm{M}(n \times n, \mathbb{C})$ heißt *Hermitesch*, wenn $a_{ji} = \overline{a}_{ij}$ für alle i, j, d.h., wenn $A = \overline{A}^t$.

Satz 21.7 *Sei $f : V \to V$ ein Endomorphismus und α eine orthonormale Basis von V. Dann ist die Matrix A von f bezüglich α symmetrisch bzw. Hermitesch genau, wenn f selbstadjungiert ist.*

Beweis Dies folgt unmittelbar aus Satz 21.4 und Lemma 8.2. □

Lemma 21.4 *Sei $f : V \to V$ ein selbstadjungierter Endomorphismus. Dann besitzt f einen Eigenwert.*

Beweis Ist $(V, \langle \cdot, \cdot \rangle)$ ein unitärer Vektorraum, dann folgt umittelbar aus Satz 16.4 und Satz 16.5, dass f einen Eigenwert besitzt, da $\chi_f \neq 0$ und $\operatorname{Grad} \chi_f \geq 1$. Nehme also an, dass $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum ist. Sei $(v_1, \ldots, v_n)$ eine orthonormale Basis von V und sei $A \in \mathrm{M}(n \times n, \mathbb{R})$ die Matrix von f bezüglich $(v_1, \ldots, v_n)$. Sei $A_{\mathbb{C}} \in \mathrm{M}(n \times n, \mathbb{C})$ die Matrix A als komplexe Matrix angesehen. Nach Satz 21.7 ist A symmetrisch und damit ist $A_{\mathbb{C}}$ Hermitesch. Daraus folgt nach Satz 21.7, dass der Endomorphismus $\varphi_{A_{\mathbb{C}}} : \mathbb{C}^n \to \mathbb{C}^n$ selbstadjungiert ist. Insbesondere besitzt $\varphi_{A_{\mathbb{C}}}$ einen Eigenwert λ und nach Satz 21.1 ist $\lambda \in \mathbb{R}$. Nun ist λ (per Definition) ein Eigenwert von $A_{\mathbb{C}}$ und daraus ergibt sich, dass λ auch ein Eigenwert von A ist. (Sei $z = x + iy \in \mathbb{C}^n$ ein Eigenvektor von $A_{\mathbb{C}}$ zum Eigenwert λ, wobei x, $y \in \mathbb{R}^n$. Dann ist $A_{\mathbb{C}} z = \lambda z$ und damit $Ax = \lambda x$ und $Ay = \lambda y$, d.h., λ ist ein Eigenwert von A, da nicht beide von x und y gleich 0 sein können.) Nach Satz 9.3 (1) ist dann λ ein Eigenwert von f. □

Satz 21.8 *Sei $f : V \to V$ ein selbstadjungierter Endomorphismus von V und seien $\lambda_1, \ldots, \lambda_m$ die verschiedenen Eigenwerte von f. Dann ist V die orthogonale Summe der Eigenräume $\mathrm{E}(f, \lambda_1), \ldots, \mathrm{E}(f, \lambda_m)$.*

Beweis Sei $U = \mathrm{E}(f, \lambda_1) + \cdots + \mathrm{E}(f, \lambda_m)$; nach Satz 21.3 ist U die orthogonale Summe von $\mathrm{E}(f, \lambda_1), \ldots, \mathrm{E}(f, \lambda_m)$ und also bleibt nur zu zeigen, dass $U = V$. Nehme an, dass $U \neq V$ und setze $W = U^\perp$. Da $f(\mathrm{E}(f, \lambda_j)) \subset \mathrm{E}(f, \lambda_j)$ für jedes j, ist auch $f(U) \subset U$, und daraus folgt nach Lemma 21.1, dass $f(W) \subset W$. Sei $g : W \to W$ die Einschränkung von f auf W. Dann sieht man leicht, dass g selbstadjungiert ist (bezüglich $(W, \langle \cdot, \cdot \rangle_W)$, wobei $\langle \cdot, \cdot \rangle_W$ die Einschränkung von $\langle \cdot, \cdot \rangle$ auf $W \times W$ ist). Da $\dim W \geq 1$, besitzt g nach Lemma 21.4 einen Eigenwert λ. Sei $v \in V$ ein Eigenvektor von g zum Eigenwert λ. Dann ist $f(v) = g(v) = \lambda v$, d.h., λ ist auch ein Eigenwert von f und damit ist $\lambda = \lambda_j$ für ein j und v ist ein Eigenvektor von f zum Eigenwert λ. Also ist $v \in U \cap W$ und dies ist ein Widerspruch, da $U \perp W$. Daraus ergibt sich, dass $U = V$, d.h., V ist die orthogonale Summe der Eigenräume $\mathrm{E}(f, \lambda_1), \ldots, \mathrm{E}(f, \lambda_m)$. □

Satz 21.9 *Sei $f : V \to V$ ein selbstadjungierter Endomorphismus. Dann gibt es eine aus Eigenvektoren von f bestehende orthonormale Basis von V.*

Beweis Seien $\lambda_1, \ldots, \lambda_m$ die verschiedenen Eigenwerte von f und für jedes j sei $(u_1^j, \ldots, u_{k_j}^j)$ eine orthonormale Basis von $\mathrm{E}(f, \lambda_j)$. Dann ist nach Satz 21.8 und Lemma 18.7 $(u_1^1, \ldots, u_{k_1}^1, \ldots, u_1^m, \ldots, u_{k_m}^m)$ eine orthonormale Basis von V und damit eine aus Eigenvektoren von f bestehende orthonormale Basis. □

Satz 21.10 *Jeder Eigenwert einer Hermiteschen Matrix ist reell.*

Beweis Dies folgt unmittelbar aus Satz 21.2, Satz 21.7 und Satz 9.3 (1). □

Satz 21.11 *Sei A symmetrisch bzw. Hermitesch; dann gibt es eine orthogonale bzw. eine unitäre Matrix P, so dass $P^{-1}AP$ eine Diagonalmatrix ist. Insbesondere ist jede symmetrische bzw. jede Hermitesche Matrix diagonalisierbar.*

Beweis Sei $\mathbb{K} = \mathbb{R}$, falls A symmetrisch ist, und $\mathbb{K} = \mathbb{C}$, falls A Hermitesch ist. Nach Satz 21.7 ist der Endomorphismus $\varphi_A : \mathbb{K}^n \to \mathbb{K}^n$ selbstadjungiert (bezüglich $\mathbb{K}^n$ mit dem üblichen Skalarprodukt) und folglich gibt es nach Satz 21.9 eine aus Eigenvektoren von φ_A bestehende orthonormale Basis $(v_1, \ldots, v_n)$ von $\mathbb{K}^n$. Die Matrix D von φ_A bezüglich $(v_1, \ldots, v_n)$ ist also eine Diagonalmatrix. Sei P die Matrix für den Wechsel von der kanonischen Basis $(e_1, \ldots, e_n)$ von $\mathbb{K}^n$ nach $(v_1, \ldots, v_n)$. Nach Satz 20.4 bzw. Satz 20.12 ist P orthogonal bzw. unitär, da $(e_1, \ldots, e_n)$ eine orthonormale Basis von $\mathbb{K}^n$ ist. Da ferner A die Matrix von φ_A bezüglich $(e_1, \ldots, e_n)$ ist, ist nach Satz 8.8 $D = P^{-1}AP$, d.h., $P^{-1}AP$ ist eine Diagonalmatrix. □

Sei A eine Hermitesche Matrix; nach Satz 21.11 gibt es dann eine unitäre Matrix P, so dass $P^{-1}AP$ eine Diagonalmatrix D ist, und die Einträge auf der Diagonalen von D müssen alle reell sein, da sie Eigenwerte von A sind.

Sei A eine symmetrische bzw. eine Hermitesche Matrix; es gibt nun das folgende Verfahren zur Bestimmung einer orthogonalen bzw. einer unitären Matrix P, so dass $P^{-1}AP$ eine Diagonalmatrix ist:

(1) Man bestimme die Nullstellen $\lambda_1, \ldots, \lambda_m$ des charakteristischen Polynoms χ_A. Nach Satz 16.3 sind dann $\lambda_1, \ldots, \lambda_m$ die Eigenwerte von A.

(2) Für jedes $j = 1, \ldots, m$ ist nach Lemma 9.4

$$\mathrm{E}(A, \lambda_j) = \mathrm{Lös}(A - \lambda_j E_n, 0)$$

und also kann mit Hilfe von Satz 9.1 und des Gaußschen Algorithmus eine Basis $(u_1^j, \ldots, u_{p_j}^j)$ von $\mathrm{E}(A, \lambda_j)$ konstruiert werden.

(3) Für jedes j wende das Gram-Schmidtsche Verfahren auf $(u_1^j, \ldots, u_{p_j}^j)$ an, um eine orthonormale Basis $(v_1^j, \ldots, v_{p_j}^j)$ von $\mathrm{E}(A, \lambda_j)$ zu konstruieren.

(4) Sei P die Matrix, die $w_1, \ldots, w_n$ als Spalten hat, wobei

$$(w_1, \ldots, w_n) = (v_1^1, \ldots, v_{p_1}^1, v_1^2, \ldots, v_{p_2}^2, \ldots, v_1^m, \ldots, v_{p_m}^m) \; .$$

Dann ist P orthogonal bzw. unitär und $P^{-1}AP = D$, wobei D die Diagonalmatrix ist mit den Einträgen $\lambda_1, \ldots, \lambda_1, \ldots, \lambda_m, \ldots, \lambda_m$ auf der Diagonalen, wobei für jedes j der Eigenwert λ_j genau p_j-mal vorkommt.

Im Folgenden sei $(V, \langle \cdot, \cdot \rangle)$ ein endlichdimensionaler unitärer Vektorraum. Eine Klasse Endomorphismen, die normalen Endomorphismen, wird nun eingeführt, und es wird gezeigt, dass ein Endomorphismus f genau dann normal ist, wenn es eine aus Eigenvektoren von f bestehende orthonormale Basis von V gibt.

Ein Endomorphismus $f : V \to V$ von V heißt *normal*, wenn $f \circ f^{\mathsf{ad}} = f^{\mathsf{ad}} \circ f$. Insbesondere ist ein unitärer Endomorphismus f normal, da nach Satz 21.6

$$f \circ f^{\mathsf{ad}} = f \circ f^{-1} = \mathrm{id}_V = f^{-1} \circ f = f^{\mathsf{ad}} \circ f \; .$$

Ferner ist ein selbstadjungierter Endomorphismus f auch normal, da

$$f \circ f^{\mathsf{ad}} = f \circ f = f^{\mathsf{ad}} \circ f \; .$$

Lemma 21.5 *Sei $f : V \to V$ ein normaler Endomorphismus. Dann gilt:*

(1) $\|f(v)\| = \|f^{\mathsf{ad}}(v)\|$ *für alle* $v \in V$.

(2) $\mathrm{Kern}\, f = \mathrm{Kern}\, f^{\mathsf{ad}}$.

(3) $\mathrm{Kern}\,(f - \lambda\, \mathrm{id}_V) = \mathrm{Kern}\,(f^{\mathsf{ad}} - \overline{\lambda}\, \mathrm{id}_V)$ *für alle* $\lambda \in \mathbb{C}$.

