LINEARE ALGEBRA II

OLIVER C. SCHNÜRER

Zusammenfassung. Bei diesem Manuskript handelt es sich um Notizen zu einer Vorlesung Lineare Algebra II (B2) an der Universität Konstanz im Sommersemester 2011 und im Sommersemester 2013.

Vielen Dank an Mario Kummer, Wolfgang Maurer und Olaf Schnürer für Korrekturen und Verbesserungsvorschläge.

Inhaltsverzeichnis

Wir benutzen [1, 2, 4, 6, 10, 11] und für den Teil über Ringe und Moduln insbesondere [8].

1. Der Dualraum

1.1. **Kovariante Schreibweise.**

Bemerkung 1.1.1. Seien F ein Körper und V, W endlichdimensionale F-Vektorräume. Sei $f\colon V \to W$ linear. Seien $v_1, \dots, v_n$ eine Basis von V und $w_1, \dots, w_m$ eine Basis von W. Dann gibt es $\left(a_i^j\right)_{\substack{1\le j\le m\\ 1\le i\le n}}$ mit

$$f(v_i) = \sum_{j=1}^{m} a_i^j w_j.$$

Sei $\xi \in V$ mit $\xi = \sum_{i=1}^{n} \xi^i v_i$. Dann folgt

$$f(\xi) = f\left(\sum_{i=1}^{n} \xi^i v_i\right) = \sum_{i=1}^{n} \xi^i f(v_i) = \sum_{i=1}^{n} \xi^i \sum_{j=1}^{m} a_i^j w_j = \sum_{j=1}^{m} \left(\sum_{i=1}^{n} \xi^i a_i^j\right) w_j.$$

Date: 6. Oktober 2017.

Somit sind die Koordinaten $\left(\zeta^j\right)_{1\le j\le m}$ von $f(\xi)$ bezüglich der Basis $w_1,\dots,w_m$, so dass also $f(\xi)=\sum\limits_{j=1}^{m}\zeta^j w_j$ gilt, durch $\zeta^j=\sum\limits_{i=1}^{n}\xi^i a_i^j=\sum\limits_{i=1}^{n}a_i^j\xi^i$ gegeben. In Matrixschreibweise erhalten wir

$$\begin{pmatrix}\zeta^1\\ \zeta^2\\ \vdots\\ \zeta^m\end{pmatrix}=\begin{pmatrix}a_1^1 & a_2^1 & \dots & a_n^1\\ a_1^2 & a_2^2 & \dots & a_n^2\\ \vdots & \vdots & \ddots & \vdots\\ a_1^m & a_2^m & \dots & a_n^m\end{pmatrix}\begin{pmatrix}\xi^1\\ \xi^2\\ \vdots\\ \xi^n\end{pmatrix}.$$

In dieser Schreibweise beinhalten Summationen stets einen oberen und einen unteren Index.

1.2. **Dualraum.**

Definition 1.2.1. Abbildungen $f\in\operatorname{Hom}(V,F)$ heißen (lineare) Funktionale auf V.

Wir schreiben $V^*:=\operatorname{Hom}(V,F)$. V^* heißt der zu V duale Raum.

Bemerkung 1.2.2.

(i) Nach Wahl einer Basis $(a_1,\dots,a_n)$ von V und bezüglich der Basis 1 von F hat ein lineares Funktional die Gestalt

$$f\left(\sum_{i=1}^{n}x^i a_i\right)=\sum_{i=1}^{n}x^i f(a_i).$$

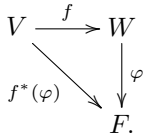
Wir bezeichnen die Abbildung $F^n\ni(x^1,\dots,x^n)\mapsto\sum\limits_{i=1}^{n}x^i f(a_i)$ auch als Linearform.

(ii) Für $0\neq f\in V^*=\operatorname{Hom}(V,F)$ mit $\dim V=n$ gilt $\dim\ker f=n-1$, da $\dim\operatorname{im}f=1$ und $\dim\operatorname{im}f+\dim\ker f=n$ gelten.

(iii) Habe V die Basis $(a_1,\dots,a_n)$. Dann bilden die Vektoren $f^i\in V^*$ mit $f^i(a_k)=\delta_k^i$ eine Basis von V^*. Wir schreiben auch $(a^*)^i=f^i$. Für $b=\sum\limits_{i=1}^{n}x^i a_i$ gilt

$$(a^*)^j(b)=\sum_{i=1}^{n}x^i(a^*)^j(a_i)=\sum_{i=1}^{n}x^i\delta_i^j=x^j.$$

Definition 1.2.3 (Duale Abbildung)**.** Sei $f\colon V\to W$ linear. Dann definieren wir die duale Abbildung $f^*\colon W^*\to V^*$ zu f durch $f^*(\varphi):=\psi$ mit $\psi(\xi)=\varphi(f(\xi))$ für $\varphi\in W^*$ und $\xi\in V$.



(Es folgt aus der Definition, dass $f^*(\varphi)\in V^*$ gilt und dass f^* linear ist.)

Lemma 1.2.4. *Sei $f\colon V\to W$ durch die $(m\times n)$-Matrix $A=\left(a_j^i\right)_{\substack{1\le i\le m\\ 1\le j\le n}}$ bezüglich Basen $(a_i)_{1\le i\le n}$ und $(b_j)_{1\le j\le m}$ von V bzw. W dargestellt. Dann ist $f^*\colon W^*\to V^*$ durch die $(m\times n)$-Matrix $\left(b_j^i\right)_{\substack{1\le i\le m\\ 1\le j\le n}}$ mit $b_j^i=a_j^i$ bezüglich der Basen $\left((b^*)^j\right)_{1\le j\le m}$ und $\left((a^*)^i\right)_{1\le i\le n}$ in dem Sinne dargestellt, dass $f^*\left((b^*)^j\right)=\sum\limits_{i=1}^{n}b_i^j(a^*)^i$ für alle $1\le j\le m$ gilt.*

Für die Koordinaten erhalten wir die Abbildung

$$(\xi_1, \dots, \xi_m) \mapsto \left(\sum_{j=1}^{m} \xi_j b_i^j \right)_{1 \le i \le n}.$$

Beweis. Nach Definition gilt

$$\begin{aligned} \left(f^* \left((b^*)^j \right) \right)(a_k) = \left((b^*)^j \circ f \right)(a_k) = (b^*)^j \left(\sum_{i=1}^{m} a_k^i b_i \right) \\ = \sum_{i=1}^{m} a_k^i (b^*)^j b_i = \sum_{i=1}^{m} a_k^i \delta_i^j = a_k^j. \end{aligned}$$

Aus $f^* \left((b^*)^j \right) = \sum_{l=1}^{n} b_l^j (a^*)^l$ folgt andererseits

$$\left(f^* \left((b^*)^j \right) \right)(a_k) = \sum_{l=1}^{n} b_l^j (a^*)^l (a_k) = \sum_{l=1}^{n} b_l^j \delta_k^l = b_k^j.$$

Die Behauptung für die Koordinaten ergibt sich direkt aus der Linearität. □

Bemerkung 1.2.5. Achtung, in der Darstellung als Matrix haben wir soeben über andere Indices als für Abbildungen $f\colon V \to W$ summiert. Die Koordinaten verändern sich folglich nach der Regel

$$(\xi_1, \dots, \xi_m) \mapsto (\xi_1, \dots, \xi_m) \begin{pmatrix} b_1^1 & \dots & b_n^1 \\ \vdots & & \vdots \\ b_1^m & \dots & b_n^m \end{pmatrix}.$$

Beachte dazu, dass wir die Basisvektoren von V^* anders als die von V oben indiziert haben. Die Koordinaten haben wir folglich unten indiziert. Somit handelt es sich um Zeilenvektoren.

In der Literatur weit verbreitet ist aber die (aus Kovarianzgründen nicht so saubere) Variante, auch diese Koordinaten als Spaltenvektoren zu schreiben, Indices generell unten anzubringen und den oberen Index einer Matrix an die erste Stelle zu senken, also a_{ij} statt a_j^i zu verwenden. Dann gilt für die Komponenten unter f^* die Abbildungsregel

$$\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_m \end{pmatrix} \mapsto \begin{pmatrix} b_{11} & \dots & b_{m1} \\ \vdots & & \vdots \\ b_{1n} & \dots & b_{mn} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_m \end{pmatrix}.$$

Wie man sich leicht überzeugt sind die Einträge im so erhaltenen Spaltenvektor dieselben wie beim Zeilenvektor. Hier ist nun die darstellende Matrix die Matrix A^T, die transponierte Matrix, die für $A = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}}$ durch $A^T := (b_{ji})_{\substack{1 \le j \le n \\ 1 \le i \le m}}$ mit $b_{ji} := a_{ij}$ definiert ist. Graphisch erhält man

$$A^T = \begin{pmatrix} a_{11} & a_{21} & a_{31} & \dots & a_{m1} \\ a_{12} & a_{22} & a_{32} & \dots & a_{m2} \\ a_{13} & a_{23} & a_{33} & \dots & a_{m3} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & a_{3n} & \dots & a_{mn} \end{pmatrix} \quad \text{aus} \quad A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix}.$$

2. Vektorräume mit Skalarprodukt

Sobald wir Skalarprodukte verwenden, wollen wir stets annehmen, dass wir $\mathbb{R}$- oder $\mathbb{C}$-Vektorräume betrachten.

2.1. **Euklidische Vektorräume.** Ein euklidischer Vektorraum ist ein Vektorraum mit einem reellen Skalarprodukt.

Definition 2.1.1. Sei V ein $\mathbb{R}$-Vektorraum. Ein (reelles) Skalarprodukt auf V ist eine Funktion $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{R}$, die folgendes erfüllt:

(i) $\langle a, b \rangle = \langle b, a \rangle$ für alle $a, b \in V$ (Symmetrie)
(ii) $\langle \lambda a + \mu b, c \rangle = \lambda \langle a, c \rangle + \mu \langle b, c \rangle$ für alle $a, b, c \in V$ und $\lambda, \mu \in \mathbb{R}$ (Linearität)
(iii) $\langle a, a \rangle \geq 0$ und $\langle a, a \rangle = 0 \iff a = 0$ für alle $a \in V$ (positive Definitheit)

Bemerkung 2.1.2.

(i) Aus des Symmetrie und der Linearität im ersten Argument folgt auch die Linearität im zweiten Argument

$$\langle c, \lambda a + \mu b \rangle = \langle \lambda a + \mu b, c \rangle = \lambda \langle a, c \rangle + \mu \langle b, c \rangle = \lambda \langle c, a \rangle + \mu \langle c, b \rangle.$$

Eine Funktion in zwei Argumenten, die in beiden Argumenten linear ist, heißt bilinear.

(ii) Per Induktion zeigt man für beliebige Linearkombinationen

$$\left\langle \sum_{i=1}^{m} \lambda^i a_i, \sum_{j=1}^{n} \mu^j b_j \right\rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} \lambda^i \mu^j \langle a_i, b_j \rangle.$$

Somit ist ein Skalarprodukt durch seine Werte auf einer Basis eindeutig bestimmt.

(iii) Sei $V = \mathbb{R}^n$. Für Vektoren $\xi = (\xi^1, \ldots, \xi^n)$ und $\eta = (\eta^1, \ldots, \eta^n)$ definieren wir

$$\langle \xi, \eta \rangle := \sum_{i=1}^{n} \xi^i \eta^i.$$

Dies ist ein Skalarprodukt auf $\mathbb{R}^n$ (Übung), das Standard-Skalarprodukt.

(iv) Sei $V = \mathbb{R}^2$. Definiere für $\xi = (\xi^1, \xi^2)$ und $\eta = (\eta^1, \eta^2)$

$$\langle \xi, \eta \rangle := \xi^1 \eta^1 + 5\xi^1 \eta^2 + 5\xi^2 \eta^2 + 26\xi^2 \eta^2.$$

Dies ist ein weiteres Skalarprodukt auf $\mathbb{R}^2$. Es stimmt nicht mit dem Standardskalarprodukt auf $\mathbb{R}^2$ überein.

(v) Sei V der Vektorraum der auf $[a, b] \subset \mathbb{R}$ stetigen reellwertigen Funktionen, $V = C^0([a, b])$. Gelte bei diesen Beispielen stets $a < b$. Definiere für $f, g \in V$

$$\langle f, g \rangle = \int_a^b f(x) g(x) \, dx.$$

Dies ist ein Skalarprodukt auf V (vgl. Analysis-Vorlesung).

Definition 2.1.3. Sei $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{R}$ ein Skalarprodukt und sei $a_1, \ldots, a_n$ eine Basis von V. Dann definieren wir

$$c_{ij} := \langle a_i, a_j \rangle \quad \text{für } 1 \leq i, j \leq n.$$

$C = (c_{ij})_{1 \leq i,j \leq n}$ heißt Matrix zum Skalarprodukt $\langle \cdot, \cdot \rangle$.

Beachte, dass wir hier beide Indices unten schreiben. In Matrizenform stellen wir C durch

$$\begin{pmatrix} c_{11} & c_{12} & \ldots & a_{1n} \\ c_{21} & c_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \ldots & a_{nn} \end{pmatrix}$$

dar. Es ist nicht ideal, Matrizen der Form $\left(c^i_j\right)$ und (c_{ij}) graphisch gleich darzustellen, jedoch üblich.

Theorem 2.1.4. *Die Matrix $C = (c_{ij})_{1\leq i,j\leq n}$ eines reellen Skalarproduktes ist symmetrisch, d. h. es gilt $c_{ij} = c_{ji}$ für alle $1 \leq i, j \leq n$ oder $C = C^T$ mit $C^T := (b_{ij})_{1\leq i,j\leq n}$ mit $b_{ij} = c_{ji}$.*

Beweis. Sei $a_1, \ldots, a_n$ eine Basis. Dann gilt

$$c_{ij} = \langle a_i, a_j\rangle = \langle a_j, a_i\rangle = c_{ji}. \qquad \square$$

Beispiele 2.1.5.

(i) Bezüglich der Standardbasis ist die Matrix des Standardskalarproduktes auf $\mathbb{R}^n$ gleich $\mathbb{1} = I = (\delta_{ij})_{1\leq i,j\leq n}$.

(ii) Das Skalarprodukt mit

$$\langle \xi, \eta\rangle := \xi^1\eta^1 + 5\xi^1\eta^2 + 5\xi^2\eta^1 + 26\xi^2\eta^2$$

ist bezüglich der Standardbasis des $\mathbb{R}^2$ durch die Matrix

$$\begin{pmatrix} 1 & 5 \\ 5 & 26 \end{pmatrix}$$

und bezüglich der Basis aus den Vektoren $(1, 0)$ und $(-5, 1)$ durch die Matrix $\mathbb{1}$ dargestellt.

(iii) Sei V der Vektorraum der Polynome vom Grad ≤ 3 mit Basis $(1, x, x^2, x^3)$. Dann ist das Skalarprodukt

$$\langle p, q\rangle = \int_0^1 p(x)q(x)\,dx$$

durch die Matrix

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \end{pmatrix}$$

dargestellt.