(4) $\lambda \in \mathbb{C}$ *ist ein Eigenwert von* f *genau dann, wenn* $\overline{\lambda}$ *ein Eigenwert von* f^{ad} *ist.*

(5) $\mathrm{E}(f,\lambda) = \mathrm{E}(f^{\mathsf{ad}},\overline{\lambda})$ *für jeden Eigenwert* λ *von* f*.*

(6) Sind λ_1 *und* λ_2 *Eigenwerte von* f *mit* $\lambda_1 \neq \lambda_2$*, so sind die Eigenräume* $\mathrm{E}(f,\lambda_1)$ *und* $\mathrm{E}(f,\lambda_2)$ *orthogonal.*

Beweis (1) Für jedes $v \in V$ gilt

$$\begin{aligned}\|f(v)\|^2 &= \langle f(v), f(v)\rangle = \langle f^{\mathsf{ad}}(f(v)), v\rangle = \langle f(f^{\mathsf{ad}}(v)), v\rangle \\ &= \overline{\langle v, f(f^{\mathsf{ad}}(v))\rangle} = \overline{\langle f^{\mathsf{ad}}(v), f^{\mathsf{ad}}(v)\rangle} = \langle f^{\mathsf{ad}}(v), f^{\mathsf{ad}}(v)\rangle = \|f^{\mathsf{ad}}(v)\|^2 .\end{aligned}$$

(2) Nach (1) ist

$$\operatorname{Kern} f = \{v \in V : \|f(v)\| = 0\} = \{v \in V : \|f^{\mathsf{ad}}(v)\| = 0\} = \operatorname{Kern} f^{\mathsf{ad}} .$$

(3) Es gilt $(f - \lambda \operatorname{id}_V)^{\mathsf{ad}} = f^{\mathsf{ad}} - \overline{\lambda}\operatorname{id}_V$, da

$$\begin{aligned}\langle (f^{\mathsf{ad}} - \overline{\lambda}\operatorname{id}_V)(u), v\rangle &= \langle f^{\mathsf{ad}}(u), v\rangle - \overline{\lambda}\langle u, v\rangle \\ &= \langle u, f(v)\rangle - \langle u, \lambda v\rangle = \langle u, (f - \lambda \operatorname{id}_V)(v)\rangle .\end{aligned}$$

Ferner ist $f - \lambda \operatorname{id}_V$ normal, da

$$\begin{aligned}(f - \lambda \operatorname{id}_V) \circ (f - \lambda \operatorname{id}_V)^{\mathsf{ad}} &= (f - \lambda \operatorname{id}_V) \circ (f^{\mathsf{ad}} - \overline{\lambda}\operatorname{id}_V) \\ &= f \circ f^{\mathsf{ad}} - \lambda f^{\mathsf{ad}} - \overline{\lambda} f + \lambda\overline{\lambda}\operatorname{id}_V = f^{\mathsf{ad}} \circ f - \overline{\lambda} f - \lambda f^{\mathsf{ad}} + \overline{\lambda}\lambda \operatorname{id}_V \\ &= (f^{\mathsf{ad}} - \overline{\lambda}\operatorname{id}_V) \circ (f - \lambda \operatorname{id}_V) = (f - \lambda \operatorname{id}_V)^{\mathsf{ad}} \circ (f - \lambda \operatorname{id}_V) .\end{aligned}$$

Daraus ergibt sich nach (2), dass

$$\operatorname{Kern}(f - \lambda \operatorname{id}_V) = \operatorname{Kern}(f - \lambda \operatorname{id}_V)^{\mathsf{ad}} = \operatorname{Kern}(f^{\mathsf{ad}} - \overline{\lambda}\operatorname{id}_V) .$$

(4) Sei $\lambda \in \mathbb{C}$; dann ist λ ein Eigenwert von f genau, wenn $\operatorname{Kern}(f - \lambda \operatorname{id}_V) \neq \{0\}$ und $\overline{\lambda}$ ist ein Eigenwert von f^{ad} genau dann, wenn $\operatorname{Kern}(f^{\mathsf{ad}} - \overline{\lambda}\operatorname{id}_V) \neq \{0\}$. Aber nach (3) ist $\operatorname{Kern}(f - \lambda \operatorname{id}_V) = \operatorname{Kern}(f^{\mathsf{ad}} - \overline{\lambda}\operatorname{id}_V)$.

(5) Dies folgt unmittelbar aus (3).

(6) Seien λ_1, λ_2 verschiedene Eigenwerte von f und $v_1 \in \mathrm{E}(f,\lambda_1)$, $v_2 \in \mathrm{E}(f,\lambda_2)$. Nach (5) gilt auch $v_1 \in \mathrm{E}(f^{\mathsf{ad}},\overline{\lambda}_1)$ und folglich ist

$$\overline{\lambda}_2\langle v_1, v_2\rangle = \langle v_1, \lambda_2 v_2\rangle = \langle v_1, f(v_2)\rangle = \langle f^{\mathsf{ad}}(v_1), v_2\rangle = \langle \overline{\lambda}_1 v_1, v_2\rangle = \overline{\lambda}_1\langle v_1, v_2\rangle .$$

Damit ist $\langle v_1, v_2\rangle = 0$, d.h., $v_1 \perp v_2$, da $\overline{\lambda}_1 \neq \overline{\lambda}_2$. □

Satz 21.12 *Sei $f : V \to V$ ein Endomorphismus. Dann sind die folgenden drei Aussagen äquivalent:*

(1) f ist normal.

(2) V ist die orthogonale Summe der Eigenräume $\mathrm{E}(f, \lambda_1), \ldots, \mathrm{E}(f, \lambda_m)$, wobei $\lambda_1, \ldots, \lambda_m$ die verschiedenen Eigenwerte von f sind.

(3) Es gibt eine aus Eigenvektoren von f bestehende orthonormale Basis von V.

Beweis (2) $\Rightarrow$ (3): Dies ist identisch mit dem Beweis für Satz 21.9.

(3) $\Rightarrow$ (1): Sei $(v_1, \ldots, v_n)$ eine aus Eigenvektoren von f bestehende orthonormale Basis von V, wobei v_j ein Eigenvektor zum Eigenwert λ_j ist für jedes j. Dann ist die Matrix von f bezüglich $(v_1, \ldots, v_n)$ die Diagonalmatrix D mit Einträgen $\lambda_1, \ldots, \lambda_n$ auf der Diagonalen. Daraus folgt nach Satz 21.4, dass $\overline{D}^t$ die Matrix von f^{ad} bezüglich $(v_1, \ldots, v_n)$ ist, und damit ist $D\overline{D}^t$ bzw. $\overline{D}^t D$ die Matrix von $f \circ f^{\mathsf{ad}}$ bzw. $f^{\mathsf{ad}} \circ f$ bezüglich $(v_1, \ldots, v_n)$. Aber $\overline{D}^t$ ist die Diagonalmatrix mit Einträgen $\overline{\lambda}_1, \ldots, \overline{\lambda}_n$ auf der Diagonalen und folglich ist $D\overline{D}^t = \overline{D}^t D = D'$, wobei D' die Diagonalmatrix mit Einträgen $|\lambda_1|^2, \ldots, |\lambda_n|^2$ auf der Diagonalen ist. Also folgt aus Lemma 8.2, dass $f \circ f^{\mathsf{ad}} = f^{\mathsf{ad}} \circ f$, d.h., f ist normal.

(1) $\Rightarrow$ (2): Seien $\lambda_1, \ldots, \lambda_m$ die verschiedenen Eigenwerte von f; dann ist nach Lemma 21.5 (6) die Summe $U = \mathrm{E}(f, \lambda_1) + \cdots + \mathrm{E}(f, \lambda_m)$ eine orthogonale Summe und also genügt es zu zeigen, dass $U = V$. Nehme an, dass $U \neq V$ und setze $W = U^\perp$. Sei $w \in W$; dann gilt $\langle w, u \rangle = 0$ für jedes $u \in \mathrm{E}(f, \lambda_j)$ und folglich ist

$$0 = \lambda_j \langle w, u \rangle = \langle w, \overline{\lambda}_j u \rangle = \langle w, f^{\mathsf{ad}}(u) \rangle = \langle f(w), u \rangle \, ,$$

da nach Lemma 21.5 (4) $\mathrm{E}(f, \lambda_j) = \mathrm{E}(f^{\mathsf{ad}}, \overline{\lambda}_j)$. Dies bedeutet, dass $f(w) \perp u$ für alle $u \in \mathrm{E}(f, \lambda_j)$, und daher ist $f(w) \perp u$ für alle $w \in W$ und alle $u \in U$, d.h., $f(W) \subset W$. Sei g die Einschränkung von f auf W. Da $\dim W \geq 1$, ist $\chi_g \neq 0$ ein Polynom mit Grad $\chi_g \geq 1$ und folglich gibt es nach Satz 16.4 und Satz 16.5 einen Eigenwert $\lambda \in \mathbb{C}$ von g. Sei $v \in V$ ein Eigenvektor von g zum Eigenwert λ. Dann ist $f(v) = g(v) = \lambda v$, d.h., λ ist auch ein Eigenwert von f und damit ist $\lambda = \lambda_j$ für ein j und v ist ein Eigenvektor von f zum Eigenwert λ. Also ist $v \in U \cap W$ und dies ist ein Widerspruch, da $U \perp W$. Daraus ergibt sich, dass $U = V$, d.h., V ist die orthogonale Summe der Eigenräume $\mathrm{E}(f, \lambda_1), \ldots, \mathrm{E}(f, \lambda_m)$. □

Satz 21.12 liefert einen alternativen Beweis für Satz 19.6 und für Satz 21.9.

22 Bilinearformen

Im Folgenden sei K ein Körper und seien V und W Vektorräume über K.

Eine Abbildung $s : V \times W \to K$ heißt *Bilinearform*, wenn für alle v_1, v_2, $v \in V$, w_1, w_2, $w \in W$ und alle λ_1, $\lambda_2 \in K$ gilt:

$$s(\lambda_1 v_1 + \lambda_2 v_2, w) = \lambda_1 s(v_1, w) + \lambda_2 s(v_2, w) \; ,$$
$$s(v, \lambda_1 w_1 + \lambda_2 w_2) = \lambda_1 s(v, w_1) + \lambda_2 s(v, w_2) \; .$$

Beispiele (1) Ist $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum, so ist das Skalarprodukt $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ eine Bilinearform.

(2) Sei $A = (a_{ij}) \in \mathrm{M}(m \times n, K)$ und definiere $s : K^m \times K^n \to K$ durch

$$s(x, y) = \sum_{j=1}^{m} \sum_{k=1}^{n} a_{jk} x_j y_k \; ,$$

wobei $x = (x_1, \ldots, x_m)$ und $y = (y_1, \ldots, y_n)$; dann ist s eine Bilinearform.

(3) Für a, $b \in \mathbb{R}$ mit $a < b$ sei $\mathrm{C}([a, b], \mathbb{R})$ der reelle Vektorraum aller stetigen Abbildungen von $[a, b]$ nach $\mathbb{R}$. Sei $K : [a, b] \times [c, d] \to \mathbb{R}$ stetig und definiere $s : \mathrm{C}([a, b], \mathbb{R}) \times \mathrm{C}([c, d], \mathbb{R}) \to \mathbb{R}$ durch

$$s(f, g) = \int_a^b \int_c^d K(x, y) f(x) g(y) \, dx \, dy \; ;$$

dann ist s eine Bilinearform.

Für jede Bilinearform $s : V \times W \to K$ gibt es die *transponierte* Bilinearform $s^{\mathrm{t}} : W \times V \to K$, die für alle $w \in W$, $v \in V$ gegeben ist durch

$$s^{\mathrm{t}}(w, v) = s(v, w) \; .$$

Seien nun V und W endlichdimensional mit $\dim V = m \geq 1$ und $\dim W = n \geq 1$.

Sei $s : V \times W \to K$ eine Bilinearform, sei $(v_1, \ldots, v_m)$ eine Basis von V und $(w_1, \ldots, w_n)$ eine Basis von W. Definiere eine Matrix $A = (a_{ij}) \in \mathrm{M}(m \times n, K)$ durch $a_{ij} = s(v_i, w_j)$; dann heißt A die *Matrix von s bezüglich $(v_1, \ldots, v_m)$ und $(w_1, \ldots, w_n)$*. Ist A die Matrix von s bezüglich $(v_1, \ldots, v_m)$ und $(w_1, \ldots, w_n)$, so ist die transponierte Matrix $A^{\mathrm{t}} \in \mathrm{M}(n \times m, K)$ die Matrix von s^{t} bezüglich $(w_1, \ldots, w_n)$ und $(v_1, \ldots, v_m)$.

Lemma 22.1 *Sei $(v_1, \ldots, v_m)$ eine Basis von V und $(w_1, \ldots, w_n)$ eine Basis von W. Dann gibt es zu jeder Matrix $A \in \mathrm{M}(m \times n, K)$ genau eine Bilinearform s, so dass A die Matrix von s bezüglich $(v_1, \ldots, v_m)$ und $(w_1, \ldots, w_n)$ ist.*

Beweis Übung. □

Im Folgenden sei $s : V \times W \to K$ eine Bilinearform.