Lemma 2.1.6. *Sei $\langle\cdot,\cdot\rangle$ ein Skalarprodukt, dem bezüglich der Basis $a_1, \ldots, a_n$ die Matrix $(c_{ij})_{1\leq i,j\leq n}$ zugeordnet ist. Seien $\xi = \sum_{i=1}^n \xi^i a_i$ und $\eta = \sum_{j=1}^n \eta^i a_i$ beliebig. Dann gilt*

$$\langle \xi, \eta\rangle = \sum_{i,j=1}^n \xi^i c_{ij}\eta^j.$$

Beweis. Benutze die Linearität in beiden Argumenten wie in Bemerkung 2.1.2 (ii). $\square$

Bemerkung 2.1.7. Wir wollen schließlich noch das Verhalten einer ein Skalarprodukt darstellenden Matrix unter Basistransformationen untersuchen.

Sei V ein reeller Vektorraum mit Basen S und T, $S = (a_1, \ldots, a_n)$ und $T = (b_1, \ldots, b_n)$. Gelte

$$b_k = \sum_{i=1}^n d_k^i a_i.$$

Das Skalarprodukt $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{R}$ werde bezüglich S durch die Matrix $C = (c_{ij})_{1 \le i,j \le n}$ beschrieben. Dann gilt

$$\begin{aligned} m_{kl} := \langle b_k, b_l \rangle &= \left\langle \sum_{i=1}^{n} d_k^i a_i, \sum_{j=1}^{n} d_l^j a_j \right\rangle \\ &= \sum_{i,j=1}^{n} d_k^i d_l^j \langle a_i, a_j \rangle = \sum_{i,j=1}^{n} d_k^i d_l^j c_{ij}. \end{aligned}$$

Setze $M := (m_{ij})_{1 \le i,j \le n}$ und $D := (d_k^i)_{1 \le i,k \le n}$. Dann gilt

$$M = D^T C D.$$

2.2. **Unitäre Vektorräume.** Mit $\bar{z}$ bezeichnen wir die komplex konjugierte Zahl zu z. Ist also $z = x + iy$ mit $x, y \in \mathbb{R}$, so ist $\bar{z} = x - iy$.

Ein Vektorraum mit einem unitären Skalarprodukt heißt unitärer Vektorraum.

Definition 2.2.1. Sei V ein $\mathbb{C}$-Vektorraum. Ein unitäres Skalarprodukt auf V ist eine Funktion $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{C}$, die die folgenden Eigenschaften erfüllt:

(i) $\langle a, b \rangle = \overline{\langle b, a \rangle}$ für alle $a, b \in V$ (hermitesch)

(ii) $\langle \lambda a + \mu b, c \rangle = \lambda \langle a, c \rangle + \mu \langle b, c \rangle$ für alle $a, b, c \in V$ und alle $\lambda, \mu \in \mathbb{C}$ (Linearität im ersten Argument)

(iii) $\langle a, a \rangle \ge 0$ und $\langle a, a \rangle = 0 \iff a = 0$ für alle $a \in V$ (positiv definit)

Bemerkung 2.2.2. Bei der positiven Definitheit dürfen wir $\langle a, a \rangle \ge 0$ schreiben, da $\langle a, a \rangle$ aufgrund der Hermitizität für alle $a \in V$ reell ist.

Folgende Eigenschaften und Beispiele sind analog zum reellen Fall

(i) Seien $a, b, c \in V$ und $\lambda, \mu \in \mathbb{C}$. Dann gilt

$$\langle c, \lambda a + \mu b \rangle = \overline{\langle \lambda a + \mu b, c \rangle} = \bar{\lambda} \overline{\langle a, c \rangle} + \bar{\mu} \overline{\langle b, c \rangle} = \bar{\lambda} \langle c, a \rangle + \bar{\mu} \langle c, b \rangle.$$

Linearität im ersten Argument und dieses Verhalten im zweiten Argument bezeichnet man als Sesquilinearität.

Es gibt auch die umgekehrte Konvention, d. h. man definiert ein unitäres Skalarprodukt so, dass es im zweiten Argument statt im ersten Argument linear ist und dass die übrigen Eigenschaften unverändert gelten. Im ersten Argument werden dann Skalare komplex konjugiert nach außen gezogen.

(ii) Per Induktion folgt hieraus für beliebige Linearkombinationen

$$\left\langle \sum_{i=1}^{m} \lambda^i a_i, \sum_{j=1}^{n} \mu^j b_j \right\rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} \lambda^i \overline{\mu^j} \langle a_i, b_j \rangle.$$

Daher ist ein unitäres Skalarprodukt durch seine Werte auf einer Basis bereits eindeutig bestimmt.

(iii) Ist $V = \mathbb{C}^n$ und seien $x = (x^1, \ldots, x^n)$ und $y = (y^1, \ldots, y^n)$ Vektoren in V, so ist durch

$$\langle x, y \rangle := \sum_{i=1}^{n} x^i \overline{y^i}$$

ein unitäres Skalarprodukt auf $\mathbb{C}^n$ definiert.

(iv) Sei V der komplexe Vektorraum der auf $[a, b]$ komplexwertigen stetigen Funktionen einer reellen Variablen. Dann definiert

$$\langle f, g \rangle := \int_a^b f(t) \overline{g(t)} \, dt$$

ein unitäres Skalarprodukt auf V. Vergleiche wieder eine Analysis-Vorlesung für die positive Definitheit.

Definition 2.2.3. Sei V ein komplexer Vektorraum und $\langle\cdot,\cdot\rangle\colon V\times V\to\mathbb{C}$ ein unitäres Skalarprodukt. Sei $(a_1,\dots,a_n)$ eine Basis von V. Dann heißt die Matrix $C=(c_{ij})_{1\le i,j\le n}$ mit

$$c_{ij}=\langle a_i,a_j\rangle,\quad 1\le i,j,\le n$$

die Matrix des Skalarproduktes $\langle\cdot,\cdot\rangle$ bezüglich der Basis $a_1,\dots,a_n$.

Theorem 2.2.4. *Die Matrix C eines unitären Skalarproduktes ist hermitesch, d. h. es gilt $C^T=\bar{C}$.*

Beweis. Es gilt

$$c_{ij}=\langle a_i,a_j\rangle=\overline{\langle a_j,a_i\rangle}=\overline{c_{ji}}.$$

Es folgt $C^T=\overline{C}$ wie behauptet. □

Wie im reellen Fall zeigt man:

Lemma 2.2.5. *Sei $\langle\cdot,\cdot\rangle$ ein unitäres Skalarprodukt, dem bezüglich einer Basis $a_1,\dots,a_n$ die Matrix $(c_{ij})_{1\le i,j\le n}$ zugeordnet ist. Seien $\xi=\sum_{i=1}^{n}\xi^i a_i$ und $\eta=\sum_{j=1}^{n}\eta^i a_i$ beliebig. Dann gilt*

$$\langle\xi,\eta\rangle=\sum_{i,j=1}^{n}\xi^i c_{ij}\bar{\eta^j}.$$

Bemerkung 2.2.6. Wir wollen wiederum das Verhalten einer ein Skalarprodukt darstellenden Matrix unter Basistransformationen untersuchen.

Sei V ein komplexer Vektorraum mit Basen S und T, $S=(a_1,\dots,a_n)$ und $T=(b_1,\dots,b_n)$. Gelte

$$b_k=\sum_{i=1}^{n}d_k^i a_i.$$

Ein Skalarprodukt $\langle\cdot,\cdot\rangle\colon V\times V\to\mathbb{C}$ werde bezüglich S durch die Matrix $C=(c_{ij})_{1\le i,j\le n}$ beschrieben. Dann gilt

$$\begin{aligned}m_{kl}:=\langle b_k,b_l\rangle&=\left\langle\sum_{i=1}^{n}d_k^i a_i,\sum_{j=1}^{n}d_l^j a_j\right\rangle\\&=\sum_{i,j=1}^{n}d_k^i\overline{d_l^j}\langle a_i,a_j\rangle=\sum_{i,j=1}^{n}d_k^i\overline{d_l^j}c_{ij}.\end{aligned}$$

Setze $M:=(m_{ij})_{1\le i,j\le n}$ und $D:=(d_k^i)_{1\le i,k\le n}$. Dann gilt

$$M=D^TC\bar{D}.$$

Bemerkung 2.2.7. Reelles und komplexes Skalarprodukt verhalten sich sehr ähnlich. Daher werden wir häufiger nur den komplexen Sachverhalt untersuchen. Ein analoges reelles Resultat folgt dann analog.

2.3. **Norm.**

Theorem 2.3.1 (Cauchy-Schwarzsche Ungleichung)**.** *Sei V ein Vektorraum mit Skalarprodukt. Dann gilt*

$$|\langle a,b\rangle|^2\le\langle a,a\rangle\cdot\langle b,b\rangle$$

für alle $a,b\in V$. Gleichheit gilt genau dann, wenn a und b linear abhängig sind.

Beweis. Der Fall $b=0$ ist einfach (Übung).

Sei $\lambda\in\mathbb{C}$ beliebig. Dann gilt

$$0\le\langle a-\lambda b,a-\lambda b\rangle=\langle a,a\rangle-\lambda\overline{\langle a,b\rangle}-\bar{\lambda}\langle a,b\rangle+\lambda\bar{\lambda}\langle b,b\rangle.$$

Setzen wir speziell $\lambda = \frac{\langle a,b\rangle}{\langle b,b\rangle}$, so folgt nach Multiplikation mit $\langle b,b\rangle$

$$0 \le \langle a,a\rangle \cdot \langle b,b\rangle - 2|\langle a,b\rangle|^2 + |\langle a,b\rangle|^2.$$

Die behauptete Ungleichung folgt.

Gilt Gleichheit, so gilt insbesondere auch in der ersten Ungleichung Gleichheit, also $0 = a - \lambda b$ aufgrund der positiven Definitheit. □

Alternativbeweis über $\mathbb{R}$. Sei $\lambda \in \mathbb{R}$ beliebig. Dann gilt

$$\begin{aligned} 0 &\le \langle \lambda a + b, \lambda a + b\rangle \\ &= \lambda^2\langle a,a\rangle + 2\lambda\langle a,b\rangle + \langle b,b\rangle. \end{aligned}$$

Dies ist für alle $\lambda \in \mathbb{R}$ aber nur möglich, wenn die Diskriminante der zugehörigen quadratischen Gleichung nicht positiv ist, wenn also

$$\langle a,b\rangle^2 - \langle a,a\rangle\langle b,b\rangle \le 0$$

gilt. Dies ist gerade die Behauptung. □

Korollar 2.3.2.

(i) Seien $a_1,\dots,a_n$ und $b_1,\dots,b_n$ reelle Zahlen. Dann gilt

$$\sum_{i=1}^{n} a_i b_i \le \left(\sum_{i=1}^{n} a_i^2\right)^{1/2} \cdot \left(\sum_{j=1}^{n} b_j^2\right)^{1/2}.$$

(ii) Seien $f,g\colon [a,b] \to \mathbb{C}$ stetig. Dann gilt

$$\left|\int_a^b f(t)\bar{g}(t)\,dt\right|^2 \le \int_a^b |f(t)|^2\,dt \cdot \int_a^b |g(t)|^2\,dt.$$

Beweis. Dies folgt direkt aus der Cauchy-Schwarzschen Ungleichung für den Vektorraum

(i) $\mathbb{R}^n$ mit Standardskalarprodukt.
(ii) der stetigen Funktionen auf $[a,b]$ mit $\int f\bar{g}$ als Skalarprodukt. □

Definition 2.3.3. Sei V ein reeller oder komplexer Vektorraum. Eine Funktion $\|\cdot\|\colon V \to \mathbb{R}$ heißt Norm auf V, wenn sie die folgenden Bedingungen erfüllt

(i) $\|\lambda a\| = |\lambda| \cdot \|a\|$ für alle $\lambda \in F$ und $a \in V$,
(ii) $\|a+b\| \le \|a\| + \|b\|$ für alle $a,b \in V$, (Dreiecksungleichung)
(iii) $\|a\| = 0 \iff a = 0$.

Theorem 2.3.4. *Sei V ein Vektorraum mit Skalarprodukt $\langle\cdot,\cdot\rangle$. Dann ist die Funktion $\|\cdot\|\colon V \to \mathbb{R}$, durch $\|a\| := \sqrt{\langle a,a\rangle}$ definiert, eine Norm.*

Beweis.

(i) $\|\lambda a\| = \sqrt{\langle \lambda a, \lambda a\rangle} = \sqrt{\lambda\bar{\lambda}\langle a,a\rangle} = |\lambda|\sqrt{\langle a,a\rangle} = |\lambda| \cdot \|a\|$.

(ii)
$$\begin{aligned} \|a+b\|^2 &= \langle a+b, a+b\rangle = \langle a,a\rangle + \langle a,b\rangle + \langle b,a\rangle + \langle b,b\rangle \\ &\le \langle a,a\rangle + 2|\langle a,b\rangle| + \langle b,b\rangle \\ &\le \|a\|^2 + 2\cdot\|a\|\cdot\|b\| + \|b\|^2 \qquad \text{(Cauchy-Schwarz)} \\ &= (\|a\| + \|b\|)^2. \end{aligned}$$

(iii) Die positive Definitheit der Norm folgt aus der positiven Definitheit des Skalarproduktes. □

In einem Skalarproduktraum (=Vektorraum mit Skalarprodukt) werden wir unter einer Norm immer $\|a\| := \sqrt{\langle a,a\rangle}$ verstehen.

2.4. **Orthonormalbasen.**

Definition 2.4.1. Sei V ein Skalarproduktraum. Dann heißt $a \in V$ senkrecht zu $b \in V$, wenn $\langle a, b\rangle = 0$ gilt. Wir schreiben $a \perp b$. Eine Teilmenge $S \subset V$ heißt orthogonal zu einer Teilmenge $T \subset V$, wenn $\langle s, t\rangle = 0$ für alle $(s, t) \in S \times T$ gilt. Hier schreiben wir auch $S \perp T$.

Bemerkung 2.4.2.

(i) Ist a orthogonal zu b, so ist b orthogonal zu a.

(ii) Der Nullvektor ist zu allen anderen Vektoren orthogonal, da aus

$$\langle 0, a\rangle = \langle 0 + 0, a\rangle = \langle 0, a\rangle + \langle 0, a\rangle$$

$\langle 0, a\rangle = 0$ folgt. Dies ist auch der einzige Vektor mit dieser Eigenschaft, da aus $\langle a, a\rangle = 0$ bereits $a = 0$ folgt.

(iii) Zwei Teilmengen S und T sind genau dann orthogonal, wenn $\langle S\rangle$ und $\langle T\rangle$ orthogonal sind (Übung).

Definition 2.4.3.

(i) Eine Familie/Teilmenge S von V heißt orthogonal, wenn je zwei verschiedene Elemente aus S orthogonal zueinander sind.

(ii) Eine orthogonale Familie S heißt orthonormiert, wenn $\langle a, a\rangle = 1$ für alle $a \in S$ gilt.

(iii) Eine orthonormierte Familie, die zugleich Basis von V ist, heißt orthonormierte Basis oder Orthonormalbasis von V.

Bemerkung 2.4.4.

(i) Eine orthogonale Familie S von Vektoren mit $0 \notin S$ kann man zu einer orthonormalen Familie machen, indem man jeden Vektor $a \in S$ durch $\frac{a}{\|a\|}$ ersetzt. Die neue Familie ist dann durch Normieren aus der alten hervorgegangen (Übung).