Lemma 22.2 *Seien α, α' zwei Basen von V und β, β' zwei Basen von W. Sei A bzw. B die Matrix von s bezüglich α und β bzw. bezüglich α' und β'. Dann gilt $B = P^t AQ$, wobei P die Matrix für den Wechsel von α nach α' und Q die Matrix für den Wechsel von β nach β' ist.*

Beweis Setze $A = (a_{ij})$, $B = (b_{ij})$, $P = (p_{ij})$, $Q = (q_{ij})$; seien $\alpha = (v_1, \dots, v_m)$, $\alpha' = (v'_1, \dots, v'_m)$, $\beta = (w_1, \dots, w_n)$ und $\beta' = (w'_1, \dots, w'_n)$. Dann gilt

$$\begin{aligned} b_{ij} &= s(v'_i, w'_j) = s\Big(\sum_{k=1}^{m} p_{ki} v_k, \sum_{\ell=1}^{n} q_{\ell j} w_\ell\Big) \\ &= \sum_{k=1}^{m}\sum_{\ell=1}^{n} p_{ki} q_{\ell j}, s(v_k, w_\ell) = \sum_{k=1}^{m}\sum_{\ell=1}^{n} p_{ki} q_{\ell j} a_{k\ell} = \sum_{k=1}^{m}\sum_{\ell=1}^{n} p^t_{ik} a_{k\ell} q_{\ell j} \end{aligned}$$

für alle $1 \le i \le m$, $1 \le j \le n$ und folglich ist $B = P^t AQ$. □

Sind $P \in \mathrm{M}(m \times m, K)$ und $Q \in \mathrm{M}(n \times n, K)$ invertierbare Matrizen, so gilt nach Lemma 7.4 (2) und (3), dass $\operatorname{rang} P^t AQ = \operatorname{rang} A$ für jedes $A \in \mathrm{M}(m \times n, K)$. Ist also A die Matrix von s bezüglich Basen α und β, so folgt aus Lemma 22.2, dass $\operatorname{rang} A$ allein durch s bestimmt wird, und folglich setzt man $\operatorname{rang} s = \operatorname{rang} A$; $\operatorname{rang} s$ heißt natürlich der *Rang* von s. Es gilt $0 \le \operatorname{rang} s \le \min\{m, n\}$.

Da $\operatorname{rang} A^{\mathrm{t}} = \operatorname{rang} A$ für jedes $A \in \mathrm{M}(m \times n, K)$, ist $\operatorname{rang} s^{\mathrm{t}} = \operatorname{rang} s$.

Erinnerung: Sei r mit $0 \le r \le \min\{m, n\}$; die folgende $m \times n$ Matrix

$$E^r_{m,n} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \leftarrow r\text{-te Zeile}$$

wurde die *Sub-Einheitsmatrix vom Rang r* genannt. Also ist $E^r_{m,n} = (\delta_{ij})$, wobei

$$\delta_{ij} = \begin{cases} 1 & \text{falls } i = j \text{ und } 1 \le i \le r\ , \\ 0 & \text{sonst}\ , \end{cases}$$

und $\operatorname{rang} E^r_{m,n} = r$.

Satz 22.1 *Sei $r = \operatorname{rang} s$; dann gibt es eine Basis α von V und eine Basis β von W, so dass $E^r_{m,n}$ die Matrix von s bezüglich α und β ist.*

Beweis Sei α' eine beliebige Basis von V und β' eine beliebige Basis von W, und sei A die Matrix von s bezüglich α' und β'. Da $\operatorname{rang} A = r$, gibt es nach Satz 7.6 invertierbare Matrizen $R \in \mathrm{M}(m \times m, K)$ und $Q \in \mathrm{M}(n \times n, K)$, so dass $RAQ = E^r_{m,n}$; setze $P = R^t$. Sei nun α bzw. β die Basis von V bzw. W, so dass P bzw. Q die Matrix für den Wechsel von α' nach α bzw. für den Wechsel von β' nach β ist. Nach Lemma 22.2 ist dann $E^r_{m,n} = P'AQ = P^tAQ$ die Matrix von s bezüglich α und β. □

Definiere nun $\mathrm{L}(s) \subset V$ und $\mathrm{R}(s) \subset W$ durch

$$\begin{aligned} \mathrm{L}(s) &= \{v \in V : s(v, w) = 0 \text{ für alle } w \in W\} , \\ \mathrm{R}(s) &= \{w \in W : s(v, w) = 0 \text{ für alle } v \in V\} . \end{aligned}$$

Es ist klar, dass $\mathrm{L}(s)$ ein Untervektorraum von V und $\mathrm{R}(s)$ ein Untervektorraum von W ist. Ferner gilt $\mathrm{L}(s^{\mathrm{t}}) = \mathrm{R}(s)$ und $\mathrm{R}(s^{\mathrm{t}}) = \mathrm{L}(s)$.

Satz 22.2 (Dimensionsformel) *Es gilt*

$$\dim V - \dim \mathrm{L}(s) = \operatorname{rang} s = \dim W - \dim \mathrm{R}(s) .$$

Beweis Sei $r = \operatorname{rang} s$; nach Satz 22.1 gibt es dann eine Basis $\alpha = (v_1, \ldots, v_m)$ von V und eine Basis $\beta = (w_1, \ldots, w_n)$ von W, so dass $E^r_{m,n}$ die Matrix von s bezüglich α und β ist. Es gilt also

$$s(v_j, w_k) = \begin{cases} 1 & \text{falls } j = k \text{ und } 1 \le j \le r , \\ 0 & \text{sonst} , \end{cases}$$

Da $(w_1, \ldots, w_n)$ eine Basis von W ist, gilt

$$\mathrm{L}(s) = \{v \in V : s(v, w_k) = 0 \text{ für jedes } k = 1, \ldots, n\} ,$$

und ist $v \in V$ mit $v = \lambda_1 v_1 + \cdots + \lambda_m v_m$, so gilt

$$\begin{aligned} s(v, w_k) &= s(\lambda_1 v_1 + \cdots + \lambda_m v_m, w_k) = \lambda_1 s(v_1, w_k) + \cdots + \lambda_m s(v_m, w_k) \\ &= \begin{cases} \lambda_k & \text{falls } k \le r , \\ 0 & \text{sonst} . \end{cases} \end{aligned}$$

Daraus ergibt sich, dass $\mathrm{L}(s) = \mathrm{L}(v_{r+1}, \ldots, v_m)$ und folglich ist

$$\dim \mathrm{L}(s) = \dim \mathrm{L}(v_{r+1}, \ldots, v_m) = m - r = \dim V - \operatorname{rang} s .$$

Genauso gilt dann

$$\dim \mathrm{R}(s) = \dim \mathrm{L}(s^{\mathrm{t}}) = \dim W - \operatorname{rang} s^{\mathrm{t}} = \dim W - \operatorname{rang} s \ . \ \square$$

Ab jetzt wird nur der Fall mit $V = W$ betrachtet, und also ist s eine bilineare Abbildung von $V \times V$ nach K. Hier wird *eine* Basis verwendet: Ist $(v_1, \ldots, v_n)$ eine Basis von V, so wird $(a_{jk}) \in \mathrm{M}(n \times n, K)$ mit $a_{jk} = s(v_j, v_k)$ die Matrix von s bezüglich $(v_1, \ldots, v_n)$ genannt.

Die Bilinearform $s : V \times V \to K$ heißt *symmetrisch*, wenn $s^{\mathrm{t}} = s$, d.h., wenn $s(v, w) = s(w, v)$ für alle v, $w \in V$. Ist A die Matrix von s bezüglich einer Basis von V, so ist s symmetrisch genau dann, wenn A symmetrisch ist, d.h., genau dann, wenn $A^{\mathrm{t}} = A$.

Sind α, α' zwei Basen von V und ist A bzw. B die Matrix von s bezüglich α bzw. bezüglich α', so gilt nach Lemma 22.2, dass $B = P^t AP$, wobei P die Matrix für den Wechsel von α nach α' ist.

Nach Satz 22.2 gilt nun

$$\dim \mathrm{L}(s) = \dim \mathrm{R}(s) = \dim V - \operatorname{rang} s \ ,$$

und insbesondere ist $\operatorname{rang} s = \dim V$ genau dann, wenn $\mathrm{L}(s) = \mathrm{R}(s) = \{0\}$. In diesem Fall sagt man, dass s *nicht ausgeartet* ist. Ist A die Matrix von s bezüglich einer Basis von V, so ist s nicht ausgeartet genau dann, wenn A invertierbar ist.

Satz 22.3 *Sei s nicht ausgeartet und $\hat{s} : V \times V \to K$ eine beliebige Bilinearform. Dann gibt es eindeutige Endomorphismen f, $g : V \to V$, so dass*

$$\hat{s}(v, w) = s(v, g(w)) = s(f(v), w)$$

für alle v, $w \in V$. Ist ferner α eine Basis von V, $A \in \mathrm{M}(n \times n, K)$ die Matrix von s und $B \in \mathrm{M}(n \times n, K)$ die Matrix von $\hat{s}$ bezüglich α, so ist $(BA^{-1})^{\mathrm{t}}$ die Matrix von f und $A^{-1}B$ die Matrix von g bezüglich α.

Beweis Sei $\alpha = (v_1, \ldots, v_n)$ eine Basis von V und sei $A = (a_{ij})$ die Matrix von s und $B = (b_{ij})$ die Matrix von $\hat{s}$ bezüglich α. Da s nicht ausgeartet ist, ist A invertierbar. Sei $f : V \to V$ ein Endomorphismus und sei $C = (c_{ij})$ die Matrix von f bezüglich α. Dann gilt

$$\begin{aligned} s(f(v_j), v_k) &= s(c_{1j}v_1 + \cdots + c_{nj}v_n, v_k) \\ &= c_{1j}s(v_1, v_k) + \cdots + c_{nj}s(v_n, v_k) = c_{1j}a_{1k} + \cdots + c_{nj}a_{nk} = (C^{\mathrm{t}}A)_{jk} \end{aligned}$$

und damit gilt $\hat{s}(v_j, v_k) = b_{jk} = s(f(v_j), v_k)$ für alle j, k genau dann, wenn $B = C^{\mathrm{t}}A$. Folglich gilt $\hat{s}(v, w) = s(f(v), w)$ für alle v, $w \in V$ genau dann, wenn

$C = (BA^{-1})^{\mathrm{t}}$. Aber es gibt einen eindeutigen Endomorphismus $f : V \to V$, für den $(BA^{-1})^{\mathrm{t}}$ die Matrix von f bezüglich α ist und daher gibt es einen eindeutigen Endomorphismus $f : V \to V$, so dass $\hat{s}(v, w) = s(f(v), w)$ für alle v, $w \in V$.

Sei nun $g : V \to V$ ein weiterer Endomorphismus mit $C' = (c'_{ij})$ die Matrix von g bezüglich α. Dann gilt

$$\begin{aligned} s(v_j, g(v_k)) &= s(v_j, c'_{1k}v_1 + \cdots + c'_{nk}v_n) \\ &= c'_{1k}s(v_j, v_1) + \cdots + c'_{nk}s(v_j, v_n) = c'_{1k}a_{j1} + \cdots + c'_{nk}a_{jn} = (AC')_{jk} \end{aligned}$$

für alle j, k und also gilt $\hat{s}(v, w) = s(v, g(w))$ für alle v, $w \in V$ genau dann, wenn $C' = A^{-1}B$. Folglich gibt es einen eindeutigen Endomorphismus $g : V \to V$, so dass $\hat{s}(v, w) = s(v, g(w))$ für alle v, $w \in V$. □

Im Folgenden sei nun $(V, \langle \cdot, \cdot \rangle)$ ein endlichdimensionaler euklidischer Vektorraum mit $\dim V = n \geq 1$. Die Bilinearform $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ ist nicht ausgeartet: Ist $(v_1, \ldots, v_n)$ eine orthonormale Basis von V, so ist die Matrix von $\langle \cdot, \cdot \rangle$ bezüglich $(v_1, \ldots, v_n)$ per Definition die Einheitsmatrix E_n und $\operatorname{rang} E_n = n$.

Sei $s : V \times V \to \mathbb{R}$ eine Bilinearform.