(ii) Sei V der Vektorraum der auf $[-\pi, \pi]$ stetigen reellwertigen Funktionen mit Skalarprodukt

$$\langle f, g\rangle = \int_{-\pi}^{\pi} f(x)g(x)\,dx.$$

Dann ist

$$\{1, \sin x, \cos x, \sin 2x, \cos 2x, \sin 3x, \cos 3x, \ldots\}$$

eine orthogonale Familie (Übung). Dies spielt bei Fourierreihen eine Rolle.

Theorem 2.4.5. *Sei S eine orthogonale Familie mit $0 \notin S$. Dann ist S linear unabhängig.*

Beweis. Gelte

$$\sum_{i=1}^{n} \lambda^i a_i = 0$$

mit $\lambda^i \in F$ und $a_i \in S$ für $1 \le i \le n$ und ein $n \in \mathbb{N}^+$ sowie $a_i \neq a_j$ für $i \neq j$. Dann folgt für $1 \le j \le n$

$$0 = \langle 0, a_j\rangle = \left\langle \sum_{i=1}^{n} \lambda^i a_i, a_j \right\rangle = \sum_{i=1}^{n} \lambda^i \langle a_i, a_j\rangle = \lambda^j \langle a_j, a_j\rangle.$$

Wegen $a_j \neq 0$ folgt also $\lambda^j = 0$. Somit ist S linear unabhängig. □

Theorem 2.4.6. *Sei V ein endlichdimensionaler Skalarproduktraum. Sei $a_1, \dots, a_n$ eine Orthonormalbasis. Dann gilt für jedes $b \in V$*

$$b = \sum_{i=1}^{n} \langle b, a_i \rangle a_i.$$

Beweis. Wir wissen bereits, dass sich b als Linearkombination der Form

$$b = \sum_{i=1}^{n} \lambda^i a_i$$

darstellen lässt. Hieraus folgt

$$\langle b, a_j \rangle = \left\langle \sum_{i=1}^{n} \lambda^i a_i, a_j \right\rangle = \sum_{i=1}^{n} \lambda^i \langle a_i, a_j \rangle = \lambda^j.$$

Wir erhalten die Behauptung. □

Der folgende Algorithmus erlaubt es, aus einer Basis eine Orthogonalbasis zu gewinnen.

Theorem 2.4.7 (Gram-Schmidtsches Orthogonalisierungsverfahren)**.** *Jeder endlichdimensionale Skalarproduktraum besitzt eine Orthonormalbasis.*

Beweis. Sei $a_1, \dots, a_n$ eine beliebige Basis. Daraus konstruieren wir induktiv eine orthogonale Basis vermöge

$$b_{k+1} := a_{k+1} - \sum_{j=1}^{k} \frac{\overline{\langle b_j, a_{k+1} \rangle}}{\langle b_j, b_j \rangle} b_j.$$

Nach Definition handelt es sich weiterhin um eine Basis. Dabei folgen die lineare Unabhängigkeit und die Erzeugniseigenschaft induktiv. Insbesondere gilt also $b_j \neq 0$ für alle j. Die Familie der b_j ist orthogonal, da wir für $i \in \{1, \dots, k\}$ induktiv

$$\begin{aligned} \langle b_i, b_{k+1} \rangle &= \langle b_i, a_{k+1} \rangle - \sum_{j=1}^{k} \frac{\langle b_j, a_{k+1} \rangle}{\langle b_j, b_j \rangle} \langle b_i, b_j \rangle \\ &= \langle b_i, a_{k+1} \rangle - \langle b_i, a_{k+1} \rangle \\ &= 0 \end{aligned}$$

erhalten. Wir normieren nun die Vektoren b_i. Nach Theorem 2.4.5 sind sie linear unabhängig. Aufgrund ihrer Anzahl handelt es sich somit um eine Orthonormalbasis. □

Bemerkung 2.4.8. Um die Fälle $F = \mathbb{R}$ und $F = \mathbb{C}$ gleichzeitig behandeln zu können, setzen wir $A^* := \bar{A}^T$ für $A \in F^{n \times n}$. Im Reellen gilt $A^* = A^T$.

Theorem 2.4.9. *Sei V ein Skalarproduktraum,* $\dim V < \infty$. *Sei $\{b_1, \dots, b_n\}$ eine Orthonormalbasis. Definiere Vektoren d_k durch*

$$d_k = \sum_{i=1}^{n} a_k^i b_i.$$

Setze $A := \left(a_j^i\right)_{1 \leq i,j \leq n}$. Dann ist $\{d_1, \dots, d_n\}$ genau dann eine Orthonormalbasis von V, wenn

$$A^* A = I$$

gilt.

Beweis. Wie angekündigt zeigen wir nur den unitären Fall. Es gilt

$$\langle d_k, d_l\rangle = \left\langle \sum_{i=1}^{n} a_k^i b_i, \sum_{j=1}^{n} a_l^j b_j \right\rangle = \sum_{i,j=1}^{n} a_k^i \overline{a_l^j} \langle b_i, b_j\rangle$$
$$= \sum_{i,j=1}^{n} a_k^i \overline{a_l^j} \delta_{ij} = \sum_{i=1}^{n} a_k^i \overline{a_l^i}.$$

Sei also $\{d_1, \ldots, d_n\}$ orthonormal. Dann folgt

$$\delta_{kl} = \sum_{i=1}^{n} a_k^i \overline{a_l^i} \quad \text{oder} \quad A^T \bar{A} = I.$$

Durch komplexes Konjugieren erhalten wir $A^* A = I$.

Gelte umgekehrt $A^* A = I$. Mit der obigen Rechnung erhalten wir daher $\langle d_k, d_l\rangle = \delta_{kl}$. Somit ist $\{d_1, \ldots, d_n\}$ orthonormiert und daher auch linear unabhängig; es handelt sich somit um eine Orthonormalbasis von V. □

Definition 2.4.10. Sei A eine $(n \times n)$-Matrix mit $A^* = A^{-1}$. Ist $A \in \mathbb{R}^{n\times n}$, so heißt A orthogonal; ist $A \in \mathbb{C}^{n\times n}$, so heißt A unitär.

Theorem 2.4.11. *Sei $A \in \mathbb{R}^{n\times n}$. Dann sind die folgenden Aussagen äquivalent:*

(i) *A ist orthogonal*
(ii) *$A^T A = I$*
(iii) *$AA^T = I$*
(iv) *A vermittelt eine Basistransformation zwischen orthonormierten Basen eines n-dimensionalen euklidischen Vektorraumes.*
(v) *Die Spalten der Matrix A bilden eine orthonormierte Basis des Vektorraumes $\mathbb{R}^n$ mit Standardskalarprodukt.*
(vi) *Die Zeilen der Matrix A bilden eine orthonormierte Basis des Vektorraumes $\mathbb{R}^n$ mit Standardskalarprodukt.*

Beweis. Nach Definition sind die Aussagen (i), (ii) und (iii) äquivalent. Theorem 2.4.9 impliziert, dass (ii) und (iv) äquivalent sind.

Sei $A = \left(a_j^i\right)_{1\le i,j\le n}$. Dann folgt aus (ii)

$$\sum_{i=1}^{n} a_k^i a_l^i = \delta_{kl} \quad \text{für } 1 \le k, l \le n.$$

(v) folgt. Die Umkehrung „(v) $\Longrightarrow$ (ii)“ folgt ebenso aus dieser Gleichung.

Genauso wie man die Äquivalenz zwischen (ii) und (v) zeigt, erhält man auch die Äquivalenz von (iii) und (vi). □

Theorem 2.4.12. *Sei $A \in \mathbb{C}^{n\times n}$. Dann sind die folgenden Aussagen äquivalent:*

(i) *A ist unitär*
(ii) *$A^* A = I$*
(iii) *$AA^* = I$*
(iv) *A vermittelt eine Basistransformation zwischen orthonormierten Basen eines n-dimensionalen unitären Vektorraumes.*
(v) *Die Spalten der Matrix A bilden eine orthonormierte Basis des Vektorraumes $\mathbb{C}^n$ mit Standardskalarprodukt.*
(vi) *Die Zeilen der Matrix A bilden eine orthonormierte Basis des Vektorraumes $\mathbb{C}^n$ mit Standardskalarprodukt.*

Beweis. Vollständig analog zum reellen Fall. □

2.5. **Orthogonale und unitäre Endomorphismen.**

Definition 2.5.1. Sei V ein Skalarproduktraum. Dann heißt ein Endomorphismus $f\colon V \to V$ orthogonal (bzw. unitär), falls

$$\langle f(a), f(b)\rangle = \langle a, b\rangle \quad \text{für alle } a, b \in V$$

gilt.

Theorem 2.5.2. *Sei V ein endlichdimensionaler Skalarproduktraum. Sei $f\colon V \to V$ ein orthogonaler (bzw. unitärer) Endomorphismus. Dann gilt*

(i) $\|f(a)\| = \|a\|$
(ii) Sind a und b orthogonal, so auch $f(a)$ und $f(b)$.
(iii) Ist λ ein Eigenwert von f, so gilt $|\lambda| = 1$.
(iv) f ist eine Isomorphismus.

Beweis.

(i) Es ist
$$\|f(a)\| = \sqrt{\langle f(a), f(a)\rangle} = \sqrt{\langle a, a\rangle} = \|a\|.$$

(ii) Dies folgt aus
$$0 = \langle a, b\rangle = \langle f(a), f(b)\rangle.$$

(iii) Ist a ein Eigenvektor zum Eigenwert λ, so gilt nach nach (i)
$$\|a\| = \|f(a)\| = \|\lambda a\| = |\lambda| \cdot \|a\|.$$
Wegen $a \neq 0$ folgt die Behauptung.

(iv) Wegen (iii) ist 0 kein Eigenwert. Also folgt $\ker f = \{0\}$, f ist also injektiv. Da V endlichdimensional ist, ist f auch surjektiv und damit ein Isomorphismus.

(Wir bemerken, dass (iv) für unendlichdimensionale Skalarproduktträume nicht aus unserer Definition folgt: Betrachte $f\colon l^2 \to l^2$ mit $\langle (a_i)_{i\in\mathbb{N}}, (b_j)_{j\in\mathbb{N}}\rangle := \sum\limits_{i\in\mathbb{N}} a_i\bar{b}_i$ und $l^2 \equiv l^2(\mathbb{N}, \mathbb{C}) \subset \mathbb{C}^{\mathbb{N}}$ der Teilmenge, auf der die zugehörige Norm endlich ist, sowie
$$f(a_0, a_1, a_2, a_3, \ldots) := (0, a_0, a_1, a_2, a_3, \ldots).$$
Dann ist f nicht surjektiv. Die anderen Teilaussagen gelten mit demselben Beweis auch für unendlichdimensionale Skalarproduktträume.) □

Theorem 2.5.3. *Sei $\{a_1, \ldots, a_n\}$ eine Orthonormalbasis von V. Dann ist ein Endomorphismus $f\colon V \to V$ genau dann orthogonal (bzw. unitär), falls*

$$\langle f(a_i), f(a_j)\rangle = \delta_{ij}$$

für alle $1 \leq i, j \leq n$ gilt.

Beweis. Es ist klar, dass aus der Orthogonalität (bzw. Unitarität) $\langle f(a_i), f(a_j)\rangle = \delta_{ij}$ folgt.

Gelte also $\langle f(a_i), f(a_j)\rangle = \delta_{ij}$. Wir wollen nachweisen, dass f „das Skalarprodukt erhält“: Seien $a = \sum\limits_{i=1}^{n} \lambda^i a_i$ und $b = \sum\limits_{j=1}^{n} \mu^j a_j$. Wir erhalten

$$\begin{aligned}
\langle f(a), f(b)\rangle &= \left\langle f\left(\sum_{i=1}^{n} \lambda^i a_i\right), f\left(\sum_{j=1}^{n} \mu^j a_j\right)\right\rangle \\
&= \sum_{i,j=1}^{n} \lambda^i \overline{\mu^j} \underbrace{\langle f(a_i), f(a_j)\rangle}_{=\delta_{ij}=\langle a_i, a_j\rangle} \\
&= \left\langle \sum_{i=1}^{n} \lambda^i a_i, \sum_{j=1}^{n} \mu^j a_j \right\rangle = \langle a, b\rangle.
\end{aligned}$$

Somit ist f orthogonal (bzw. unitär). □

Theorem 2.5.4. *Sei V ein endlichdimensionaler Skalarproduktraum. Sei S eine Orthonormalbasis. Dann ist ein Endomorphismus $f\colon V \to V$ genau dann orthogonal (bzw. unitär), wenn die Matrix A von f bezüglich S orthogonal (bzw. unitär) ist, d. h. wenn $A^*A = I$ gilt.*

Beweis. Sei $S = \{b_1, \dots, b_n\}$ und $A = \left(a_j^i\right)_{1\le i,j\le n}$, d. h. es gelte

$$f(b_k) = \sum_{i=1}^{n} a_k^i b_i.$$

Wir erhalten

$$\langle f(b_k), f(b_l)\rangle = \sum_{i,j=1}^{n} a_k^i \overline{a_l^j} \underbrace{\langle b_i, b_j\rangle}_{=\delta_{ij}} = \sum_{i=1}^{n} \overline{a_l^i} a_k^i.$$

Nach Theorem 2.5.3 ist f genau dann orthogonal, wenn $\langle f(b_k), f(b_l)\rangle = \delta_{kl}$ gilt. Aufgrund unserer Rechnung ist dies aber äquivalent zu $\delta_{kl} = \sum_{i=1}^{n} \overline{a_l^i} a_k^i$ und somit auch zu $A^*A = I$. Die Behauptung folgt. □

Beispiel 2.5.5. Bezüglich des Standardskalarproduktes des $\mathbb{R}^2$ sind die Matrizen

$$\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

für beliebige $\varphi \in \mathbb{R}$ orthogonal, insbesondere also auch I und $-I$.

Theorem 2.5.6. *Seien A, B orthogonale (bzw. unitäre) $(n \times n)$-Matrizen. Dann gelten*

(i) A^{-1} *ist orthogonal (bzw. unitär).*
(ii) AB *ist orthogonal (bzw. unitär).*
(iii) $|\det A| = 1$.

Die orthogonalen (bzw. unitären) $(n \times n)$-Matrizen bilden somit eine Untergruppe der invertierbaren Matrizen $Gl_n(\mathbb{K})$,

- *die orthogonale Gruppe $O(n)$ bzw.*
- *die unitäre Gruppe $U(n)$.*

Weiterhin definiert man:

- Die spezielle lineare Gruppe $SL(n, F) \equiv SL_n(F)$ besteht aus den Elementen der "general linear group" $GL(n, F) \equiv Gl_n(F)$ mit Determinante 1.
- Die spezielle orthogonale Gruppe $SO(n)$ besteht aus den orthogonalen Matrizen A mit $\det A = 1$.

Der Determinantenmultiplikationssatz liefert, dass es sich damit tatsächlich um Untergruppen von $Gl_n(F)$ bzw. $O(n)$ handelt.

Beweis.

(i) Nach Definition ist $A^*A = I$. Weiterhin gilt $AA^* = I$. Also ist A^* die Inverse zu A, $A^{-1} = A^*$. Es folgt aus

$$\left(A^{-1}\right)^* A^{-1} = (A^*)^* A^* = AA^* = I,$$

dass auch A^{-1} orthogonal ist.

(ii) Dies folgt aus

$$(AB)^*(AB) = B^* (A^*A) B = B^* I B = B^* B = I.$$

(iii) Es gilt

$$\begin{aligned} 1 = \det I = \det(A^* A) &= \det A^* \cdot \det A = \det \overline{A}^T \cdot \det A \\ &= \det \overline{A} \cdot \det A = \overline{\det A} \cdot \det A = |\det A|^2. \end{aligned}$$

Die Behauptung folgt. □

2.6. **Orthogonale Komplemente.**

Definition 2.6.1. Sei V ein Skalarproduktraum und sei $S \subset V$ beliebig. Dann heißt

$$S^\perp := \{a \in V : \langle a, b\rangle = 0 \quad \text{für alle } b \in S\}$$

das orthogonale Komplement von S in V.

Theorem 2.6.2. *Sei V ein Skalarproduktraum und sei $S \subset V$. Dann gelten*

(i) $S^\perp$ *ist ein Unterraum von* V,
(ii) $\langle S\rangle \cap S^\perp = \{0\}$,
(iii) $\left(S^\perp\right)^\perp \supset \langle S\rangle$ *und*
(iv) $S^\perp = \langle S\rangle^\perp$.

Beweis.

(i) Seien $a, b \in S^\perp$ und $\lambda, \mu \in F$. Dann gilt für alle $c \in S$

$$\langle \lambda a + \mu b, c\rangle = \lambda\langle a, c\rangle + \mu\langle b, c\rangle = 0.$$

Somit ist auch $\lambda a + \mu b \in S^\perp$. Wegen $0 \in S^\perp$ ist $S^\perp$ nicht leer.

(ii) Sei $a \in S^\perp \cap \langle S\rangle$. Dann gibt es $\lambda^i \in F$ und $a_i \in S$ mit

$$\sum_{i=1}^{n} \lambda^i a_i = a.$$

Wegen $a \in S^\perp$ folgt hieraus

$$\langle a, a\rangle = \sum_{i=1}^{n} \overline{\lambda^i}\langle a, a_i\rangle = 0.$$

Somit ist $a = 0$ und wir erhalten die Behauptung.

(iii) Seien $a \in S$ und $b \in S^\perp$. Dann gilt $\langle a, b\rangle = \overline{\langle b, a\rangle} = 0$. Somit ist $S \subset \left(S^\perp\right)^\perp$. Da $\left(S^\perp\right)^\perp$ ein Unterraum ist, folgt sogar $\langle S\rangle \subset \left(S^\perp\right)^\perp$.

(Ist $\dim V < \infty$, so gilt sogar Gleichheit (s. u.). Im Raum $l^2 \equiv l^2(\mathbb{N}) \equiv l^2(\mathbb{N}, \mathbb{R}) \subset \mathbb{R}^{\mathbb{N}}$ ist $S = \{e_0, e_1, e_2, \ldots\}$ mit $S^\perp = \{0\}$ aber $\langle S\rangle \subsetneq l^2$ ein Gegenbeispiel, da $l^2(\mathbb{N}, \mathbb{R})$ auch Folgen mit unendlich vielen von Null verschiedenen Einträgen enthält.)

(iv) Wegen $S \subset \langle S\rangle$ ist klar, dass $\langle S\rangle^\perp \subset S^\perp$ gilt. Sei also $a \in S^\perp$ und $b = \sum_{i=1}^{n} \lambda^i b_i$ mit $b_i \in S$ ein beliebiges Element in $\langle S\rangle$. Dann gilt

$$\langle a, b\rangle = \sum_{i=1}^{n} \overline{\lambda^i}\langle a, b_i\rangle = 0.$$

Somit gilt auch $a \in \langle S\rangle^\perp$. □

Theorem 2.6.3. *Sei V ein endlichdimensionaler Skalarproduktraum. Ist $U \subset V$ ein Unterraum, so gilt $V = U \oplus U^\perp$.*

Beweis. Nach Theorem 2.6.2 gilt $U \cap U^\perp = \{0\}$. Es genügt also nachzuweisen, dass sich jeder Vektor $c \in V$ in der Form $c = a + b$ mit $a \in U$ und $b \in U^\perp$ darstellen lässt. Sei $\{a_1, \ldots, a_r\}$ eine Orthonormalbasis von U. Definiere a durch

$$a := \langle c, a_1\rangle a_1 + \ldots + \langle c, a_r\rangle a_r$$

und setze $b := c - a$. Dann ist offensichtlicherweise $c = a + b$. Wir müssen also noch zeigen, dass $b \in U^\perp$ gilt: Es ist

$$\begin{aligned}\langle b, a_i\rangle &= \langle c - a, a_i\rangle \\ &= \langle c, a_i\rangle - \left\langle \sum_{j=1}^{r} \langle c, a_j\rangle a_j, a_i \right\rangle \\ &= \langle c, a_i\rangle - \sum_{j=1}^{r} \langle c, a_j\rangle \delta_{ji} = 0.\end{aligned}$$

Somit steht b orthogonal zu einer Basis von U. Wir erhalten $b \in U^\perp$ und die Behauptung folgt. □

Theorem 2.6.4. *Sei V ein endlichdimensionaler Skalarproduktraum. Ist $U \subset V$ ein Unterraum, so gelten*

(i) $\dim V = \dim U + \dim U^\perp$.
(ii) $\left(U^\perp\right)^\perp = U$,

Beweis.

(i) Dies folgt direkt aus Theorem 2.6.3 und der Dimensionsformel für Vektorräume.
(ii) Es gilt $V = U \oplus U^\perp$ und $V = U^\perp \oplus \left(U^\perp\right)^\perp$. Nach Theorem 2.6.2 gilt aber $U \subset \left(U^\perp\right)^\perp$. Dies ist aber nur möglich, wenn bereits $U = \left(U^\perp\right)^\perp$ gilt. □

Beispiel 2.6.5. Sei V der Vektorraum der auf $[-a, a]$, $a > 0$, stetigen reellwertigen Funktionen mit (L^2-)Skalarprodukt

$$\langle f, g\rangle := \int_{-a}^{a} f(x)g(x)\,dx.$$

Sei $U := \{f \in V\colon f(-x) = -f(x) \text{ für alle } x \in [-a, a]\}$ der Unterraum der ungeraden Funktionen und $G := \{f \in V\colon f(-x) = f(x) \text{ für alle } x \in [-a, a]\}$ der Unterraum der geraden Funktionen in V. Wir behaupten, dass $U^\perp = G$ gilt.

Beweis. Zunächst einmal ist klar, dass $G \subset U^\perp$ gilt. Sei nun $h \in U^\perp$ beliebig. Wir setzen $h_1(x) := \frac{1}{2}(h(x) + h(-x))$ und $h_2(x) := \frac{1}{2}(h(x) - h(-x))$. Dann gilt $h = h_1 + h_2$. Es ist $h_1 \in G$ und $h_2 \in U$. Wir sind fertig, wenn wir $h_2 \equiv 0$ zeigen können. Aus $h \in U^\perp$ und $h_2 \in U$ erhalten wir

$$0 = \langle h, h_2\rangle = \int_{-a}^{a} h(x)h_2(x)\,dx = \int_{-a}^{a} h_1(x)h_2(x)\,dx + \int_{-a}^{a} h_2(x)h_2(x)\,dx.$$

Das erste Integral auf der rechten Seite mit einer geraden und einer ungeraden Funktion verschwindet. Da h und somit auch h_2 stetig ist, folgt also $h_2 \equiv 0$ wie behauptet. □

2.7. **Adjungierte Abbildungen.** Zu einem Endomorphismus $f\colon V \to V$ wollen wir eine lineare Abbildung $g\colon V \to V$ mit

$$\langle f(a), b\rangle = \langle a, g(b)\rangle$$

finden.

Definition 2.7.1. Sei V ein Skalarproduktraum. Zu $a \in V$ definieren wir $a^*\colon V \to \mathbb{R}$ (bzw. $a^*\colon V \to \mathbb{C}$) durch

$$a^*(b) = \langle b, a\rangle \quad \text{für } b \in V.$$

Bemerkung: Die Abbildung a^* ist linear, denn es gilt

$$a^*(\lambda b + \mu c) = \langle \lambda b + \mu c, a\rangle = \lambda\langle b, a\rangle + \mu\langle c, a\rangle = \lambda a^*(b) + \mu a^*(c).$$

Theorem 2.7.2. *Sei V ein endlichdimensionaler Skalarproduktraum. Dann gibt es zu jedem Funktional φ ein eindeutig bestimmtes $a \in V$ mit $\varphi(\xi) = \langle \xi, a\rangle$ für alle $\xi \in V$. Somit ist $\varphi = a^*$.*

Beweis. Sei $\{a_1, \dots, a_n\}$ eine orthonormierte Basis von V. Definiere a durch

$$a := \sum_{i=1}^{n} \overline{\varphi(a_i)} a_i.$$

Wir erhalten

$$\langle a_j, a\rangle = \sum_{i=1}^{n} \overline{\overline{\varphi(a_i)}}\langle a_j a_i\rangle = \varphi(a_j).$$

Also stimmen a^* und φ auf einer Basis von V überein und sind daher gleich.

Es bleibt noch zu zeigen, dass a eindeutig bestimmt ist. Gelte also $\langle \xi, a\rangle = \langle \xi, a'\rangle$ für alle $\xi \in V$. Dies ist äquivalent zu $\langle \xi, a - a'\rangle = 0$ für alle $\xi \in V$ und gilt insbesondere für $\xi = a - a'$. Also folgt $a = a'$. □

Bemerkung 2.7.3. Aus diesem Beweis folgt insbesondere, dass $\langle \xi, a\rangle = 0$ für alle $\xi \in V$ nur für $a = 0$ erfüllt sein kann.

Bemerkung 2.7.4. Sei $f\colon V \to V$ ein Endomorphismus. Ist $a \in V$ fest, so definieren wir durch

$$\varphi(\xi) := \langle f(\xi), a\rangle$$

ein lineares Funktional $\varphi\colon V \to F$: Es gilt nämlich

$$\varphi(\lambda b + \mu c) = \langle f(\lambda b + \mu c), a\rangle = \lambda\langle f(b), a\rangle + \mu\langle f(c), a\rangle = \lambda\varphi(b) + \mu\varphi(c).$$

Nach Theorem 2.7.2 gibt es zu φ ein eindeutig bestimmtes $a_0 \in V$ mit $\varphi(\xi) = \langle \xi, a_0\rangle$ für alle $\xi \in V$. Für fixiertes f ordnen wir auf diese Weise jedem $a \in V$ ein $a_0 \in V$ zu. Wir bezeichnen diese Zuordnung als $g\colon V \to V$, $g(a) := a_0$. g ist durch die Gleichung

$$\langle f(\xi), a\rangle = \langle \xi, g(a)\rangle \quad \text{für alle } a, \xi \in V$$

festgelegt. Wir behaupten, dass g eine lineare Abbildung ist. Aus der definierenden Gleichung erhalten wir

$$\begin{aligned}\langle \xi, g(\lambda a + \mu b)\rangle &= \langle f(\xi), \lambda a + \mu b\rangle = \overline{\lambda}\langle f(\xi), a\rangle + \overline{\mu}\langle f(\xi), b\rangle \\ &= \overline{\lambda}\langle \xi, g(a)\rangle + \overline{\mu}\langle \xi, g(b)\rangle = \langle \xi, \lambda g(a) + \mu g(b)\rangle.\end{aligned}$$

Da dies für alle $\xi \in V$ gilt, erhalten wir $g(\lambda a + \mu b) = \lambda g(a) + \mu g(b)$, also die Linearität von g.

Wir sagen, dass g die zu f adjungierte Abbildung ist und schreiben $g = f^*$.

Definition 2.7.5. Die zum Endomorphismus $f\colon V \to V$ adjungierte Abbildung $f^*\colon V \to V$ ist durch

$$\langle f(\xi), \eta\rangle = \langle \xi, f^*(\eta)\rangle \quad \text{für alle } \xi, \eta \in V$$

definiert.

Beispiele 2.7.6.

(i) Sei $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ durch

$$f\left(x^1, x^2, x^3\right) := \left(x^1 - x^2, -x^1 + x^2 + 2x^3, x^2 + x^3\right)$$

gegeben. Nach Definition gilt $\langle f(x), y\rangle = \langle x, g(y)\rangle$. Somit erhalten wir aus

$$\begin{aligned}\langle f(x), y\rangle &= \left(x^1 - x^2\right) y^1 + \left(-x^1 + x^2 + 2x^3\right) y^2 + \left(x^2 + x^3\right) y^3\\ &= x^1\left(y^1 - y^2\right) + x^2\left(-y^1 + y^2 + y^3\right) + x^3\left(2y^2 + y^3\right)\\ &= \langle x, f^*(y)\rangle\end{aligned}$$

die adjungierte Abbildung

$$f^*\left(y^1, y^2, y^3\right) := \left(y^1 - y^2, -y^1 + y^2 + y^3, 2y^2 + y^3\right).$$

Die darstellenden Matrizen bezüglich der Standardbasis sind

$$A_f = \begin{pmatrix} 1 & -1 & 0\\ -1 & 1 & 2\\ 0 & 1 & 1\end{pmatrix} \quad \text{sowie} \quad A_{f^*} = \begin{pmatrix} 1 & -1 & 0\\ -1 & 1 & 1\\ 0 & 2 & 1\end{pmatrix}.$$

(ii) Im Raum $V = L^2([a,b])$ der auf $[a,b]$ stetigen reellwertigen Funktionen mit L^2-Skalarprodukt ist

$$\begin{aligned}\Phi&\colon V \to V,\\ \Phi(f)(x) &:= x \cdot f(x)\end{aligned}$$

selbstadjungiert, d. h. es gilt $\Phi^* = \Phi$, denn es gilt

$$\langle \Phi(f), g\rangle = \langle f, \Phi(g)\rangle.$$

Den Zusammenhang zwischen der f und der f^* darstellenden Matrix haben wir bereits im Beispiel gesehen.