Satz 22.4 *Es gibt einen eindeutigen Endomorphismus $f : V \to V$, so dass $s(u, v) = \langle u, f(v) \rangle$ für alle u, $v \in V$. Ist ferner α eine orthonormale Basis von V und ist A die Matrix von f bezüglich α, so ist A auch die Matrix von s bezüglich α.*

Beweis Dies ist ein Spezialfall von Satz 22.3. □

Lemma 22.3 *Sei f der eindeutige Endomorphismus mit $s(u, v) = \langle u, f(v) \rangle$ für alle u, $v \in V$. Dann ist s symmetrisch genau, wenn f selbstadjungiert ist.*

Beweis Ist s symmetrisch, dann gilt für alle u, $v \in V$

$$\langle u, f(v) \rangle = s(u, v) = s(v, u) = \langle v, f(u) \rangle = \langle f(u), v \rangle$$

und damit ist f selbstadjugiert. Ist umgekehrt f selbstadjugiert, so ist

$$s(u, v) = \langle u, f(v) \rangle = \langle f(u), v \rangle = \langle v, f(u) \rangle = s(v, u)$$

für alle u, $v \in V$ und also ist s symmetrisch. □

Satz 22.5 *Sei s symmetrisch. Dann gibt es eine orthonormale Basis α von V, so dass die Matrix von s bezüglich α eine Diagonalmatrix D ist. Ist $\alpha = (v_1, \ldots, v_n)$ und sind $\lambda_1, \ldots, \lambda_n$ die Einträge auf der Diagonalen, so gilt für alle u, $v \in V$:*

$$s(u, v) = \sum_{j=1}^{n} \lambda_j \langle u, v_j \rangle \langle v, v_j \rangle \ .$$

Beweis Sei $f : V \to V$ der eindeutige Endomorphismus mit $s(u, v) = \langle u, f(v)\rangle$ für alle u, $v \in V$. Nach Lemma 22.3 ist f selbstadjungiert, da s symmetrisch ist, und folglich gibt es nach Satz 21.9 eine aus Eigenvektoren von f bestehende orthonormale Basis $(v_1, \ldots, v_n)$ von V. Sei D die Diagonalmatrix mit Einträgen $\lambda_1, \ldots, \lambda_n$ auf der Diagonalen, wobei v_j ein Eigenvektor zum Eigenwert λ_j ist für jedes j. Dann ist D die Matrix von f bezüglich $(v_1, \ldots, v_n)$ und damit ist nach Satz 22.4 D auch die Matrix von s bezüglich $(v_1, \ldots, v_n)$. Nach Satz 18.5 ist nun

$$\begin{aligned} s(u, v) &= s\Big(\sum_{j=1}^{n}\langle u, v_j\rangle v_j, \sum_{k=1}^{n}\langle v, v_k\rangle v_k\Big) = \sum_{j=1}^{n}\sum_{k=1}^{n}\langle u, v_j\rangle\langle v, v_k\rangle s(v_j, v_k) \\ &= \sum_{j=1}^{n}\lambda_j\langle u, v_j\rangle\langle v, v_j\rangle \end{aligned}$$

für alle u, $v \in V$, da $s(v_j, v_k) = \begin{cases} \lambda_j & \text{falls } j = k\,, \\ 0 & \text{falls } j \neq k\,. \end{cases}$ □

Sei s symmetrisch. Nach Satz 22.5 gibt es eine orthonormale Basis α von V, so dass die Matrix von s bezüglich α eine Diagonalmatrix D ist. Sei $\alpha = (v_1, \ldots, v_n)$ und seien $\lambda_1, \ldots, \lambda_n$ die Einträge auf der Diagonalen von D; sei $\sigma : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ gegeben durch

$$\sigma(x, y) = \sum_{j=1}^{n}\lambda_j x_j y_j$$

für alle $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$. Dann ist σ eine symmetrische Bilinearform auf $\mathbb{R}^n$ und nach Satz 22.5 gilt

$$s(\psi(x), \psi(y)) = \sigma(x, y)$$

für alle x, $y \in \mathbb{R}^n$, wobei $\psi : \mathbb{R}^n \to V$ der durch

$$\psi((x_1, \ldots, x_n)) = x_1 v_1 + \cdots + x_n v_n$$

gegebene Isomorphismus ist, (da $\langle\psi((x_1, \ldots, x_n)), v_j\rangle = x_j$ für jedes j). Daraus ergibt sich, dass für alle u, $v \in V$

$$s(u, v) = \sigma(\psi^{-1}(u), \psi^{-1}(v))\,.$$

Betrachte nun den speziellen Fall mit $V = \mathbb{R}^n$ (mit dem üblichen Skalarprodukt). Sei $A = (a_{ij}) \in \mathrm{M}(n \times n, \mathbb{R})$ eine symmetrische Matrix und sei $s : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ die symmetrische Bilinearform auf $\mathbb{R}^n$, die gegeben ist durch

$$s(x, y) = \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk} x_j y_k$$

für alle $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$. Natürlich ist A die Matrix von s bezüglich der kanonischen Basis $(e_1, \ldots, e_n)$ von $\mathbb{R}^n$. Ferner ist

$$s(x,y) = \sum_{j=1}^{n} \sum_{k=1}^{n} a_{jk} x_j y_k = \sum_{j=1}^{n} x_j \Big(\sum_{k=1}^{n} a_{jk} y_k \Big) = x \cdot \varphi_A(y)$$

für alle $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$, und damit ist φ_A der eindeutige Endomorphismus von $\mathbb{R}^n$ mit $s(x,y) = x \cdot \varphi_A(y)$ für alle x, $y \in \mathbb{R}^n$.

Für alle $(x_1, \ldots, x_n)$, $(y_1, \ldots, y_n) \in \mathbb{R}^n$ ist

$$\sum_{j=1}^{n} \sum_{k=1}^{n} a_{jk} x_j y_k = \sum_{j=1}^{n} x_j \sum_{k=1}^{n} a_{jk} y_k = (x_1, \ldots, x_n) A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

und folglich gilt $s(x,y) = x^t A y$ für alle x, $y \in \mathbb{R}^n$, wenn die Elemente von $\mathbb{R}^n$ als Spaltenvektoren betrachtet werden.

Satz 22.6 *Es gibt einen orthogonalen Endomorphismus $\psi : \mathbb{R}^n \to \mathbb{R}^n$ von $\mathbb{R}^n$ und Elemente $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, so dass*

$$s(x,y) = \sigma(\psi(x), \psi(y))$$

für alle x, $y \in \mathbb{R}^n$, wobei $\sigma : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ wieder die durch

$$\sigma(x,y) = \sum_{j=1}^{n} \lambda_j x_j y_j$$

für $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$ definierte symmetrische Bilinearform auf $\mathbb{R}^n$ ist.

Beweis Nach Satz 21.11 gibt es eine orthogonale Matrix $P \in \mathrm{M}(n \times n, \mathbb{R})$, so dass $P^{-1}AP$ eine Diagonalmatrix D ist. Seien $\lambda_1, \ldots, \lambda_n$ die Einträge auf der Diagonalen von D. Dann gilt

$$\begin{aligned} s(x,y) = x^t A y = x^t (PDP^{-1}) y &= x^t (PDP^t) y = (x^t P) D (P^t y) \\ &= (P^t x)^t D (P^t y) = \sigma(P^t x, P^t y) = \sigma(\psi(x), \psi(y)) \end{aligned}$$

für alle x, $y \in \mathbb{R}^n$, wobei $\psi = \varphi_{P^t}$. Aber $P^t = P^{-1}$ ist eine orthogonale Matrix und daraus folgt nach Satz 20.1, dass ψ ein orthogonaler Endomorphismus von $\mathbb{R}^n$ ist. □

Satz 22.6 ist eigentlich nur ein Spezialfall von Satz 22.5: Sei $P \in \mathrm{M}(n \times n, \mathbb{R})$ die orthogonale Matrix, die im Beweis für Satz 22.6 vorkommt, und seien $u_1, \ldots, u_n$

die Spalten von P. Dann ist $(u_1, \ldots, u_n)$ eine orthonormale Basis von $\mathbb{R}^n$ und D ist die Matrix von s bezüglich $(u_1, \ldots, u_n)$.

Es wird nun eine 'quadratische' Abbildung auf $\mathbb{R}^n$ untersucht: Seien a, b_j $c_{k\ell}$, d_j, $1 \leq j \leq n$, $1 \leq k < \ell \leq n$, Elemente aus $\mathbb{R}$ und definiere $q : \mathbb{R}^n \to \mathbb{R}$ durch

$$q((x_1, \ldots, x_n)) = \sum_{j=1}^{n} d_j x_j^2 + \sum_{i<j} c_{ij} x_i x_j + \sum b_j x_j + a \; .$$

Sei $\gamma : \mathbb{R}^n \to \mathbb{R}^{n+1}$ die Abbildung, die gegeben ist durch

$$\gamma((x_1, \ldots, x_n)) = (x_1, \ldots, x_n, 1)$$

für alle $(x_1, \ldots, x_n) \in \mathbb{R}^n$.

Satz 22.7 *Es gibt einen orthogonalen Endomorphismus $\psi : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ von $\mathbb{R}^{n+1}$ und Elemente $\lambda_1, \ldots, \lambda_{n+1} \in \mathbb{R}$, so dass*

$$q = \tau \circ \psi \circ \gamma \; ,$$

wobei $\tau : \mathbb{R}^{n+1} \to \mathbb{R}$ die 'quadratische' Abbildung ist, die gegeben ist durch

$$\tau(x) = \sum_{j=1}^{n+1} \lambda_j x_j^2$$

für alle $x = (x_1, \ldots, x_{n+1}) \in \mathbb{R}^{n+1}$.

Beweis Sei $A = (a_{ij})) \in \mathrm{M}((n+1) \times (n+1), \mathbb{R})$ die durch

$$a_{ij} = \begin{cases} d_j & \text{falls } i = j \text{ und } 1 \leq j \leq n \; , \\ c_{ij}/2 & \text{falls } 1 \leq i < j \leq n \; , \\ c_{ji}/2 & \text{falls } 1 \leq j < i \leq n \; , \\ b_j/2 & \text{falls } 1 \leq j \leq n \text{ und } i = n+1 \; , \\ b_i/2 & \text{falls } 1 \leq i \leq n \text{ und } j = n+1 \; , \\ a & \text{falls } i = j = n+1 \end{cases}$$

definierte Matrix und sei $s : \mathbb{R}^{n+1} \times \mathbb{R}^{n+1} \to \mathbb{R}$ die durch

$$s(x, y) = \sum_{j=1}^{n+1} \sum_{k=1}^{n+1} a_{jk} x_j y_k$$

für alle $x = (x_1, \ldots, x_{n+1})$, $y = (y_1, \ldots, y_{n+1}) \in \mathbb{R}^{n+1}$ definierte Bilinearform. Da die Matrix A symmetrisch ist, ist s symmetrisch, und man sieht leicht, dass

$$q((x_1, \ldots, x_n)) = s((x_1, \ldots, x_n, 1), (x_1, \ldots, x_n, 1))$$

für alle $(x_1, \ldots, x_n) \in \mathbb{R}^n$. Nach Satz 22.6 gibt es nun Elemente $\lambda_1, \ldots, \lambda_{n+1} \in \mathbb{R}$ und einen orthogonalen Endomorphismus $\psi : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$, so dass

$$s(x, y) = \sigma(\psi(x), \psi(y))$$

für alle $x, y \in \mathbb{R}^{n+1}$, wobei $\sigma : \mathbb{R}^{n+1} \times \mathbb{R}^{n+1} \to \mathbb{R}$ die durch

$$\sigma(x, y) = \sum_{j=1}^{n+1} \lambda_j x_j y_j$$

für alle $x = (x_1, \ldots, x_{n+1})$, $y = (y_1, \ldots, y_{n+1}) \in \mathbb{R}^{n+1}$ gegebene symmetrische Bilinearform auf $\mathbb{R}^{n+1}$ ist. Für alle $x \in \mathbb{R}^n$ gilt also

$$q(x) = s(\gamma(x), \gamma(x)) = \sigma(\psi(\gamma(x)), \psi(\gamma(x))) = \tau(\psi(\gamma(x))) \ ,$$

d.h.. $q = \tau \circ \psi \circ \gamma$. □

Im Folgenden sei V wieder ein endlichdimensionaler reeller Vektorraum und sei $s : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform.

Satz 22.8 *Es gibt eine Basis α von V, so dass die Matrix von s bezüglich α eine Diagonalmatrix ist.*

Beweis Sei $(u_1, \ldots, u_n)$ eine beliebige Basis von V und sei $A \in \mathrm{M}(n \times n, \mathbb{R})$ die Matrix von s bezüglich $(u_1, \ldots, u_n)$. Da A symmetrisch ist, gibt es nach Satz 21.11 eine orthogonale Matrix $P \in \mathrm{M}(n \times n, \mathbb{R})$, so dass $P^t AP = P^{-1}AP$ eine Diagonalmatrix D ist. Nach Lemma 6.6 gibt es eine Basis $(v_1, \ldots, v_n)$ von V, so dass P die Matrix für den Wechsel von $(u_1, \ldots, u_n)$ nach $(v_1, \ldots, v_n)$ ist. Daraus ergibt sich nach Lemma 22.2, dass $P^t AP = D$ die Matrix von s bezüglich $(v_1, \ldots, v_n)$ ist. □

Satz 22.9 *Es gibt eine Basis α von V, so dass die Matrix von s bezüglich α folgende Gestalt hat:*

$$\begin{pmatrix} E_p & & \\ & -E_q & \\ & & 0 \end{pmatrix} \ ,$$

wobei $p \geq 0$, $q \geq 0$ und $p + q \leq n$. Ist $\alpha = (v_1, \ldots, v_v)$, so gilt also

$$s\Big(\sum_{j=1}^{n} \lambda_j v_j, \sum_{k=1}^{n} \mu_k v_k\Big) = (\lambda_1\mu_1 + \cdots + \lambda_p\mu_p) - (\lambda_{p+1}\mu_{p+1} + \cdots + \lambda_{p+q}\mu_{p+q})$$

für alle $\lambda_1, \ldots, \lambda_n, \mu_1, \ldots, \mu_n \in \mathbb{R}$.

Beweis Nach Satz 22.8 gibt es eine Basis $\alpha = (v_1, \ldots, v_n)$ von V, so dass die Matrix von s bezüglich α eine Diagonalmatrix ist. Seien $\lambda_1, \ldots, \lambda_n$ die Einträge auf der Diagonalen von D. Für jede Permutation $\pi \in S_n$ ist $\alpha_\pi = (v_{\pi(1)}, \ldots, v_{\pi(n)})$ auch eine Basis von V und die Matrix von s bezüglich α_π ist die Diagonalmatrix D_π mit Einträgen $\lambda_{\pi(1)}, \ldots, \lambda_{\pi(n)}$ auf der Diagonalen. Es gibt also $p \geq 0$, $q \geq 0$ mit $p + q \leq n$ und eine Permutation π, so dass $\lambda_{\pi(j)} > 0$ für $j = 1, \ldots, p$, $\lambda_{\pi(j)} < 0$ für $j = p+1, \ldots, p+q$ und $\lambda_{\pi(j)} = 0$ für $j = p+q+1, \ldots, n$. Setze nun $(v_{\pi(1)}, \ldots, v_{\pi(n)}) = (u_1, \ldots, u_n)$ und $(\lambda_{\pi(1)}, \ldots, \lambda_{\pi(n)}) = (\mu_1, \ldots, \mu_n)$, sei

$$\varepsilon_j = \begin{cases} \left(\sqrt{\mu_j}\right)^{-1} & \text{falls } 1 \leq j \leq p\,, \\ \left(\sqrt{-\mu_j}\right)^{-1} & \text{falls } p+1 \leq j \leq p+q\,, \\ 1 & \text{falls } p+q+1 \leq j \leq n\,. \end{cases}$$

und für $j = 1, \ldots, n$ sei $w_j = \varepsilon_j u_j$. Dann ist $(w_1, \ldots, w_n)$ eine Basis von V und

$$s(w_j, w_k) = s(\varepsilon_j u_j, \varepsilon_k u_k) = \varepsilon_j \varepsilon_k s(u_j, u_k) = \begin{cases} \varepsilon_j^2 \mu_j & \text{falls } j = k\,, \\ 0 & \text{falls } j \neq k\,, \end{cases}$$

und ferner ist

$$\varepsilon_j^2 \mu_j = \begin{cases} 1 & \text{falls } 1 \leq j \leq p\,, \\ -1 & \text{falls } p+1 \leq j \leq p+q\,, \\ 0 & \text{falls } p+q+1 \leq j \leq n\,. \end{cases}$$

Daraus ergibt sich, dass

$$\begin{pmatrix} E_p & & \\ & -E_q & \\ & & 0 \end{pmatrix}$$

die Matrix von s bezüglich $(w_1, \ldots, w_n)$ ist. □

Satz 22.10 (Sylvestersches Trägheitsgesetz) *Seien α und β zwei Basen von V und sei A bzw. B die Matrix von s bezüglich α bzw. β. Nehme an, dass*

$$A = \begin{pmatrix} E_p & & \\ & -E_q & \\ & & 0 \end{pmatrix} \quad \textit{und} \quad B = \begin{pmatrix} E_{p'} & & \\ & -E_{q'} & \\ & & 0 \end{pmatrix}\,.$$

Dann ist $p = p'$ und $q = q'$.

Beweis Sei $\alpha = (u_1, \ldots, u_n)$, $\beta = (v_1, \ldots, v_n)$ und sei P die Matrix für den Wechsel von α nach β; nach Lemma 22.2 ist $B = P^{\mathrm{t}} A P$ und daraus folgt, dass $\operatorname{rang} A = \operatorname{rang} B$, da P und P^{t} invertierbar sind. Damit ist

$$p + q = \operatorname{rang} A = \operatorname{rang} B = p' + q'\,.$$

Setze $U = \mathrm{L}(u_1, \ldots, u_p)$ und $W = \mathrm{L}(v_{p'+1}, \ldots, v_n)$. Dann ist

$$s\Big(\sum_{j=1}^{p} \lambda_j u_j, \sum_{j=1}^{p} \lambda_j u_j\Big) = \sum_{j=1}^{p} \sum_{k=1}^{p} \lambda_j \lambda_k s(u_j, u_k) = \sum_{j=1}^{p} \lambda_j^2 \ ,$$

und damit ist $s(u, u) > 0$ für alle $u \in U \setminus \{0\}$. Genauso gilt

$$s\Big(\sum_{j=p'+1}^{n} \mu_j v_j, \sum_{j=p'+1}^{n} \mu_j v_j\Big) = \sum_{j=p'+1}^{n} \sum_{k=p'+1}^{n} \mu_j \mu_k s(v_j, v_k) = -\sum_{j=p'+1}^{n} \mu_j^2 \ ,$$

und damit ist $s(w, w) \leq 0$ für alle $w \in W$. Folglich ist $U \cap W = \{0\}$, und daraus ergibt sich nach Satz 4.8, dass

$$p + (n - p') = \dim U + \dim W = \dim(U + W) \leq n \ ,$$

d.h., $p \leq p'$. Aber das gleiche Argument zeigt, dass $p' \leq p$ und damit ist $p = p'$. Nun gilt auch, dass $q = (p + q) - p = (p' + q') - p' = q'$. □

Sei $s : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform und seien $p \geq 0$, $q \geq 0$ mit $p + q \leq n = \dim V$ die durch Satz 22.9 und Satz 22.10 eindeutig bestimmten Zahlen. Man nennt $p + q$ den *Rang* von s, p den *Index* von s und $p - q$ die *Signatur* von s.

23 Quotientenräume und das Tensor-Produkt

Im folgenden sei K ein Körper; ein Vektorraum bedeutet stets ein Vektorraum über K.

Sei V ein Vektorraum und sei U ein Untervektorraum von V. Zunächst wird ein neuer Vektorraum, der Quotientenraum von V durch U, konstruiert.

Für jedes $v \in V$ sei $U + v = \{w \in V : w = u + v \text{ für ein } u \in U\}$. Insbesondere ist $v \in U + v$ und $U + 0 = U$.

Lemma 23.1 *(1) Für alle v_1, $v_2 \in V$ gilt entweder $(U + v_1) \cap (U + v_2) = \varnothing$ oder $U + v_1 = U + v_2$. Ferner gilt $U + v_1 = U + v_2$ genau dann, wenn $v_1 - v_2 \in U$.*

(2) Seien v_1, v_1', v_2, $v_2' \in V$ mit $U + v_1 = U + v_1'$ und $U + v_2 = U + v_2'$; dann gilt $U + (v_1 + v_2) = U + (v_1' + v_2')$.

(3) Seien v_1, $v_1' \in V$ mit $U + v_1 = U + v_1'$; dann gilt $U + \lambda v_1 = U + \lambda v_1'$ für alle $\lambda \in K$.

Beweis (1) Nehme an, dass $(U + v_1) \cap (U + v_2) \neq \varnothing$; dann gibt es u_1, $u_2 \in U$ mit $u_1 + v_1 = u_2 + v_2$ und folglich ist $v_1 - v_2 = u_2 - u_1 \in U$. Sei nun $v \in U + v_1$; dann gibt es $u \in U$ mit $v = u + v_1$ und damit ist $v = u + (v_1 - v_2) + v_2 \in U + v_2$, da $u + (v_1 - v_2) \in U$. Folglich ist $U + v_1 \subset U + v_2$ und genauso gilt $U + v_2 \subset U + v_1$, d.h., $U + v_1 = U + v_2$. Dies zeigt, dass entweder $(U + v_1) \cap (U + v_2) = \varnothing$ oder $U + v_1 = U + v_2$ und auch dass $v_1 - v_2 \in U$, falls $U + v_1 = U + v_2$. Ferner zeigt der Beweis, dass $U + v_1 = U + v_2$, falls $v_1 - v_2 \in U$.

(2) Gilt $U + v_1 = U + v_1'$ und $U + v_2 = U + v_2'$, so ist nach (1) $v_1 - v_1' \in U$ und $v_2 - v_2' \in U$. Damit ist $(v_1 + v_2) - (v_1' + v_2') = (v_1 - v_1') + (v_2 - v_2') \in U$ und daraus folgt nach (1), dass $U + (v_1 + v_2) = U + (v_1' + v_2')$.

(3) Gilt $U + v_1 = U + v_1'$, so ist nach (1) $v_1 - v_1' \in U$. Daraus folgt nach (1), dass $U + \lambda v_1 = U + \lambda v_1'$, da $\lambda v_1 - \lambda v_1' = \lambda(v_1 - v_1') \in U$. □

Eine Teilmenge $N \subset V$ wird *U-Nebenklasse* genant, wenn $W = U + v$ für ein $v \in V$. Insbesondere ist U selbst eine Nebenklasse. Da $v \in U + v$, liegt jedes $v \in V$ in mindestens einer U-Nebenklasse, und daraus ergibt sich nach Lemma 23.1 (1), dass jeder Vektor v in genau einer U-Nebenklasse liegt.

Die Menge der U-Nebenklassen wird mit V/U bezeichnet. Nach Lemma 23.1 (2) kann eine Addition $+ : V/U \times V/U \to V/U$ durch

$$(U + v_1) + (U + v_2) = U + (v_1 + v_2)$$

definiert werden. Es ist klar, dass diese Addition assoziativ und kommutativ ist. Ferner ist $U + N = N$ für jedes $N \in V/U$ und zu jedem $N \in V/U$ gibt es ein

$-N \in V/U$ mit $-N+N = U$, da $(U+(-v))+(U+v) = U$. Nach Lemma 23.1 (3) kann ebefalls eine Multiplikation mit Skalaren $\cdot : K \times V/U \to V/U$ durch

$$\lambda(U+v) = U + \lambda v$$

definiert werden. Mit diesen Verknüpfungen ist V/U ein Vektorraum über K, der der *Quotientenraum* von V durch U heißt. Die Nebenklasse U ist das Nullelement in V/U.

Sei nun $\pi : V \to V/U$ die Abbildung mit $\pi(v) = U + v$ für jedes $v \in V$. Da

$$\pi(v_1+v_2) = U + (v_1+v_2) = (U+v_1)+(U+v_2) = \pi(v_1)+\pi(v_2)$$

für alle $v_1, v_2 \in V$ und $\pi(\lambda v) = U + \lambda v = \lambda(U+v) = \lambda\pi(v)$ für alle $v \in V$, $\lambda \in K$, ist π eine lineare Abbildung. Ferner ist nach Lemma 23.1 (1)

$$\operatorname{Kern}\pi = \{v \in V : \pi(v) = U\} = \{v \in V : U+v = U+0\} = U$$

und π ist surjektiv. Die Abbildung π wird *kanonische Abbildung* genannt.