Theorem 2.7.7. *Sei V ein endlichdimensionaler Skalarproduktraum. Sei S eine Orthonormalbasis. Wird $f\colon V \to V$ bezüglich S durch die Matrix A dargestellt, so wird die adjungierte Abbildung $f^*\colon V \to V$ bezüglich S durch die Matrix A^* dargestellt.*

Beweis. Seien $S = (d_1, \ldots, d_n)$, $A = \left(a^i_j\right)_{1\le i,j\le n}$ und $A_{f^*} = \left(b^i_j\right)_{1\le i,j\le n}$. Wir erhalten

$$\begin{aligned} f(d_k) &= \sum_{i=1}^n a^i_k d_i,\\ f^*(d_l) &= \sum_{j=1}^n b^j_l d_j,\\ \langle f(d_k), d_l\rangle &= \langle d_k, f^*(d_l)\rangle,\\ \langle f(d_k), d_l\rangle &= \left\langle \sum_{i=1}^n a^i_k d_i, d_l \right\rangle = \sum_{i=1}^n a^i_k \langle d_i, d_l\rangle = \sum_{i=1}^n a^i_k \delta_{il} = a^l_k,\end{aligned}$$

$$\langle d_k, f^*(d_l)\rangle = \left\langle d_k, \sum_{j=1}^{n} b_l^j d_j \right\rangle = \sum_{j=1}^{n} \overline{b_l^j}\delta_{kj} = \overline{b_l^k}.$$

(In den letzten beiden Zeilen sind die jeweils letzten Gleichheiten aus Kovarianzgründen unschön geschrieben. Für die zugehörigen reellen bzw. komplexen Zahlen gilt die Gleichheit jedoch.) Also ist $a_k^l = \overline{b_l^k}$ für alle $1 \leq k, l \leq n$. Die Behauptung folgt. □

Theorem 2.7.8. *Sei V ein endlichdimensionaler Skalarproduktraum. Sei $f\colon V \to V$ linear. Dann gilt*

$$\ker f^* = (\operatorname{im} f)^\perp.$$

Beweis. Es gilt

$$\begin{aligned}\ker f^* &= \{\xi \in V\colon f^*(\xi) = 0\}\\ &= \{\xi \in V\colon \langle \eta, f^*(\xi)\rangle = 0 \quad \text{für alle } \eta \in V\}\\ &= \{\xi \in V\colon \langle f(\eta), \xi\rangle = 0 \quad \text{für alle } \eta \in V\}\\ &= (\operatorname{im} f)^\perp.\end{aligned}$$

□

Beispiel 2.7.9. Sei

$$\sum_{k=1}^{n} a_k^i x^k = b^i, \quad 1 \leq i \leq n$$

ein reelles lineares quadratisches Gleichungssystem. Sei $f\colon \mathbb{R}^n \to \mathbb{R}^n$ die (bezüglich der Standardbasis) durch $A = \left(a_j^i\right)$ dargestellte Abbildung. Dann ist das lineare Gleichungssystem genau dann lösbar, wenn $b = \left(b^1, \ldots, b^n\right) \in \operatorname{im} f$ gilt. Auf $\mathbb{R}^n$ führen wir das Standardskalarprodukt ein und erhalten aus Theorem 2.7.8, dass das lineare Gleichungssystem genau dann lösbar ist, wenn $b \in (\ker f^*)^\perp$ gilt. Über $\mathbb{R}$ wird f^* durch die transponierte Matrix dargestellt. Somit folgt: Das lineare Gleichungssystem ist genau dann lösbar, wenn für jede Lösung $y = (y_1, \ldots, y_n)$ von

$$\sum_{i=1}^{n} a_k^i y_i = 0 \quad \text{auch} \quad \sum_{i=1}^{n} y_i b^i = 0 \quad \text{gilt.}$$

Theorem 2.7.10. *Sei V ein endlichdimensionaler Skalarproduktraum. Sei $f\colon V \to V$ ein Endomorphismus. Dann haben f und f^* den gleichen Rang.*

Beweis. Es gilt

$$\begin{aligned}\operatorname{rang} f &= \dim \operatorname{im} f = \dim V - \dim(\operatorname{im} f)^\perp\\ &= \dim V - \dim(\ker f^*) && \text{(Theorem 2.7.8)}\\ &= \dim(\operatorname{im} f^*) = \operatorname{rang} f^*.\end{aligned}$$

(Alternativ kann man über $\mathbb{R}$ wegen $A^* = A^T$ benutzen, dass Zeilenrang und Spaltenrang übereinstimmen.) □

Korollar 2.7.11. *Sei $A \in \mathbb{C}^{n\times n}$. Dann gilt* $\operatorname{rang} A = \operatorname{rang} \overline{A}$.

Beweis. Setze $B := \overline{A}^T$. Da der Rang einer Matrix B mit dem von B^T übereinstimmt, folgt

$$\operatorname{rang} A = \operatorname{rang} B^* = \operatorname{rang} B = \operatorname{rang} B^T = \operatorname{rang} \overline{A}.$$

□

2.8. **Diagonalisierung von selbstadjungierten linearen Endomorphismen.** Wir wollen insbesondere zeigen, dass sich solche Endomorphismen sogar mit Hilfe einer Orthogonalbasis stets diagonalisieren lassen.

Definition 2.8.1. Sei V ein Skalarproduktraum. Der Endomorphismus $f\colon V \to V$ heißt selbstadjungiert, wenn

$$\langle f(\xi), \eta\rangle = \langle \xi, f(\eta)\rangle$$

für alle $\xi, \eta \in V$ gilt, also $f = f^*$ ist.

Theorem 2.8.2. *Sei V ein endlichdimensionaler Skalarproduktraum mit Orthonormalbasis S. Der Endomorphismus $f\colon V \to V$ werde bezüglich S durch die Matrix A beschrieben. Dann ist $f\colon V \to V$ genau dann selbstadjungiert, wenn $A = A^*$ gilt.*

Beweis. Dies folgt direkt aus Theorem 2.7.7. □

Theorem 2.8.3. *Sei V ein n-dimensionaler Skalarproduktraum. Sei $f\colon V \to V$ ein selbstadjungierter Endomorphismus. Dann sind alle Nullstellen des charakteristischen Polynoms $\chi_f(X)$ reell, es gilt also*

$$\chi_f(X) = \pm(X - \lambda_1)\cdots(X - \lambda_n)$$

für $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$.

Beweis. Wir betrachten zunächst den unitären Fall. Wie jedes Polynom zerfällt das charakteristische Polynom in Linearfaktoren und wir müssen lediglich nachweisen, dass die Eigenwerte von f reell sind. Sei also λ ein Eigenwert von f zum Eigenvektor ξ. Wir erhalten

$$\lambda\langle\xi,\xi\rangle = \langle\lambda\xi,\xi\rangle = \langle f(\xi),\xi\rangle = \langle\xi, f(\xi)\rangle = \langle\xi,\lambda\xi\rangle = \overline{\lambda}\langle\xi,\xi\rangle.$$

Wegen $\xi \neq 0$ bzw. $\langle\xi,\xi\rangle \neq 0$ folgt $\lambda = \overline{\lambda}$. Somit ist λ reell.

Im reellen Fall benutzen wir die sogenannte Komplexifizierung. Die zu $f\colon V \to V$ gehörige reelle symmetrische Matrix definiert eine Abbildung $\tilde{f}\colon \mathbb{C}^n \to \mathbb{C}^n$. Eine reelle symmetrische Matrix ist hermitesch, wenn wir sie als komplexe Matrix auffassen. Somit ist $\tilde{f}\colon \mathbb{C} \to \mathbb{C}$ bezüglich des Standardskalarproduktes auf $\mathbb{C}^n$ selbstadjungiert. Es gilt

$$\chi_f(X) = \chi_A(X) = \chi_{\tilde{f}}(X).$$

Somit stimmen die Eigenwerte von f und $\tilde{f}$ überein und sind aufgrund der obigen Überlegungen reell. □

Theorem 2.8.4. *Sei V ein endlichdimensionaler Skalarproduktraum. Sei $f\colon V \to V$ ein selbstadjungierter Endomorphismus. Dann besitzt V eine Orthonormalbasis aus Eigenvektoren von f.*

Beweis. Wir wissen, dass das charakteristische Polynom $\chi_f(X)$ in Linearfaktoren zerfällt. (Im Fall $F = \mathbb{C}$ folgt dies aus dem Fundamentalsatz der Algebra und für $F = \mathbb{R}$ (und $\mathbb{C}$) aus Theorem 2.8.3.)

Benutze Induktion nach $\dim V =: n$. Für $n = 1$ ist die Aussage wahr. Sei also $n \geq 2$. Sei a_1 ein Eigenvektor zum Eigenwert λ mit $\|a_1\| = 1$. Definiere $W := \{\xi \in V\colon \langle a_1, \xi\rangle = 0\}$. Nach Theorem 2.6.4 folgt $\dim W = n - 1$. W ist ein bezüglich f invarianter Unterraum, es gilt nämlich für $\xi \in W$

$$\langle a_1, f(\xi)\rangle = \langle f(a_1), \xi\rangle = \langle\lambda a_1, \xi\rangle = \lambda\langle a_1, \xi\rangle = 0.$$

Es folgt $f(\xi) \in W$. Nach Induktionsvoraussetzung existiert eine Orthonormalbasis $\{a_2, \ldots, a_n\}$ von W aus Eigenvektoren von $f|_W$ und somit auch von f. (Dabei verwenden wir das auf W eingeschränkte Skalarprodukt. Man überzeugt sich

leicht davon, dass diese Einschränkung wieder ein Skalarprodukt ist.) Also ist $\{a_1, a_2, \ldots, a_n\}$ wie behauptet eine Orthonormalbasis. □

Theorem 2.8.5. *Jeder selbstadjungierte Endomorphismus eines endlichdimensionalen Skalarproduktraumes ist diagonalisierbar.*

Beweis. Benutze eine Orthonormalbasis aus Eigenvektoren. Bezüglich dieser ist die zugehörige Matrix eine Diagonalmatrix mit den Eigenwerten entsprechend ihrer Vielfachheit auf der Diagonalen. □

Theorem 2.8.6. *Sei A eine reelle symmetrische (bzw. komplexe hermitesche) $(n \times n)$-Matrix. Dann gibt es eine orthogonale (bzw. unitäre) Matrix D mit*

$$DAD^* = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}$$

mit $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$.

Beweis. Dies folgt direkt aus den bisherigen Theoremen dieses Kapitels, da für orthogonale (bzw. unitäre) Matrizen $D^{-1} = D^*$ gilt. □

2.9. **Kästchenform orthogonaler Matrizen.** Wir wollen unitäre und orthogonale Endomorphismen bezüglich Orthonormalbasen durch „einfache“ Matrizen darstellen.

Theorem 2.9.1. *Sei $f\colon V \to V$ ein unitärer Endomorphismus eines endlichdimensionalen unitären Vektorraumes V. Dann gibt es eine Orthonormalbasis von V, die aus Eigenvektoren von f besteht.*

Beweis. Wir beweisen dies per Induktion nach $n = \dim V$. Der Induktionsanfang ist trivial. Sei $n \geq 2$. Das charakteristische Polynom zerfällt über $\mathbb{C}$ in Linearfaktoren

$$\chi_f(X) = \pm(X - \lambda_1)(X - \lambda_2) \cdots (X - \lambda_n).$$

Nach Theorem 2.5.2 gilt $|\lambda_i| = 1$. Sei a_1 ein Eigenvektor zum Eigenwert λ_1 mit $\|a_1\| = 1$. Sei W das orthogonale Komplement von a_1, $V = W \oplus \langle a_1 \rangle$. Für jedes $\xi \in W$ gilt

$$0 = \langle a_1, \xi \rangle = \langle f(a_1), f(\xi) \rangle = \lambda_1 \langle a_1, f(\xi) \rangle.$$

Da $\lambda_1 \neq 0$ gilt, folgt $f(\xi) \in W$. Somit induziert die unitäre Abbildung f eine unitäre Abbildung $f|_W\colon W \to W$. Nach Induktionsvoraussetzung gibt es für $f|_W$ eine Orthonormalbasis aus Eigenvektoren. Zusammen mit a_1 erhält man wie gewünscht eine Orthogonalbasis aus Eigenvektoren. □

Korollar 2.9.2. *Sei $A \in \mathbb{C}^{n \times n}$ unitär. Dann gibt es eine unitäre Matrix D mit*

$$DAD^* = \begin{pmatrix} e^{i\varphi_1} & 0 & \ldots & 0 \\ 0 & e^{i\varphi_2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & e^{i\varphi_n} \end{pmatrix} \quad \textit{und} \quad 0 \leq \varphi_i < 2\pi.$$

Lemma 2.9.3. *Sei $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ orthogonal. Fixiere eine Orthonormalbasis. Dann gibt es $0 \leq \varphi < 2\pi$, so dass f bezüglich dieser Basis durch*

$$\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \quad \textit{oder} \quad \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}$$

dargestellt wird.

Beweis. Wir benutzen Theorem 2.4.11. Die Spalten und die Zeilen der Matrix bilden also eine Orthonormalbasis. Einen beliebigen Einheitsvektor in $\mathbb{R}^2$ können wir als

$$\left(\cos\varphi, \pm\sqrt{1-\cos^2\varphi}\right) = (\cos\varphi, \pm\sin\varphi)$$

schreiben. Somit ist f durch eine Matrix der Form

$$\begin{pmatrix} \cos\varphi & \pm\sin\varphi \\ \pm\sin\varphi & \pm\cos\varphi \end{pmatrix}$$

dargestellt. Man überzeugt sich leicht, dass die Vorzeichen für $\sin\varphi \neq 0$ bzw. $\cos\varphi \neq 0$ wie angegeben gewählt werden müssen um eine orthogonale Matrix zu erhalten. Ggf. ist φ durch $-\varphi + 2\pi k$, $k \in \mathbb{Z}$, zu ersetzen um Vorzeichen abzuändern. □

Durch eine spezielle Basiswahl können wir Kästchen der zweiten Form ausschließen.

Theorem 2.9.4. *Sei $f\colon V \to V$ eine orthogonale Selbstabbildung eines euklidischen Vektorraumes mit $\dim V < \infty$. Dann gibt es eine Orthonormalbasis von V, in der sich f in der Form*

$$\begin{pmatrix}
\mathbb{1}_k & 0 & 0 & 0 & \dots & 0 & 0 \\
0 & -\mathbb{1}_l & 0 & 0 & \dots & 0 & 0 \\
0 & 0 & \begin{matrix}\cos\varphi_1 & -\sin\varphi_1 \\ \sin\varphi_1 & \cos\varphi_1\end{matrix} & 0 & \dots & 0 & 0 \\
0 & 0 & 0 & & & & 0 \\
\vdots & \vdots & \vdots & & \ddots & & \vdots \\
0 & 0 & 0 & & & & 0 \\
0 & 0 & 0 & 0 & \dots & 0 & \begin{matrix}\cos\varphi_m & -\sin\varphi_m \\ \sin\varphi_m & \cos\varphi_m\end{matrix}
\end{pmatrix}$$

mit $k + l + 2m = \dim V$ darstellen lässt.

(In „Matrizensprache“ übersetzt erhalten wir: Zu jeder orthogonalen Matrix A gibt es also eine orthogonale Matrix D, so dass DAD^T von der angegebenen Gestalt ist.)

Beweis. Vermöge einer Matrixdarstellung $A \in \mathbb{R}^{\dim V \times \dim V}$ ordnen wir f wiederum eine Abbildung $\tilde{f}\colon \mathbb{C}^n \to \mathbb{C}^n$ zu. Wegen $A^*A = A^TA = \mathbb{1}$ ist A als Matrix in $\mathbb{C}^{n\times n}$ unitär. Nach Theorem 2.9.1 gibt es eine Orthonormalbasis aus Eigenvektoren für $\tilde{f}$ zu Eigenwerten λ_i mit $|\lambda_i| = 1$. Zu jedem λ_i mit $\lambda_i = \pm 1$ finden wir einen reellen Eigenvektor zu diesem Eigenwert, den Realteil des komplexen Eigenvektors (falls dieser Realteil $\neq 0$ ist): Aus $A\xi = \lambda_i\xi$ folgt $A\overline{\xi} = \lambda_i\overline{\xi}$ und $A\left(\xi + \overline{\xi}\right) = \lambda_i\left(\xi + \overline{\xi}\right)$. Daher ist $\operatorname{Re}\xi$ ein Eigenvektor oder der Nullvektor. Ist $\operatorname{Re}\xi$ der Nullvektor, so ist $i\xi$ ein reeller Eigenvektor.