Satz 23.1 *Sei W ein Vektorraum und $f : V \to W$ eine lineare Abbildung mit $U \subset \operatorname{Kern} f$; dann gibt es eine eindeutige lineare Abbildung $g : V/U \to W$ mit $g \circ \pi = f$.*

Beweis Gilt $U + v_1 = U + v_2$, so ist nach Lemma 23.1 (1) $v_1 - v_2 \in U \subset \operatorname{Kern} f$ und damit ist $f(v_1) = f(v_1 - v_2 + v_2) = f(v_1 - v_2) + f(v_2) = 0 + f(v_2) = f(v_2)$. Folglich gibt es eine eindeutige Abbildung $g : V/U \to W$ mit $g(U+v) = f(v)$ für alle $v \in V$ und da $g(U+x) = g(\pi(v))$, ist dann g die eindeutige Abbildung mit $g \circ \pi = f$. Aber g ist linear, da

$$\begin{aligned} g((U+v_1)+(U+v_2)) &= g(U+(v_1+v_2)) \\ &= f(v_1+v_2) = f(v_1)+f(v_2) = g(U+v_1)+g(U+v_2) \end{aligned}$$

für alle $v_1, v_2 \in V$ und

$$g(\lambda(U+v)) = g(U+\lambda v) = f(\lambda v) = \lambda f(v) = \lambda g(U+v)$$

für alle $v \in V$, $\lambda \in K$. Also ist $g : V/U \to W$ die eindeutige lineare Abbildung mit $g \circ \pi = f$. □

Die Eigenschaft in Satz 23.1 führt zur folgenden Definition: Ein Paar (V_U, ϱ) bestehend aus einem Vektorraum V_U und einer linearen Abbildung $\varrho : V \to V_U$ mit $\operatorname{Kern}\varrho = U$ heißt ein *V/U-Quotientenraum*, wenn es zu jedem Vektorraum W und zu jeder linearen Abbildung $f : V \to W$ mit $U \subset \operatorname{Kern} f$ eine eindeutige lineare Abbildung $g : V_U \to W$ gibt, so dass $f = g \circ \varrho$.

Satz 23.2 *Es gibt einen V/U-Quotientenraum.*

Beweis Nach Satz 23.1 ist $(V/U, \pi)$ ein V/U-Quotientenraum. □

Lemma 23.2 *Seien (V_U, ϱ) und (V'_U, ϱ') zwei V/U-Quotientenräume; dann gibt es einen eindeutigen Isomorphismus $h : V_U \to V'_U$, so dass $h \circ \varrho = \varrho'$.*

Beweis (1) Das Paar (V_U, ϱ) ist ein V/U-Quotientenraum und $\varrho' : V \to V'_U$ ist eine lineare Abbildung mit $U \subset \operatorname{Kern} \varrho'$. Folglich gibt es eine eindeutige lineare Abbildung $h : V_U \to V'_U$ mit $h \circ \varrho = \varrho'$.

(2) Das Paar (V'_U, ϱ') ist ein V/U-Quotientenraum und $\varrho : V \to V_U$ ist eine lineare Abbildung mit $U \subset \operatorname{Kern} \varrho$. Folglich gibt es eine eindeutige lineare Abbildung $h' : V'_U \to V_U$ mit $h' \circ \varrho' = \varrho$.

(3) Das Paar (V_U, ϱ) ist ein V/U-Quotientenraum und $\varrho : V \to V_U$ ist eine lineare Abbildung mit $U \subset \operatorname{Kern} \varrho$. Folglich gibt es eine eindeutige lineare Abbildung $g : V_U \to V_U$ mit $g \circ \varrho = \varrho$. Aber $\mathrm{id}_{V_U} : V_U \to V_U$ ist eine lineare Abbildung mit $\mathrm{id}_{V_u} \circ \varrho = \varrho$ und damit ist $g = \mathrm{id}_{V_U}$. Dies bedeutet: id_{V_U} ist der eindeutige Endomorphismus von V_U mit $\mathrm{id}_{V_U} \circ \varrho = \varrho$.

(4) Genauso ist $\mathrm{id}_{V'_U}$ der eindeutige Endomorphismus von V'_U mit $\mathrm{id}_{V'_U} \circ \varrho' = \varrho'$.

(5) Andererseits ist $h' \circ h$ ein Endomorphismus von V_U mit

$$(h' \circ h) \circ \varrho = h' \circ (h \circ \varrho) = h' \circ \varrho' = \varrho$$

und daher ist $h' \circ h = \mathrm{id}_{V_U}$.

(6) Genauso ist $h \circ h' = \mathrm{id}_{V'_U}$.

(7) Da $h' \circ h = \mathrm{id}_{V_U}$ und $h \circ h' = \mathrm{id}_{V'_U}$, ist $h : V_U \to V'_U$ ein Isomorphismus, und da h die einzige lineare Abbildung ist mit $h \circ \varrho = \varrho'$, ist h insbesondere der einzige Isomorphismus, so dass $h \circ \varrho = \varrho'$. □

Den Beweis für Lemma 23.2 soll man merken; ein im Wesentlichen identischer Beweis kommt stets bei jeder Definition vor, die die gleiche Struktur hat wie in der Definition eines U/V-Quotientenraums.

In der Konstruktion eines Tensor-Produkts braucht man einen Quotientenraum und einfachhalber wird der explizite Quotientenraum $(V/U, \pi)$ verwendet. Es wird aber von V/U und π lediglich benutzt, dass $(V/U, \pi)$ ein V/U-Quotientenraum ist.

Sei X eine Menge. Ein Vektorraum V mit $X \subset V$ heißt *X-frei*, wenn es zu jedem Vektorraum W und zu jeder Abbildung $g : X \to W$ eine eindeutige lineare Abbildung $f : V \to W$ gibt, so dass $g = f_{|X}$.

Satz 23.3 *Sei V ein Vektorraum und sei X eine nichtleere endliche Teilmenge von V. Dann ist V ein X-freier Vektorraum genau, wenn die Vektoren in X eine Basis von V bilden (und insbesondere ist V endlichdimensional mit $\dim V = |X|$).*

Beweis Übung. □

Lemma 23.3 *Seien V und V' zwei X-freie Vektorräume; dann gibt es einen eindeutigen Isomorphismus $h : V \to V'$, so dass $h(x) = x$ für jedes $x \in X$.*

Beweis Im Wesentlichen identisch mit dem Beweis für Lemma 23.2. □

Satz 23.4 *Es gibt einen X-freien Vektorraum.*

Beweis Sei $\mathrm{Abb}(X, K)$ der Vektorraum aller Abbildungen von X nach K und sei

$$V = \{h \in \mathrm{Abb}(X, K) : \mathrm{s}(h) \text{ is endlich}\} ,$$

wobei $\mathrm{s}(h) = \{x \in X : h(x) \neq 0\}$. Für jedes $x \in X$ sei ferner $\varepsilon_x : X \to K$ durch

$$\varepsilon_x(y) = \begin{cases} 1 & \text{falls } y = x , \\ 0 & \text{falls } y \neq x , \end{cases}$$

definiert; also ist $\varepsilon_x \in V$. Für jedes $h \in V$ gilt dann $h = \sum_{x \in \mathrm{s}(h)} h(x)\varepsilon_x$. Setze $X' = \{\varepsilon_x : x \in X\}$. Sei nun W ein Vektorraum und $g : X' \to W$ eine Abbildung; definiere $f : V \to W$ durch

$$f(h) = \sum_{x \in \mathrm{s}(h)} h(x) g(\varepsilon_x) .$$

Dann sieht man leicht, dass f die eindeutige lineare Abbildung mit $g = f_{|X'}$ ist und folglich ist V ein X'-freier Vektorraum. Aber die Abbildung $x \mapsto \varepsilon_x$ liefert eine Bijektion zwischen X und X' und daher kann man X mit X' identifizieren. Auf diese Weise ist V auch ein X-freier Vektorraum. □

Es wird nun das Tensor-Produkt von Vektorräumen eingeführt. Dies wird nur für zwei Faktoren gemacht; bei dem allgemeinen Fall von n Faktoren tauchen aber keine zusätzliche Probleme auf.

Seien V_1, V_2 und V Vektorräume. Eine Abbildung $s : V_1 \times V_2 \to V$ heißt *bilinear*, wenn für jedes $v_1 \in V_1$ die Abbildung $s(v_1, \cdot) : V_2 \to V$ linear und für jedes $v_2 \in V_2$ die Abbildung $s(\cdot, v_2) : V_1 \to V$ linear ist.

Seien V_1 und V_2 Vektorräume. Ein Paar (V, σ) bestehend aus einem Vektorraum V und einer bilinearen Abbildung $\sigma : V_1 \times V_2 \to V$ heißt ein *Tensor-Produkt* von V_1 und V_2, wenn es zu jedem Vektorraum W und zu jeder bilinearen Abbildung $s : V_1 \times V_2 \to W$ eine eindeutige lineare Abbildung $f : V \to W$ gibt, so dass $f \circ \sigma = s$.

Lemma 23.4 *Seien (V, σ) und (V', σ') zwei Tensor-Produkte von V_1 und V_2; dann gibt es einen eindeutigen Isomorphismus $h : V \to V'$, so dass $h \circ \sigma = \sigma'$.*

Beweis Im Wesentlichen identisch mit dem Beweis für Lemma 23.2. □

Satz 23.5 *Es gibt ein Tensor-Produkt von V_1 und V_2.*

Beweis Setze $X = V_1 \times V_2$ und sei V ein X-freier Vektorraum. Sei nun $[X]$ die Teilmenge von V bestehend aus allen Elementen, die eine der folgenden Formen haben:

$(v_1 + v_1', v_2) - (v_1, v_2) - (v_1', v_2)$ mit $v_1,\, v_1' \in V_1$ und $v_2 \in V_2$,

$(v_1, v_2 + v_2') - (v_1, v_2) - (v_1, v_2')$ mit $v_1 \in V_1$ und $v_2,\, v_2' \in V_2$,

$(\lambda v_1, v_2) - \lambda(v_1, v_2)$ mit $v_1, \in V_1$, $v_2 \in V_2$ und $\lambda \in K$,

$(v_1, \lambda v_2) - \lambda(v_1, v_2)$ mit $v_1, \in V_1$, $v_2 \in V_2$ und $\lambda \in K$.

Sei U der kleinste Untervektorraum von V, der $[X]$ enthält (also ist U der Durchschnitt von allen Untervektorräumen, die $[X]$ enthalten).

Betrachte den Quotientenraum V/U und die kanonische Abbildung $\pi : V \to V/U$. Sei $\sigma = \pi_{|X} : X = V_1 \times V_2 \to V/U$ die Einschränkung von $\pi : V \to V/U$ auf X.

Die Abbildung $\sigma : V_1 \times V_2 \to V/U$ ist bilinear: Seien $v_1,\, v_1' \in V_1$ und $v_2 \in V_2$; dann ist $(v_1 + v_1', v_2) - (v_1, v_2) - (v_1', v_2) \in [X] \subset U = \operatorname{Kern} \pi$ und damit ist

$$\begin{aligned}\sigma(v_1 + v_1', v_2) - \sigma(v_1, v_2) - \sigma(v_1', v_2) \\ &= \pi((v_1 + v_1', v_2)) - \pi((v_1, v_2)) - \pi((v_1', v_2)) \\ &= \pi((v_1 + v_1', v_2) - (v_1, v_2) - (v_1', v_2)) = 0\ ;\end{aligned}$$

d.h., $\sigma(v_1 + v_1', v_2) = \sigma(v_1, v_2) + \sigma(v_1', v_2)$. Genauso gilt

$\sigma(v_1, v_2 + v_2') = \sigma(v_1, v_2) + \sigma(v_1, v_2')$ für alle $v_1 \in V_1,\, v_2,\, v_2' \in V_2$,

$\sigma(\lambda v_1, v_2) = \lambda\sigma(v_1, v_2)$ für alle $v_1, \in V_1$, $v_2 \in V_2$ und $\lambda \in K$, und

$\sigma(v_1, \lambda v_2) = \lambda\sigma(v_1, v_2)$ für alle $v_1, \in V_1$, $v_2 \in V_2$ und $\lambda \in K$.