Da das orthogonale Komplement eines Eigenvektors jeweils invariant unter f ist, erhalten wir induktiv die beiden Blöcke $\mathbb{1}_k$ und $\mathbb{1}_l$.

Sei ξ mit $|\xi| = 1$ ein Eigenvektor zum Eigenwert $e^{i\varphi} \notin \{-1, 1\}$, $A\xi = e^{i\varphi}\xi$. Durch komplexe Konjugation erhalten wir einen zweiten Eigenvektor: $A\overline{\xi} = e^{-i\varphi}\overline{\xi}$. Setze $u := \xi + \overline{\xi}$ und $v := -i\left(\xi - \overline{\xi}\right)$. Da ξ und $\overline{\xi}$ wegen $e^{i\varphi} \notin \{-1, 1\}$ Eigenvektoren von A zu unterschiedlichen Eigenwerten sind, gilt $\xi \neq \pm\overline{\xi}$ und somit $u \neq 0 \neq v$. Wir

erhalten

$$\begin{aligned} Au &= A\xi + A\overline{\xi} = e^{i\varphi}\xi + e^{-i\varphi}\overline{\xi} = \frac{e^{i\varphi}+e^{-i\varphi}}{2}\left(\xi+\overline{\xi}\right) + \frac{e^{i\varphi}-e^{-i\varphi}}{2}\left(\xi-\overline{\xi}\right) \\ &= \cos\varphi\cdot u - \sin\varphi\cdot v, \\ -\frac{1}{i}Av &= A\xi - A\overline{\xi} = e^{i\varphi}\xi - e^{-i\varphi}\overline{\xi} = \frac{e^{i\varphi}-e^{-i\varphi}}{2}\left(\xi+\overline{\xi}\right) + \frac{e^{i\varphi}+e^{-i\varphi}}{2}\left(\xi-\overline{\xi}\right) \\ &= i\sin\varphi\cdot u + \cos\varphi\cdot\frac{v}{-i}, \\ Av &= \sin\varphi\cdot u + \cos\varphi\cdot v. \end{aligned}$$

Wir wählen $\frac{1}{\sqrt{2}}u$ und $\frac{1}{\sqrt{2}}v$ als Teil einer reellen Orthonormalbasis. Beachte dazu, dass aus $\left\langle \xi,\overline{\xi}\right\rangle_{\mathbb{C}} = 0$ die Relationen $\left\|\xi+\overline{\xi}\right\|^2 = 2$, $\left\|-i\left(\xi-\overline{\xi}\right)\right\|^2 = 2$ und $\left\langle \xi+\overline{\xi}, -i\left(\xi-\overline{\xi}\right)\right\rangle = 0$ folgen, zunächst mit komplexen Skalarprodukten und Normen. Da die Einträge aber reell sind, gelten dieselben Beziehungen auch über $\mathbb{R}$ und zugehörigen Skalarprodukten und Normen. Dies liefert den ersten Block:

$$\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} = \begin{pmatrix} \cos\varphi_1 & -\sin\varphi_1 \\ \sin\varphi_1 & \cos\varphi_1 \end{pmatrix}, \quad \varphi_1 = -\varphi.$$

Da beide Zeilen und Spalten, die durch diesen Block verlaufen, bereits „in diesem Block Länge Eins haben“, enthalten sie sonst nur Nullen. Die restlichen Blöcke erhält man per Induktion. □

2.10. **Normale Matrizen.**

Definition 2.10.1. Sei A eine quadratische Matrix über $\mathbb{R}$ oder $\mathbb{C}$. Dann heißt A normal, falls

$$A^*A = AA^*$$

gilt.

Ein Endomorphismus f eines Skalarproduktraumes heißt normal, wenn $f\circ f^* = f^*\circ f$ gilt.

Beispiele 2.10.2.

(i) Unitäre ($A^{-1} = A^*$), hermitesche ($A = A^*$) und Matrizen mit $A = -A^*$ sind normal.

(ii) Orthogonale ($A^{-1} = A^T$), symmetrische ($A = A^T$) und schiefsymmetrische Matrizen ($A = -A^T$) sind normal.

Resultate für solche Matrizen werden wir einheitlicher für normale Matrizen behandeln.

(iii) Es gibt weitere normale Matrizen, z. B.

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

denn es gilt

$$\begin{aligned} AA^T &= \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} = A^TA. \end{aligned}$$

Theorem 2.10.3. *Sei V ein endlichdimensionaler Skalarproduktraum. Sei f ein normaler Endomorphismus. Dann gelten*

$$\ker f = \ker f^* \quad \textit{und} \quad \operatorname{im} f = \operatorname{im} f^*.$$

Insbesondere folgt $V = \ker f \oplus \operatorname{im} f$*, wobei die Summanden senkrecht zueinander sind,* $\ker f \perp \operatorname{im} f$.

Beweis. Aus $v \in \ker f$ erhalten wir

$$0 = \langle f(v), f(v)\rangle = \langle v, f^*(f(v))\rangle = \langle v, f(f^*(v))\rangle = \langle f^*(v), f^*(v)\rangle.$$

Somit gilt auch $v \in \ker f^*$.

Aus Theorem 2.7.8 wissen wir, dass $\ker f^* = (\operatorname{im} f)^\perp$ gilt. Somit folgt

$$(\operatorname{im} f)^\perp = \ker f^* = \ker f = (\operatorname{im} f^*)^\perp.$$

Da V endlichdimensional ist, folgt auch $\operatorname{im} f = \operatorname{im} f^*$. □

Korollar 2.10.4. *Sei V ein endlichdimensionaler Skalarproduktraum (über $\mathbb{R}$ oder $\mathbb{C}$). Sei $f\colon V \to V$ normal. Sei $\lambda \in \mathbb{C}$ bzw. $\in \mathbb{R}$. Dann gilt für die zugehörigen Eigenräume*

$$E_\lambda(f) = E_{\bar\lambda}(f^*).$$

Beweis. Setze $g := f - \lambda \operatorname{id}$. Dann ist $g^* = f^* - \bar\lambda \operatorname{id}$. g ist ebenfalls normal, denn es gilt

$$g \circ g^* = \underbrace{f \circ f^*}_{=f^*\circ f} - \lambda f^* - \bar\lambda f + \lambda\bar\lambda \operatorname{id} = g^* \circ g.$$

Somit erhalten wir

$$E_\lambda(f) = \ker g = \ker g^* = E_{\bar\lambda}(f^*).$$ □

Wir können normale Endomorphismen nicht nur diagonalisieren; dies ist auch eine notwendige Bedingung für die orthogonale Diagonalisierbarkeit. Der folgende Satz funktioniert aber nur für komplexe Matrizen. Drehmatrizen sind Beispiele für reelle orthogonale und daher auch normale aber nicht reell diagonalisierbare Matrizen.

Theorem 2.10.5. *Sei V ein endlichdimensionaler unitärer Skalarproduktraum mit $\dim V = n$. Sei $f\colon V \to V$ linear. Dann sind die folgenden Bedingungen äquivalent:*

(i) Es gibt eine Orthonormalbasis aus Eigenvektoren von f.
(ii) f ist normal.

Beweis.
„$\Longrightarrow$“: Sei $a_1, \dots, a_n$ eine Orthonormalbasis aus Eigenvektoren von f mit $f(a_i) = \lambda_i a_i$. Wir erhalten

$$\langle a_j, f^*(a_i)\rangle = \langle f(a_j), a_i\rangle = \lambda_j\langle a_j, a_i\rangle = \left\langle a_j, \overline{\lambda_i} a_i\right\rangle.$$

Somit ist $f^*(a_i) = \bar\lambda_i a_i$. Für diese Basisvektoren gilt

$$f(f^*(a_i)) = \bar\lambda_i f(a_i) = \bar\lambda_i\lambda_i a_i = \lambda_i f^*(a_i) = f^*(f(a_i)).$$

Da es genügt, Normalität für eine Basis nachzurechnen, ist f normal.

„$\Longleftarrow$“: Sei umgekehrt f normal. Über $\mathbb{C}$ zerfällt das charakteristische Polynom in Linearfaktoren. Wähle einen normierten Eigenvektor a_1 zum Eigenvektor λ_1 und definiere $E_{\lambda_1} := \langle a_1\rangle$. Setze $W := E_{\lambda_1}^\perp$. Es gilt $\dim W = n - 1$. Können wir zeigen, dass W ein f-invarianter Unterraum von V ist und dass $f|_W$ normal ist, so folgt die Behauptung per Induktion nach der Dimension des Vektorraumes. Es gilt für $w \in W$

$$\langle f(w), a_1\rangle = \langle w, f^*(a_1)\rangle = \langle w, \bar\lambda_1 a_1\rangle = \lambda_1\langle w, a_1\rangle = 0.$$

Somit ist W ein f-invarianter Unterraum. Da E_{λ_1} auch in einem Eigenraum der normalen Abbildung f^* enthalten ist, gilt auch $f^*(W) \subset W$. Da W für f und für f^* ein invarianter Unterraum ist und $(f|_W)^* = f^*|_W$ gilt, ist $f|_W$ ebenfalls normal und die Behauptung folgt. □

Für Matrizen umformuliert erhalten wir

Korollar 2.10.6. *Sei $A \in \mathbb{C}^{n\times n}$. Dann ist A genau dann normal, wenn es eine unitäre Matrix $S \in \mathbb{C}^{n\times n}$ gibt, so dass SAS^{-1} diagonal ist.*

Theorem 2.10.7. *Seien $A, B \in \mathbb{C}^{n\times n}$ normal und gelte $[A, B] = AB - BA = 0$. Dann gibt es eine unitäre Matrix U, so dass UAU^* und UBU^* diagonal sind.*

Beweis. Sei λ_1 ein Eigenwert von A und sei V_1 der zugehörige Eigenraum. Sei $v_1 \in V_1$. Aus $AB = BA$ folgt

$$A(Bv_1) = BAv_1 = \lambda_1(Bv_1).$$

Somit gilt $Bv_1 \in V_1$. V_1 ist also ein invarianter Unterraum für A und B.

Definiere $W_1 := V_1^\perp$. Wir behaupten, dass W_1 auch ein invarianter Unterraum für A und B ist. Wir argumentieren wie in Theorem 2.10.5. Nach Korollar 2.10.4 gilt $A^*v = \bar{\lambda}_1 v$ für alle $v \in V_1$. Seien $v \in V_1$ und $w \in W_1$ beliebig. Dann gilt

$$\langle Aw, v\rangle = \langle w, A^*v\rangle = \langle w, \bar{\lambda}_1 v\rangle = \lambda_1\langle w, v\rangle = 0.$$

Wir erhalten $AW_1 \subset V_1^\perp = W_1$.

Zu $BW_1 \subset W_1$: Nach Definition von V_1 ist $V_1 = \ker(A - \lambda_1 \mathbb{1})$, wobei wir Matrizen und zugehörige Abbildungen identifizieren. Wir haben gezeigt, dass $AW_1 \subset W_1$ gilt. Somit folgt auch $(A - \lambda_1\mathbb{1})W_1 \subset W_1$. Da $V_1 \oplus W_1 = V$ gilt und W_1 endlichdimensional ist, ist $(A - \lambda_1\mathbb{1})|_{W_1} : W_1 \to W_1$ bijektiv. Also gelten $(A - \lambda_1\mathbb{1})V = W_1$ und $W_1 = (A - \lambda_1\mathbb{1})W_1$. Wir erhalten

$$BW_1 = B(A - \lambda_1\mathbb{1})W_1 = (A - \lambda_1\mathbb{1})BW_1 \subset (A - \lambda_1\mathbb{1})V = W_1.$$

Die Zerlegung $V = V_1 \oplus W_1$ in orthogonale unter A und B invariante Unterräume impliziert, dass A und B (nach einer unitären Transformation) in Blockgestalt sind. Der zu V_1 gehörige Block ist für A ein Vielfaches der Einheitsmatrix. Die Normalität überträgt sich aufgrund der Blockgestalt auf jeden der einzelnen Blöcke, ebenso $[A, B] = 0$. Mit Theorem 2.10.5 können wir den zu V_1 gehörigen Block von B diagonalisieren: Die zugehörige unitäre Transformationsmatrix läßt sich zu einer unitären Matrix von V erweitern. Der zu W_1 gehörige Block ist kleiner als die Matrizen A und B. Somit folgt die Aussage per Induktion. □

Korollar 2.10.8. *Seien $A, B \in \mathbb{C}^{n\times n}$ normale Matrizen, die kommutieren, d. h. gelte $AB = BA$. (Oder der Kommutator $[A, B] := AB - BA$ verschwindet: $[A, B] = 0$) Dann sind AB und $A + B$ ebenfalls normal.*

Beweis. Die Eigenschaften, normal zu sein und zu kommutieren gelten genau dann für A und B, wenn sie für UAU^{-1} und UBU^{-1} gelten. Verwende nun Theorem 2.10.7. Für Diagonalmatrizen ist die Aussage klar. □

Theorem 2.10.9. *Sei $A \in \mathbb{C}^{n\times n}$. Dann sind die folgenden Aussagen sind äquivalent:*

(i) *A ist normal.*
(ii) *A lässt sich durch eine unitäre Matrix diagonalisieren.*
(iii) *Es gibt eine Orthonormalbasis aus Eigenvektoren von A.*
(iv) *Es gilt $\|Ax\| = \|A^*x\|$ für alle $x \in \mathbb{C}^n$.*
(v) *Der „hermitesche Anteil" $\frac{1}{2}(A + A^*)$ und der "anti-hermitesche Anteil" $\frac{1}{2}(A - A^*)$ kommutieren.*

Beweis. Die Äquivalenz von (i), (ii) und (iii) haben wir bereits gezeigt.

Klar ist (i) $\Longrightarrow$ (iv). Gelte umgekehrt (iv). Wir erhalten nacheinander

$$\begin{aligned}\langle Ax, Ax\rangle &= \langle A^*x, A^*x\rangle,\\ \langle x, A^*Ax\rangle &= \langle x, AA^*x\rangle,\\ 0 &= \langle x, [A, A^*]x\rangle,\end{aligned}$$

jeweils für alle $x \in \mathbb{C}^n$. Da $[A, A^*]^* = [A, A^*]$ gilt, ist $[A, A^*]$ normal und daher diagonalisierbar. Somit folgt aus der letzten Zeile $[A, A^*] = 0$, also (i).

(i) $\Longleftrightarrow$ (v) folgt direkt aus

$$[(A + A^*), (A - A^*)] = 2[A^*, A].$$

Von dieser Gleichheit überzeugt man sich durch eine einfache direkte Rechnung. □

Definition 2.10.10.

(i) Eine Matrix $A = \left(a^i_j\right)_{1\le i\le j} \in \mathbb{C}^{n\times n}$ heißt positiv semidefinit, falls sie hermitesch ist und
$$\langle x, Ax\rangle \ge 0 \quad \text{für alle } x \in \mathbb{C}^n$$
gilt. Schreibweise: $A \ge 0$, $\left(a^i_j\right) \ge 0$ oder $A \succcurlyeq 0$, $\left(a^i_j\right) \succcurlyeq 0$, auch $a^i_j \ge 0$ oder $a^i_j \succcurlyeq 0$, falls dies nicht mit der Positivität der Einträge verwechselt werden kann.

(ii) Gilt zusätzlich $\langle x, Ax\rangle > 0$ für alle $x \in \mathbb{C}^n \setminus \{0\}$, so heißt A positiv definit. Wir schreiben $A > 0$, $\left(a^i_j\right) > 0$ oder $A \succ 0$, $\left(a^i_j\right) \succ 0$ bzw. $a^i_j > 0$ oder $a^i_j \succ 0$.

(iii) Sei V ein Skalarproduktraum. Dann heißt ein Endomorphismus f (oder eine darstellende Matrix) positiv semidefinit (bezüglich dieses Skalarproduktes), falls $\langle x, f(x)\rangle \ge 0$ für alle $x \in V$ gilt. Wir schreiben $f \succcurlyeq 0$. Gilt $\langle x, f(x)\rangle > 0$ für alle $x \in V \setminus \{0\}$, so heißt f strikt positiv definit, $f \succ 0$.

Lemma 2.10.11. *Sei $A \in \mathbb{C}^{n\times n}$ positiv semidefinit. Dann gibt es eine eindeutig bestimmte positiv semidefinite Quadratwurzel, d. h. eine positiv semidefinite Matrix $\sqrt{A} \in \mathbb{C}^{n\times n}$, so dass $\sqrt{A}\sqrt{A} = A$ gilt.*

Beweis. Da A normal ist, gibt es eine unitäre Matrix U, so dass U^*AU diagonal ist. Die Diagonaleinträge von U^*AU sind nicht negativ. Somit gibt es eine (reelle) Diagonalmatrix B mit nichtnegativen Diagonaleinträgen und $BB = U^*AU$. Die gesuchte Matrix ist $\sqrt{A} = UBU^*$. Da B eine Diagonalmatrix mit nichtnegativen Einträgen ist, sieht man direkt, dass $\sqrt{A}$ positiv semidefinit ist. $\sqrt{A}\sqrt{A} = A$ ist einfach nachzurechnen.

Sei C eine weitere Quadratwurzel von A. Da C hermitesch ist, gibt es eine Orthonormalbasis $\{a_i\}$ aus Eigenvektoren von C zu Eigenwerten λ_i. Wir erhalten

$$Aa_i = CCa_i = \lambda_i^2 a_i.$$

Somit sind die Werte λ_i^2 als Eigenwerte von A eindeutig bestimmt. Aufgrund der positiven Semidefinitheit sind auch die Werte $\lambda_i \ge 0$ eindeutig bestimmt. Die obigen Gleichungen liefern, dass ein Eigenraum von A zum Eigenwert λ_i^2 ein Eigenraum von C zum Eigenwert λ_i ist. Somit ist C eindeutig bestimmt. □

Das folgende Theorem verallgemeinert die Darstellung einer komplexen Zahl in der Form $re^{i\varphi}$ auf Matrizen.

Theorem 2.10.12 (Polarzerlegung)**.** *Sei $A \in \mathbb{C}^{n\times n}$. Dann gibt es eine unitäre Matrix U und eine positiv semidefinite Matrix S, so dass $A = US$ gilt.*

S ist eindeutig bestimmt. Ist A invertierbar, so ist die Zerlegung eindeutig.

Beweis. Die Matrix A^*A ist hermitesch und positiv semidefinit. Setze $S := \sqrt{A^*A}$. Gibt es solch eine Zerlegung, so ist S eindeutig bestimmt, denn es gilt

$$A^*A = S^*U^*US = S^*S = S^2$$

und die Wurzel ist eindeutig bestimmt.

Sei $A = \tilde{U}\tilde{S}$ eine weitere solche Zerlegung. Es gilt $\tilde{S} = S$. Ist A invertierbar, so ist auch S invertierbar und aus $US = \tilde{U}S$ folgt $U = \tilde{U}$.

Sei $(a_1, \dots, a_s)$ eine Orthonormalbasis aus Eigenvektoren von S: $Sa_i = \sqrt{\lambda_i}a_i$ mit s so dass $\lambda_1, \dots, \lambda_s > 0$ und $\lambda_{s+1}, \dots, \lambda_n = 0$ gelten. Es folgt

$$\langle Aa_i, Aa_j\rangle = \langle a_i, A^*Aa_j\rangle = \langle a_i, S^2a_j\rangle = \langle a_i, \lambda_j a_j\rangle = \lambda_j\delta_{ij},$$

da $\lambda_i \in \mathbb{R}$ für $1 \le i \le n$ gilt. Somit sind die Vektoren

$$\left(\frac{Aa_1}{\sqrt{\lambda_1}}, \dots, \frac{Aa_s}{\sqrt{\lambda_s}}\right) =: (w_1, \dots, w_s)$$

orthonormiert. Ergänze $(w_1, \dots, w_s)$ zu einer Orthonormalbasis $(w_1, \dots, w_n)$ von $V = \mathbb{C}^n$. Sei U die unitäre Matrix mit $Ua_i = w_i$ für $1 \le i \le n$, d. h. es gilt $U = (w_1, \dots, w_n)(a_1, \dots, a_n)^*$, wobei wir hintereinander geschriebene Vektoren in $\mathbb{C}^n$ als Matrizen auffassen. Dann gilt für $i = 1, \dots, s$

$$USa_i = U\sqrt{\lambda_i}a_i = \sqrt{\lambda_i}w_i = Aa_i.$$

Sei nun $i \in \{s+1, \dots, n\}$. Dann ist $a_i \in \ker S = \ker A^*A$. Wir behaupten, dass bereits $Aa_i = 0$ gilt. Es gelten $Aa_i \in \ker A^*$ und $Aa_i \in \operatorname{im} A$. Nach Theorem 2.7.8 gilt $\ker A^* \perp \operatorname{im} A$. Somit ist $Aa_i = 0$. Es folgt $USa_i = 0 = Aa_i$. Da $A = US$ für beliebige Basiselemente a_i, $1 \le i \le n$, gilt, folgt die Behauptung. □

Beachte, dass U auf dem Erzeugnis der Eigenräume von S zu positiven Eigenwerten stets eindeutig bestimmt ist.

Die folgende Zerlegung wird häufig ebenfalls als Polarzerlegung bezeichnet.

Korollar 2.10.13. *Sei $A \in \mathbb{C}^{n\times n}$. Dann gibt es eine unitäre Matrix U und eine positiv semidefinite Matrix S, so dass $A = SU$ gilt. S ist eindeutig bestimmt. Ist A invertierbar, so ist die Zerlegung eindeutig.*

Beweis. Die Polarzerlegung $A^* = \tilde{U}\tilde{S}$ liefert eine Zerlegung $A = \tilde{S}^*\tilde{U}^*$, also von der angegebenen Form. Die Eindeutigkeit von S folgt aus $A^*A = SU^*US = S^2$. Die Eindeutigkeit von U folgt wie bei der Polarzerlegung. □

Lemma 2.10.14. *Sei $A \in \mathbb{C}^{n\times n}$. Dann sind die folgenden Aussagen äquivalent:*

(i) A ist normal.
(ii) In der Polarzerlegung $A = US$ kommutieren S und U.
(iii) In der Zerlegung $A = SU$ kommutieren S und U.
(iv) Es gibt eine unitäre Matrix W mit $A^* = AW$.

Beweis.
„(i)$\Longrightarrow$ (ii)“: Sei W unitär, so dass WAW^* diagonal ist,

$$WAW^* = \operatorname{diag}\left\{r_1e^{i\varphi_1}, \dots, r_ke^{i\varphi_k}, 0, \dots, 0\right\}$$

mit $r_i > 0$, $0 \le \varphi_i < 2\pi$ für $1 \le i \le k$ und ein $k \in \mathbb{N}$. Die zugehörige Polarzerlegung hat dann die Gestalt $WAW^* = \tilde{S}\tilde{U}$, wobei

$$\tilde{S} = \operatorname{diag}\{r_1, \dots, r_k, 0, \dots 0\}$$

und

$$\tilde{U} = \operatorname{diag}\left\{e^{i\varphi_1}, \dots, e^{i\varphi_k}, U_0\right\}$$

sind und U_0 eine unitäre Matrix bezeichnet, der letzte Eintrag also als Block zu verstehen ist. Dann gilt $A = \left(W^*\tilde{S}W\right)\left(W^*\tilde{U}W\right)$ und die beiden Ausdrücke in Klammern entsprechen gerade der Polarzerlegung. Man sieht direkt, dass diese Matrizen vertauschen.

„(ii)$\Longrightarrow$ (i)“: Sei $A = SU = US$. Dann folgen $A^* = S^*U^* = SU^*$, $A^* = U^*S^* = U^*S$ und

$$A^*A - AA^* = SU^*US - SUU^*S = S^2 - S^2 = 0.$$

„(ii)$\Longleftrightarrow$ (iii)“: Ist offensichtlich wenn A invertierbar und U daher eindeutig bestimmt ist.

Sonst seien $A = SU = SV$ zwei Polarzerlegungen. Aus dem Beweis der Existenz der Polarzerlegung geht hervor, dass U und V auf den Eigenräumen von S zu positiven λ_i's übereinstimmen müssen und sich nur um eine unitäre Transformation K von $\ker S \subset \mathbb{C}^n$ unterscheiden können. K sei auf $\mathbb{C}^n$ durch die Identität/Einheitsmatrix fortgesetzt. Die Matrix K entspricht den unterschiedlichen Basisergänzungen im Beweis. Es gilt also $V = KU$. Da K sich nur auf $\ker S$ von der Identität unterscheidet und $K(\ker S) \subset \ker S$ gilt, folgt $KS = SK$. Gelte also $SU = US$ und $SU = SV$. Aus $SV = SKU = KSU = KUS = VS$ und derselben Argumentation im Fall VS, angewandt auf $(US)^* = (SU)^* = (VS)^*$, erhalten wir die Behauptung.

„(iii)$\Longrightarrow$ (iv)“: Schreibe $A = SU$ mit $[U, S] = 0$. Wir erhalten $A = US$ und daraus $A^* = S^*U^* = SU^* = SUU^*U^* = AU^*U^*$. U^* ist unitär, da U unitär ist: $U^* = U^{-1}$ liefert $U^{**} = (U^*)^{-1}$.

„(iv)$\Longrightarrow$ (i)“: Aus $A^* = AU$ folgt einerseits

$$A = AUU^* = A^*U^*$$

und andererseits

$$A = (A^*)^* = (AU)^* = U^*A^*,$$

somit ergeben sich also

$$A^*U^* = U^*A^*$$

und

$$UA = AU.$$

Zusammengenommen erhalten wir

$$A^*A - AA^* = (AU)A - A(AU) = AAU - AAU = 0.$$

Somit ist A normal. □

3. Bilinearformen

3.1. **Bilinearformen, Quadratische Formen.**

Definition 3.1.1 (Bilinearform)**.**

(i) Seien $V_1, \ldots, V_n, W$ Vektorräume über F. Eine Abbildung

$$f\colon V_1 \times \ldots \times V_n \to W$$

heißt multilinear, falls für jedes $i \in \{1, \ldots, n\}$ und festes

$$(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n) \in V_1 \times \ldots \times V_{i-1} \times V_{i+1} \times \ldots \times V_n$$

die Abbildung

$$V_i \ni v_i \mapsto f(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n)$$

linear ist.

(ii) Sei V ein F-Vektorraum. Eine multilineare Abbildung $\varphi\colon V \times V \to F$ heißt Bilinearform auf V. φ heißt symmetrisch, falls $\varphi(a, b) = \varphi(b, a)$ für alle $a, b \in V$ gilt.

Beispiele 3.1.2.

(i) Reelle Skalarprodukte sind symmetrische Bilinearformen.
(ii) $\varphi\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ mit $\varphi(\xi, \eta) := \xi^1\eta^1 - \xi^2\eta^2$ ist eine symmetrische Bilinearform auf $\mathbb{R}^2$, aber kein Skalarprodukt.
(iii) Sei V ein euklidischer Vektorraum und $f \in \mathrm{End}(V)$. Dann definiert $\varphi(\xi, \eta) := \langle \xi, f(\eta)\rangle$ für $\xi, \eta \in V$ eine Bilinearform. Ist f selbstadjungiert, so ist φ symmetrisch.
(iv) Sei $A \in F^{n\times n}$. Dann ist $F^n \times F^n \ni (x, y) \mapsto x^T Ay$ eine Bilinearform.
(v) Sei $V = C^0([a, b])$ und $k \in C^0\left([a, b]^2\right)$. Dann ist $\varphi\colon V \times V \to \mathbb{R}$, definiert durch

$$\varphi(f, g) := \int_a^b \int_a^b k(s, t) \cdot f(s) \cdot g(t)\, ds\, dt$$

eine Bilinearform auf V. φ ist genau dann symmetrisch, wenn $k(s, t) = k(t, s)$ für alle $a \le s, t \le b$ gilt.

Bemerkung 3.1.3.

(i) Seien $\varphi, \psi\colon V_1 \times \ldots \times V_n \to W$ zwei multilineare Abbildungen. Der Raum aller multilinearen Abbildungen ist ein Vektorraum, $L^n(V_1 \times \ldots \times V_n, W)$, wobei wir

$$(\varphi + \psi)(v_1, \ldots, v_n) := \varphi(v_1, \ldots, v_n) + \psi(v_1, \ldots, v_n)$$

und

$$(\alpha \cdot \varphi)(v_1, \ldots, v_n) := \alpha \cdot \varphi(v_1, \ldots, v_n)$$

setzen.
(ii) Sei $(a_1, \ldots, a_n)$ eine Basis von V. Sei φ eine Bilinearform auf V. Dann folgt für

$$\xi = \sum_{i=1}^n \xi^i a_i \quad \text{und} \quad \eta = \sum_{j=1}^n \eta^j a_j$$

wie bei linearen Abbildungen

$$\varphi(\xi, \eta) = \sum_{i,j=1}^n \xi^i \eta^j \varphi(a_i, a_j).$$

Daher ist φ durch $(c_{ij})_{1\le i,j\le n}$ mit $c_{ij} := \varphi(a_i, a_j)$ eindeutig bestimmt.
(iii) Umgekehrt definiert eine Matrix (c_{ij}) unter Verwendung der obigen Formel mit c_{ij} statt $\varphi(a_i, a_j)$ eine eindeutig bestimmte Bilinearform.
(iv) Wie bei linearen Abbildungen ist die Zuordnung $\varphi \mapsto (c_{ij})$ bei fixierter Basis $(a_1, \ldots, a_n)$ ein Vektorraumisomorphismus zwischen dem Vektorraum der Bilinearformen auf V und den $(n \times n)$-Matrizen über F. Insbesondere hat der Vektorraum der Bilinearformen auf einem n-dimensionalen Vektorraum V die Dimension n^2.
(v) Eine Bilinearform φ ist genau dann symmetrisch, wenn die zugehörige Matrix symmetrisch ist.