Damit ist σ bilinear. Sei nun $s : V_1 \times V_2 \to W$ eine beliebige bilineare Abbildung. Da $X = V_1 \times V_2$ und V ein X-freier Vektorraum ist, gibt es eine eindeutige lineare Abbildung $g : V \to W$, so dass $g_{|X} = s$, und es gilt $U \subset \operatorname{Kern} g$: Seien $v_1,\, v_1' \in V_1$ und $v_2 \in V_2$; dann ist

$$\begin{aligned}g((v_1 + v_1', v_2) - (v_1, v_2) - (v_1', v_2)) \\ &= g((v_1 + v_1', v_2)) - g((v_1, v_2)) - g((v_1', v_2)) \\ &= s(v_1 + v_1', v_2) - s(v_1, v_2) - s(v_1', v_2) = 0\ ;\end{aligned}$$

d.h., $(v_1 + v_1', v_2) - (v_1, v_2) - (v_1', v_2) \in \operatorname{Kern} g$. Genauso gilt

$(v_1, v_2 + v_2') - (v_1, v_2) - (v_1, v_2') \in \operatorname{Kern} g$ für alle $v_1 \in V_1$, v_2, $v_2' \in V_2$,

$(\lambda v_1, v_2) - \lambda(v_1, v_2) \in \operatorname{Kern} g$ für alle $v_1, \in V_1$, $v_2 \in V_2$ und $\lambda \in K$, und

$(v_1, \lambda v_2) - \lambda(v_1, v_2) \in \operatorname{Kern} g$ für alle $v_1, \in V_1$, $v_2 \in V_2$ und $\lambda \in K$.

Folglich ist $[X] \subset \operatorname{Kern} g$ und daher ist $U \subset \operatorname{Kern} g$.

Da $U \subset \operatorname{Kern} g$ und $(V/U, \pi)$ ein V/U-Quotientenraum ist, gibt es eine eindeutige lineare Abbildung $f : V/U \to W$ mit $f \circ \pi = g$. Daraus ergibt sich insbesondere, dass $f \circ \pi_{|X} = g_{|X}$, d.h., $f \circ \sigma = s$.

Sei schließlich $f' : V/U \to W$ eine beliebige lineare Abbildung mit $f' \circ \sigma = s$. Dann ist $f' \circ \pi : V \to W$ eine lineare Abbildung mit $(f' \circ \pi)_{|X} = f' \circ \pi_{|X} = f' \circ \sigma = s$. Aber g ist die eindeutige lineare Abbildung mit $g_{|X} = s$ und also ist $f' \circ \pi = g$. Aber f ist die eindeutige lineare Abbildung mit $f \circ \pi = g$ und damit ist $f' = f$. Folglich ist $f : V/U \to W$ die eindeutige lineare Abbildung, so dass $f \circ \sigma = s$, und dies zeigt, dass $(V/U, \sigma)$ ein Tensor-Produkt von V_1 und V_2 ist. □

Ein Tensor-Produkt von V_1 und V_2 wird meistens mit $(V_1 \otimes V_2, \otimes)$ bezeichnet, also ist $V_1 \otimes V_2$ ein Vektorraum und $\otimes : V_1 \times V_2 \to V_1 \otimes V_2$ eine bilineare Abbildung (und man schreibt $v_1 \otimes v_2$ statt $\otimes(v_1, v_2)$). Für jede bilineare Abbildung $s : V_1 \times V_2 \to W$ gibt es also eine eindeutige lineare Abbildung $f : V_1 \otimes V_2 \to W$, so dass $f \circ \otimes = s$. Man redet hier meistens von *dem* Tensor-Produkt $V_1 \otimes V_2$, obwohl $V_1 \otimes V_2$ nur im Sinne von Lemma 23.4 eindeutig ist.

Satz 23.6 *Seien V_1 und V_2 endlichdimensionale Vektorräume mit $V_1 \neq \{0\}$ und $V_2 \neq \{0\}$, sei $(v_1, \ldots, v_m)$ eine Basis von V_1 und $(w_1, \ldots, w_n)$ eine Basis von V_2. Dann bilden die Vektoren*

$$\{v_j \otimes w_k : 1 \leq j \leq m,\, 1 \leq k \leq n\}$$

eine Basis von $V_1 \otimes V_2$. Insbesondere ist $V_1 \otimes V_2$ endlichdimensional mit

$$\dim V_1 \otimes V_2 = \dim V_1 \times \dim V_2 \,.$$

Beweis Setze $B = \{v_j \otimes w_k : 1 \leq j \leq m,\, 1 \leq k \leq n\}$. Betrachte einen beliebigen Vektorraum W und eine beliebige Abbildung $g : B \to W$. Da $(v_1, \ldots, v_m)$ und $(w_1, \ldots, w_n)$ Basen sind, kann eine Abildung $s : V_1 \times V_2 \to W$ durch

$$s\Big(\sum_{j=1}^{m} \lambda_j v_j, \sum_{k=1}^{n} \mu_k w_k\Big) = \sum_{j=1}^{m} \sum_{k=1}^{n} \lambda_j \mu_k g(v_j \otimes w_k)$$

definiert werden und es ist klar, dass s biliear ist. Es gibt also eine eindeutige lineare Abbildung $f : V_1 \otimes V_2 \to W$, so dass $f \circ \otimes = s$. Insbesondere gilt dann

$$f(v_j \otimes w_k) = s(v_j, w_k) = g(v_j \otimes w_k)$$

für alle j, k, d.h., $f_{|B} = g$. Sei umgekehrt $f' : V_1 \otimes V_2 \to W$ eine lineare Abbildung mit $f'_{|B} = g$; dann gilt $f'(v_j \otimes w_k) = g(v_j \otimes w_k) = s(v_j, w_k)$ für alle j, k. Daraus folgt, dass $f' \circ \otimes = s$ und damit ist $f' = f$. Dies zeigt: Zu jedem Vektorraum W und jeder Abbildung $g : B \to W$ gibt es eine eindeutige lineare Abbildung $f : V_1 \otimes V_2 \to W$, so dass $f_{|B} = g$, d.h., $V_1 \otimes V_2$ ist ein B-freier Vektorraum. Nach Satz 23.3 bilden dann die Vektoren in B eine Basis von $V_1 \otimes V_2$. □

Satz 23.7 *Betrachte K als eindimensionalen Vektorraum, sei V ein Vektorraum und sei $\sigma : K \times V \to V$ Multiplikation mit Skalaren (also ist $\sigma(\lambda, v) = \lambda v$ für alle $\lambda \in K$, $v \in V$). Dann ist (V, σ) ein Tensor-Produkt von K und V, und daher kann man $V = K \otimes V$ schreiben (mit $\lambda \otimes v = \lambda v$).*

Beweis Zunächst ist es klar, dass die Abbildung $\sigma : K \times V \to V$ bilinear ist. Sei also W ein Vektorraum und $s : K \times V \to W$ eine bilineare Abbildung. Definiere $f : V \to W$ durch $f(v) = s(1, v)$; also ist f linear und

$$(f \circ \sigma)(\lambda, v) = f(\sigma(\lambda, v)) = f(\lambda v) = \lambda f(v) = \lambda s(1, v) = s(\lambda, v)$$

für alle $\lambda \in K$, $v \in V$, d.h., $f \circ \sigma = s$. Ist umgekehrt $f' : V \to W$ eine beliebige Abbildung mit $f' \circ \sigma = s$, so ist $f'(v) = f'(1v) = f'(\sigma(1, v)) = s(1, v) = f(v)$ für alle $v \in V$, d.h., $f' = f$. Insbesondere ist f die eindeutige lineare Abbildung, so dass $f \circ \sigma = s$. □

Es wird nun eine wichtige Konstruktion präsentiert, die das Tensor-Produkt verwendet. Diese wird zunächst im Fall $K = \mathbb{R}$ durchgeführt.

Der reelle Vektorraum $\mathbb{R}^n$ ist Teilmenge der komplexen Vektorraum $\mathbb{C}^n$. Ferner kann man $\mathbb{C}^n$ als reellen Vektorraum betrachten: Die Addition ist die Addition in $\mathbb{C}^n$ und die Multiplikation mit Skalaren $\cdot : \mathbb{R} \times \mathbb{C}^n \to \mathbb{C}^n$ die Einschränkung der Multiplikation mit Skalaren $\cdot : \mathbb{C} \times \mathbb{C}^n \to \mathbb{C}^n$. Dieser reelle Vektorraum wird mit $\mathbb{C}^n_{\mathbb{R}}$ bezeichnet. Insbesondere (mit $n = 1$) bezeichnet $\mathbb{C}_{\mathbb{R}}$ der Körper $\mathbb{C}$ betrachtet als reeller Vektorraum. Sei $\sigma : \mathbb{C}_{\mathbb{R}} \times \mathbb{R}^n \to \mathbb{C}^n_{\mathbb{R}}$ die Abbildung mit $\sigma(z, v) = zv$ (also ist σ die Einschränkung der Multiplikation mit Skalaren $\cdot : \mathbb{C} \times \mathbb{C}^n \to \mathbb{C}^n$).

Lemma 23.5 *$(\mathbb{C}^n_{\mathbb{R}}, \sigma)$ ist ein Tensor-Produkt von $\mathbb{C}_{\mathbb{R}}$ und $\mathbb{R}^n$.*

Beweis Zunächst ist es klar, dass die Abbildung $\sigma : \mathbb{C}_{\mathbb{R}} \times \mathbb{R}^n \to \mathbb{C}^n_{\mathbb{R}}$ bilinear ist. Sei also W ein reeller Vektorraum und $s : \mathbb{C}_{\mathbb{R}} \times \mathbb{R}^n \to W$ eine bilineare Abbildung. Definiere $f : \mathbb{C}^n_{\mathbb{R}} \to W$ durch $f(v) = s(1, v_1) + s(i, v_2)$, wobei v_1, v_2 die eindeutigen Elemente von $\mathbb{R}^n$ mit $v = v_1 + iv_2$ sind. Dann sieht man leicht, dass f die eindeutige lineare Abbildung ist mit $f \circ \sigma = s$. □

Für jedes $z \in \mathbb{C}_{\mathbb{R}}$ definiere $\sigma_z : \mathbb{C}_{\mathbb{R}} \times \mathbb{R}^n \to \mathbb{C}^n_{\mathbb{R}}$ durch $\sigma_z(z', v) = \sigma(zz', v)$. Dann ist σ_z bilinear und folglich gibt es einen eindeutigen Endomorphismus α_z von $\mathbb{C}^n_{\mathbb{R}}$,

so dass $\alpha_z \circ \sigma = \sigma_z$. Aber hier ist α_z explizit gegeben durch $\alpha_z(v) = zv$ für alle $v \in \mathbb{C}^n_{\mathbb{R}}$ und dies bedeutet, dass die Abbildung $(z, v) \mapsto \alpha_z(v)$ im Wesentlichen nichts anderes ist als Multiplikation mit Skalaren $\cdot : \mathbb{C} \times \mathbb{C}^n \to \mathbb{C}^n$.

Man beachte schließlich, dass die Abbildung $\sigma(1, \cdot) : \mathbb{R}^n \to \mathbb{C}^n_{\mathbb{R}}$ injektiv ist und daher kann man $\mathbb{R}^n$ mit dem Untervektorraum Bild $\sigma(1, \cdot)$ von $\mathbb{C}^n_{\mathbb{R}}$ identifizieren.

Die obige Konstruktion für $\mathbb{R}^n$ wird nun für einen beliebigen reellen Vektorraum V durchgeführt. Sei $\mathbb{C}_{\mathbb{R}} \otimes V$ das Tensor-Produkt von $\mathbb{C}_{\mathbb{R}}$ und V.

Lemma 23.6 *(1) Für jedes $z \in \mathbb{C}_{\mathbb{R}}$ gibt es einen eindeutigen Endomorphismus α_z von $\mathbb{C}_{\mathbb{R}} \otimes V$, so dass $\alpha_z(z' \otimes v) = zz' \otimes v$ für alle $z' \in \mathbb{C}_{\mathbb{R}}$, $v \in V$.*

(2) Für alle z_1, $z_2 \in \mathbb{C}_{\mathbb{R}}$ gilt $\alpha_{z_1+z_2} = \alpha_{z_1} + \alpha_{z_2}$ und $\alpha_{z_1 z_2} = \alpha_{z_1} \circ \alpha_{z_2}$.

(3) Für jedes $x \in \mathbb{R} \subset \mathbb{C}_{\mathbb{R}}$ ist $\alpha_x = x \, \mathrm{id}_{\mathbb{C}_{\mathbb{R}} \otimes V}$.

Beweis (1) Definiere $\sigma_z : \mathbb{C}_{\mathbb{R}} \times V \to \mathbb{C}_{\mathbb{R}} \otimes V$ durch $\sigma_z(z', v) = zz' \otimes v$. Dann ist die Abbildung σ_z bilinear und folglich gibt es einen eindeutigen Endomorphismus α_z von $\mathbb{C}_{\mathbb{R}} \otimes V$, so dass $\alpha_z \circ \otimes = \sigma_z$, d.h., so dass $\alpha_z(z' \otimes v) = zz' \otimes v$ für alle $z' \in \mathbb{C}_{\mathbb{R}}$, $v \in V$.

(2) Für alle $z' \in \mathbb{C}_{\mathbb{R}}$, $v \in V$ gilt

$$(\alpha_{z_1} + \alpha_{z_2})(z' \otimes v) = \alpha_{z_1}(z' \otimes v) + \alpha_{z_2}(z' \otimes v) = z_1 z' \otimes v + z_2 z' \otimes v = (z_1 + z_2) z' \otimes v$$

und nach der Eindeutigkeit von $\alpha_{z_1+z_2}$ ist also $\alpha_{z_1} + \alpha_{z_2} = \alpha_{z_1+z_2}$. Genauso gilt

$$(\alpha_{z_1} \circ \alpha_{z_2})(z' \otimes v) = \alpha_{z_1}(\alpha_{z_2}(z' \otimes v)) = \alpha_{z_1}(z_2 z' \otimes v) = z_1 z_2 z \otimes v$$

für alle $z' \in \mathbb{C}_{\mathbb{R}}$, $v \in V$ und folglich ist $\alpha_{z_1} \circ \alpha_{z_2} = \alpha_{z_1 z_2}$.

(3) Sei $x \in \mathbb{R}$; da $\otimes$ bilinear ist, gilt für alle $z' \in \mathbb{C}_{\mathbb{R}}$, $v \in V$, dass

$$(x \, \mathrm{id}_{\mathbb{C}_{\mathbb{R}} \otimes V})(z' \otimes v) = x(z' \otimes v) = xz' \otimes v$$

und nach der Eindeutigkeit von α_x ist daher $x \, \mathrm{id}_{\mathbb{C}_{\mathbb{R}} \otimes V} = \alpha_x$. □

Für $z \in \mathbb{C}_{\mathbb{R}} = \mathbb{C}$, $w \in \mathbb{C}_{\mathbb{R}} \otimes V$ setze $z \star w = \alpha_z(w)$. Nach Lemma 23.6 gilt dann

(a) $(z_1 z_2) \star v = z_1 \star (z_2 w)$ für alle z_1, $z_2 \in \mathbb{C}$, $w \in \mathbb{C}_{\mathbb{R}} \otimes V$.

(b) $1 \star w = w$ für alle $w \in \mathbb{C}_{\mathbb{R}} \otimes V$.

(c) $z \star (w_1 + w_2) = z \star w_1 + z \star w_2$ für alle $z \in \mathbb{C}$, w_1, $w_2 \in \mathbb{C}_{\mathbb{R}} \otimes V$.

(d) $(z_1 + z_2) \star w = z_1 \star w + z_2 \star w$ für alle z_1, $z_2 \in \mathbb{C}$, $w \in \mathbb{C}_{\mathbb{R}} \otimes V$.

(Natürlich gilt (c) einfach, weil α_z linear ist.) Dies bedeutet: Mit der Addition aus $\mathbb{C}_\mathbb{R} \otimes V$ (als reellem Vektorraum) und mit $\star : \mathbb{C} \times (\mathbb{C}_\mathbb{R} \otimes V) \to \mathbb{C}_\mathbb{R} \otimes V$ als Multiplikation mit Skalaren wird $\mathbb{C}_\mathbb{R} \otimes V$ ein komplexer Vektorraum. Ferner ist nach Lemma 23.6 (3) die Verknüpfung $\star$ verträglich mit der Multiplikation mit Skalaren $\cdot : \mathbb{R} \times (\mathbb{C}_\mathbb{R} \otimes V) \to \mathbb{C}_\mathbb{R} \otimes V$: Es gilt $x \star w = xw$ für alle $x \in \mathbb{R}$, $w \in \mathbb{C}_\mathbb{R} \otimes V$.

Lemma 23.7 *Die lineare Abbildung $v \mapsto 1 \otimes v$ von V nach $\mathbb{C}_\mathbb{R} \otimes V$ ist injektiv.*

Beweis Definiere $s : \mathbb{C}_\mathbb{R} \times V \to \mathbb{R}$ durch $s((x + iy), v) = xv$ für alle $x, y \in \mathbb{R}$ und $v \in V$. Dann ist s eine bilinear und damit gibt es eine eindeutige lineare Abbildung $f : \mathbb{C}_\mathbb{R} \otimes V \to \mathbb{R}$ mit $f \circ \otimes = s$. Insbesondere ist $f(1 \otimes v) = s(1, v) = v$ für alle $v \in V$ und also ist $1 \otimes v \neq 0$, falls $v \neq 0$. □

Nach Lemma 23.7 kann man V mit dem (reellen) Untervektorraum

$$1 \otimes V = \{1 \otimes v : v \in V\}$$

von $\mathbb{C}_\mathbb{R} \otimes V$ identifizieren. Auf diese Weise ist der reelle Vektorraum V Teilmenge des komplexen Vektorraumes $\mathbb{C}_\mathbb{R} \otimes V$; ferner sind die Verknüpfungen auf V die Einschränkungen der entsprechenden Verknüpfungen auf $\mathbb{C}_\mathbb{R} \otimes V$.

Sei nun W ein weiterer reeller Vektorraum und sei $h : V \to W$ eine lineare Abbildung; sei $s_h : \mathbb{C}_\mathbb{R} \times V \to \mathbb{C}_\mathbb{R} \otimes W$ die Abbildung mit $s_h(z, v) = z \otimes h(v)$ für alle $z \in \mathbb{C}_\mathbb{R}$, $v \in V$. Dann ist s_h bilinear und folglich gibt es eine eindeutige lineare Abbildung $h_\mathbb{C} : \mathbb{C}_\mathbb{R} \otimes V \to \mathbb{C}_\mathbb{R} \otimes W$, so dass $h_\mathbb{C} \circ \otimes = s_h$. Zunächst ist $h_\mathbb{C}$ nur eine lineare Abbildung mit $\mathbb{C}_\mathbb{R} \otimes V$ und $\mathbb{C}_\mathbb{R} \otimes W$ als reellen Vektorräumen. Aber $h_\mathbb{C}$ ist auch eine lineare Abbildung zwischen den komplexen Vektorräumen $\mathbb{C}_\mathbb{R} \otimes V$ und $\mathbb{C}_\mathbb{R} \otimes W$: Für alle $z \in \mathbb{C}_\mathbb{R}$ gilt

$$(h_\mathbb{C} \circ \alpha_z) \circ \otimes = (\alpha_z \circ h_\mathbb{C}) \circ \otimes = s_h^z \ ,$$

wobei $s_h^z(z', v) = zz' \otimes h(v)$ für alle $z' \in \mathbb{C}_\mathbb{R}$, $v \in V$, und also ist $h_\mathbb{C} \circ \alpha_z = \alpha_z \circ h_\mathbb{C}$, d.h. $h_\mathbb{C}(z \star w) = z \star h_\mathbb{C}(w)$ für alle $w \in \mathbb{C}_\mathbb{R} \otimes V$. Schließlich ist

$$h_\mathbb{C}(1 \otimes v) = s_h(1, v) = 1 \otimes h(v)$$

für alle $v \in V$, und damit ist die Einschränkung von $h_\mathbb{C}$ auf $1 \otimes V$ die lineare Abbildung $1 \otimes h : 1 \otimes V \to 1 \otimes W$, wobei $(1 \otimes h)(v) = 1 \otimes h(v)$ für jedes $v \in V$.

Die obige Konstruktion kann für eine beliebige Körpererweiterung durchgeführt werden: Ein Körper F heißt *Körpererweiterung* eines Körpers K, wenn $K \subset F$ und die Addition (bzw. die Multiplikation) in K die Einschränkung der Addition (bzw. der Multiplikation) in F ist. Ist F eine Körpererweiterung von K, so ist die Null (bzw. die Eins) in K auch die Null (bzw. die Eins) in F.

Im Folgenden sei F eine Körpererweiterung eines Körpers K. Dann kann F als Vektorraum über K betrachtet werden; dieser K-Vektorraum wird mit F_K bezeichnet. Sei V ein K-Vektorraum; genauso wie im Spezialfall $K = \mathbb{R}/F = \mathbb{C}$ gibt es eine Verknüpfung $\star : F \times (F_K \otimes V) \to F_K \otimes V$, die den K-Vektorraum $F_K \otimes V$ in einen F-Vektorraum umwandelt.

Das einzige Problem hier ist das Ergebnis, das dem Lemma 27.7 entspricht. Dieses ist zwar richtig, aber im Allgemeinen ist der Beweis dafür nicht so trivial:

Lemma 23.8 *Die lineare Abbildung $v \mapsto 1 \otimes v$ von V nach $F_K \otimes V$ ist injektiv.*

Beweis Dies folgt unmittelbar aus Satz 23.8. □

Satz 23.8 *Sei $V_1 \otimes V_2$ das Tensor-Produkt von Vektorräumen V_1 und V_2. Dann ist $v_1 \otimes v_2 \neq 0$ für alle $v_1 \in V_1 \setminus \{0\}$, $v_2 \in V_2 \setminus \{0\}$.*

Beweis Sei $v_1 \in V_1 \setminus \{0\}$; nach Lemma 23.9 gibt es eine Linearform $\varphi : V_1 \to K$ mit $\varphi(v_1) \neq 0$. Definiere $s : V_1 \times V_2 \to V_2$ durch $s(u_1, u_2) = \varphi(u_1)u_2$; dann ist s bilinear und folglich gibt es eine eindeutige lineare Abbildung $f : V_1 \otimes V_2 \to V_2$, so dass $f \circ \otimes = s$. Also ist $f(v_1 \otimes v_2) = \varphi(v_1)v_2 \neq 0$ für alle $v_2 \neq 0$ und damit ist $v_1 \otimes v_2 \neq 0$ für alle $v_2 \in V_2 \setminus \{0\}$. □

Lemma 23.9 *Zu jedem Vektorraum V und jedem Vektor $v \in V \setminus \{0\}$ gibt es eine Linearform $\varphi : V \to K$ mit $\varphi(v) \neq 0$.*

Beweis Nehme zunächst an, dass V endlichdimensional ist; dann gibt es eine Basis $(v_1, \ldots, v_n)$ von V mit $v_1 = v$ und es gibt eine Linearform $\varphi : V \to K$ mit

$$\varphi(\lambda_1 v_1 + \cdots + \lambda_n v_n) = \lambda_1$$

für alle $\lambda_1, \ldots, \lambda_n \in K$. Insbesondere ist $\varphi(v) = 1 \neq 0$.

Im Allgemeinen braucht man den folgenden Fakt: Es gibt stets eine Teilmenge X von V mit $v \in X$, so dass V ein X-freier Vektorraum ist. Per Definition gibt es insbesondere dann eine Linearform $\varphi : V \to K$ mit $\varphi(v) = 1$. □

Literatur

[1] Beutelspacher, Albrecht (2001): Lineare Algebra. Vieweg

[2] Brieskorn, Egbert (1985): Lineare Algebra und analytische Geometrie. Vierweg

[3] Fischer, Gerd (1997): Lineare Algebra. Vierweg

[4] Halmos, Paul (1995): Linear algebra problem book. Math. Assoc. Am.

[5] Jänich, Klaus (2002): Lineare Algebra. Springer

[6] Klingenberg, Wilhelm (1992): Lineare Algebra und Geometrie. Springer

[7] Kowalsky, Hans-Joachim (1998): Lineare Algebra. De-Gruyter-Lehrbuch

[8] Lang, Serge (1987); Introduction to linear algebra. Springer

[9] Lorenz, Falko (1996): Lineare Algebra. Spektrum Verlag

Index