Bemerkung 3.1.4 (Transformationsverhalten bilinearer Abbildungen)**.**
Seien $(a_1, \ldots, a_n)$ und $(b_1, \ldots, b_n)$ Basen eines Vektorraumes V. Sei φ eine Bilinearform auf V. Sei $A = (\alpha_{ij})$ die zu φ bezüglich der Basis aus den a_i's gehörige Matrix und $B = (\beta_{ij})$ die zu den b_i's. Gelte

$$b_k = \sum_{i=1}^n d_k^i a_i$$

für $1 \leq k \leq n$. Setze $D := (d_j^i)$. Dann folgt

$$\beta_{kl} = \varphi(b_k, b_l) = \varphi\left(\sum_{i=1}^n d_k^i a_i, \sum_{j=1}^n d_l^j a_j\right) = \sum_{i,j=1}^n d_k^i d_l^j \varphi(a_i, a_j) = \sum_{i,j=1}^n d_k^i d_l^j \alpha_{ij}.$$

In Matrixform gilt also $B = D^T A D$. Dies ist dasselbe Transformationsverhalten wie beim reellen Skalarprodukt.

Definition 3.1.5 (Quadratische Form)**.** Sei $\varphi\colon V \times V \to F$ eine Bilinearform. Dann heißt $Q\colon V \to F$, definiert durch

$$Q(v) := \varphi(v, v) \quad \text{für } v \in V,$$

die zu φ zugehörige quadratische Form.

Beispiele 3.1.6.

(i) Sei φ die zur Matrix $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ gehörige Bilinearform. Dann gilt für die zugehörige quadratische Form Q die Identität $Q(v) = 0$ für beliebige $v \in V$.

(ii) Sei $\varphi = \langle \cdot, \cdot \rangle$ ein reelles Skalarprodukt. Dann gilt $Q = \|\cdot\|^2$ für die zugehörige quadratische Form Q.

Theorem 3.1.7. *Sei F ein Körper mit Charakteristik* $\operatorname{char} F \neq 2$. *Sei $Q\colon V \to F$ die zur Bilinearform $\varphi\colon V \times V \to F$ gehörige quadratische Form. Definiere $\psi\colon V \times V \to F$ durch*

$$\psi(\xi, \eta) := \tfrac{1}{2}(Q(\xi + \eta) - Q(\xi) - Q(\eta)) \quad \textit{für alle } \xi, \eta \in V.$$

(i) Es gilt

$$\psi(\xi, \eta) = \tfrac{1}{2}(\varphi(\xi, \eta) + \varphi(\eta, \xi)).$$

(ii) ψ ist eine weitere Bilinearform, zu der die quadratische Form Q gehört.

(iii) Ist φ symmetrisch, so gilt $\varphi = \psi$.

(iv) Ist φ bezüglich einer Basis die Matrix (a_{ij}) zugeordnet, so ist ψ bezüglich dieser Basis die Matrix $D = (d_{ij}) = \left(\frac{1}{2}(a_{ij} + a_{ji})\right)_{1 \leq i,j \leq n}$ zugeordnet.

Beweis. Wir zeigen nur die ersten beiden Teile.

(i) Es gilt

$$\begin{aligned} Q(\xi + \eta) &= \varphi(\xi + \eta, \xi + \eta) = \varphi(\xi, \xi) + \varphi(\xi, \eta) + \varphi(\eta, \xi) + \varphi(\eta, \eta) \\ &= Q(\xi) + \varphi(\xi, \eta) + \varphi(\eta, \xi) + Q(\eta). \end{aligned}$$

Wir erhalten daraus

$$\psi(\xi, \eta) = \tfrac{1}{2}(\varphi(\xi, \eta) + \varphi(\eta, \xi)).$$

(ii) Dies folgt aus

$$\psi(\xi, \xi) = \tfrac{1}{2}(\varphi(\xi, \xi) + \varphi(\xi, \xi)) = \varphi(\xi, \xi) = Q(\xi). \qquad \square$$

Korollar 3.1.8. *In einem reellen Skalarproduktraum V gilt die Polarisationsformel*

$$\langle x, y \rangle = \tfrac{1}{2}\left(\|x + y\|^2 - \|x\|^2 - \|y\|^2\right) \quad \textit{für alle } x, y \in V.$$

Wir wollen noch Bilinearformen über beliebigen Körpern durch Basiswechsel diagonalisieren. Dazu benutzen wir [7].

Definition 3.1.9. Sei V ein endlichdimensionaler F-Vektorraum und $B\colon V \times V \to F$ eine symmetrische Bilinearform. Dann heißt $(a_1, \ldots, a_n)$ eine orthogonale Basis von V bezüglich B, falls $(a_1, \ldots, a_n)$ eine Basis von V ist und

$$B(a_i, a_j) = 0 \quad \text{für alle } 1 \leq i < j \leq n$$

gilt.

Theorem 3.1.10. *Sei F ein Körper mit* $\operatorname{char} F \neq 2$. *Sei V ein endlichdimensionaler F-Vektorraum und $B\colon V \times V \to F$ eine symmetrische Bilinearform. Dann besitzt V eine orthogonale Basis bezüglich B.*

Beweis. Gilt $B \equiv 0$ oder $\dim V = 1$, so ist die Behauptung klar. Sonst finden wir ein $x \in V$ mit $B(x,x) \neq 0$. (Ist B bezüglich eines Basis $(b_1, \dots, b_n)$ die symmetrische Matrix $A = (a_{ij})_{1 \leq i,j \leq n}$ zugeordnet, so können wir $x = b_i$ wählen, falls $a_{ii} \neq 0$ gilt. Gilt $a_{ij} = a_{ji} \neq 0$ für feste i, j mit $i \neq j$ und $a_{kk} = 0$ für alle $1 \leq k \leq n$, so wählen wir $x = b_i + b_j$.) Wir ergänzen $x \neq 0$ zu einer Basis $(x, x_2, \dots, x_n)$ von V. Definiere $\tilde{x}_i := x_i - \frac{B(x,x_i)}{B(x,x)} x$ für $2 \leq i \leq n$. Dann gilt

$$B(x, \tilde{x}_i) = B(x, x_i) - \frac{B(x,x_i)}{B(x,x)} B(x,x) = 0$$

für alle $2 \leq i \leq n$. Nach Konstruktion ist $(x, \tilde{x}_2, \dots, \tilde{x}_n)$ ebenfalls eine Basis von V.

Die Behauptung folgt nun per Induktion nach der Dimension: Für den Vektorraum $W := \langle \tilde{x}_2, \dots, \tilde{x}_n \rangle$ finden wir nach Induktionsannahme eine bezüglich $B|_{W \times W}$ orthogonale Basis. Zusammen mit x erhalten wir die gesuchte orthogonale Basis von V. □

3.2. **Bilinearformen in euklidischen Vektorräumen.** Beliebige Bilinearformen lassen sich mit Hilfe des Skalarproduktes und mit linearen Abbildungen wie folgt darstellen:

Theorem 3.2.1. *Sei φ eine Bilinearform auf einem endlichdimensionalen reellen Vektorraum V. Sei $\langle \cdot, \cdot \rangle$ ein Skalarprodukt auf V, so gibt es eindeutig bestimmte lineare Abbildungen $f, g\colon V \to V$ mit*

$$\varphi(\xi, \eta) = \langle \xi, f(\eta) \rangle = \langle g(\xi), \eta \rangle$$

für alle $\xi, \eta \in V$. Die Abbildung f (und damit auch g) ist genau dann selbstadjungiert, wenn φ symmetrisch ist.

Beweis. Sei $\{a_1, \dots, a_n\}$ eine Orthonormalbasis von V. Definiere

$$c_{ik} := \varphi(a_i, a_k) \quad \text{für } 1 \leq i, k \leq n.$$

Definiere die linearen Abbildungen $f, g\colon V \to V$ durch

$$f(a_k) := \sum_{j=1}^{n} c_{jk} a_j = \sum_{j,l=1}^{n} c_{lk} \delta^{lj} a_j, \quad 1 \leq k \leq n,$$

$$g(a_i) := \sum_{j=1}^{n} c_{ij} a_j = \sum_{j,l=1}^{n} c_{il} \delta^{lj} a_j, \quad 1 \leq i \leq n.$$

Dann folgt, da die Vektoren a_i eine Orthonormalbasis bilden,

$$\langle a_i, f(a_k) \rangle = \left\langle a_i, \sum_{j=1}^{n} c_{jk} a_j \right\rangle = \sum_{j=1}^{n} c_{jk} \delta_{ij} = c_{ik} = \varphi(a_i, a_k),$$

$$\langle g(a_i), a_k \rangle = \left\langle \sum_{j=1}^{n} c_{ij} a_j, a_k \right\rangle = \sum_{j=1}^{n} c_{ij} \delta_{jk} = c_{ik} = \varphi(a_i, a_k).$$

Da beide Seiten in beiden Gleichungen bilinear sind, erhalten wir

$$\varphi(\xi, \eta) = \langle \xi, f(\eta) \rangle = \langle g(\xi), \eta \rangle$$

für alle $\xi, \eta \in V$.

Die Abbildung f ist eindeutig bestimmt; sei nämlich $\tilde{f}$ eine weitere solche Abbildung, so erhalten wir $\varphi(\xi, \eta) = \langle \xi, f(\eta) \rangle = \left\langle \xi, \tilde{f}(\eta) \right\rangle$ für alle $\xi, \eta \in V$. Weiterhin

folgt $0 = \left\langle \xi, f(\eta) - \tilde{f}(\eta) \right\rangle$ für alle $\xi, \eta \in V$ und damit $f(\eta) - \tilde{f}(\eta) = 0$ für alle $\eta \in V$. Also gilt $f = \tilde{f}$. Ebenso folgt die Eindeutigkeit für g.

Sei φ symmetrisch. Dann erhalten wir für $\xi, \eta \in V$

$$\langle \xi, f(\eta) \rangle = \varphi(\xi, \eta) = \varphi(\eta, \xi) = \langle \eta, f(\xi) \rangle = \langle f(\xi), \eta \rangle.$$

Somit ist f selbstadjungiert. Sei umgekehrt f selbstadjungiert. Es folgt

$$\varphi(\xi, \eta) = \langle \xi, f(\eta) \rangle = \langle f(\xi), \eta \rangle = \langle \eta, f(\xi) \rangle = \varphi(\eta, \xi).$$

Somit ist φ symmetrisch. □

Bemerkung 3.2.2.

(i) Den obigen Satz kann man umformulieren, indem man statt eines Skalarproduktes eine Orthogonalbasis vorgibt.

(ii) f und g hängen vom Skalarprodukt ab. Die Eigenschaft, selbstadjungiert zu sein oder nicht, ist jedoch davon unabhängig, da sie äquivalent zur Symmetrie von φ ist und um die Symmetrie von φ zu überprüfen, benötigt man kein Skalarprodukt.

(iii) Die f und die die Bilinearform bezüglich einer festen Orthonormalbasis darstellenden Matrizen stimmen überein. Zu g gehört die dazu transponierte Matrix.

Definition 3.2.3. Sei $\varphi\colon V \times V \to F$ eine symmetrische Bilinearform. Dann definieren wir den Nullraum von φ durch

$$N(\varphi) := \{\xi \in V\colon \varphi(\xi, \eta) = 0 \quad \text{für alle } \eta \in V\}.$$

Eine symmetrische Bilinearform heißt ausgeartet, falls $N(\varphi) \neq \{0\}$ gilt.

Bemerkung 3.2.4.

(i) $N(\varphi)$ ist ein Unterraum von V.

(ii) Da φ symmetrisch ist, gilt auch

$$N(\varphi) = \{\xi \in V\colon \varphi(\eta, \xi) = 0 \quad \text{für alle } \eta \in V\}.$$

(iii) Eine symmetrische Bilinearform φ ist genau dann nicht ausgeartet, wenn es zu jedem $v \in V$ mit $v \neq 0$ ein $w \in V$ mit $\varphi(v, w) \neq 0$ gibt. (Kleine Übung)

Beispiele 3.2.5.

(i) Fassen wir ein Skalarprodukt als Bilinearform auf, so gilt $N = \{0\}$.

(ii) Sei die bilineare Abbildung $\varphi\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ bezüglich der Standardbasis durch die Matrix

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

gegeben. Dann gilt $N(\varphi) = \langle \binom{1}{1} \rangle$.

Theorem 3.2.6. *Sei $\varphi\colon V \times V \to \mathbb{R}$ eine symmetrische Bilinearform auf einem euklidischen Vektorraum V. Gilt $\varphi(\xi, \eta) = \langle \xi, f(\eta) \rangle$ für alle $\xi, \eta \in V$, so folgt auch $N = N(\varphi) = \ker f$.*

Beweis. Sei $0 = \varphi(\xi, \eta)$ für alle $\eta \in V$. Dann folgt $0 = \varphi(\xi, \eta) = \varphi(\eta, \xi) = \langle \eta, f(\xi) \rangle$ auch $f(\xi) = 0$. Die umgekehrte Inklusion funktioniert analog. □

Der Kern einer einer symmetrischen Bilinearform zugeordneten linearen Abbildung f ist daher im Gegensatz zur Abbildung f selbst vom verwendeten Skalarprodukt unabhängig. Daher ist die folgende Definition gerechtfertigt

Definition 3.2.7. Sei φ eine symmetrische Bilinearform. Dann definieren wir den Rang von φ, $\operatorname{rang} \varphi$ als den Rang einer zugehörigen linearen Abbildung f (bezüglich eines beliebigen Skalarproduktes). Wir nennen φ regulär (bzw. singulär), wenn f regulär (bzw. singulär) ist.

Bemerkung 3.2.8.

(i) Es gilt $\operatorname{rang}\varphi = \dim V - \dim N(\varphi)$.
(ii) Ist φ durch C dargestellt, so gilt $\operatorname{rang}\varphi = \operatorname{rang} C$.

Beispiele 3.2.9.

(i) Die zu einem Skalarprodukt gehörige Bilinearform ist regulär.
(ii) Bezüglich der Standardbasis des $\mathbb{R}^3$ definiert die Matrix

$$C = \begin{pmatrix} 10 & 4 & 8 \\ 4 & -8 & -4 \\ 8 & -4 & 1 \end{pmatrix}$$

eine Bilinearform φ. Es gilt $\operatorname{rang} C = 2 = \operatorname{rang}\varphi$ sowie

$$N(\varphi) = \left\langle \begin{pmatrix} 2 \\ 3 \\ -4 \end{pmatrix} \right\rangle.$$

3.3. **Hauptachsentransformation.**

Sei V ein euklidischer Raum. Das Gram-Schmidtsche Orthogonalisierungsverfahren liefert eine Orthogonalbasis von V. Bezüglich dieser Basis ist das Skalarprodukt durch die Einheitsmatrix dargestellt.

Ziel dieses Abschnittes ist es zu zeigen, dass es zu jeder symmetrischen Bilinearform auf V eine Basis von V gibt, so dass die Bilinearform durch eine Diagonalmatrix dargestellt wird. Wir können sogar eine Orthogonalbasis wählen.

Theorem 3.3.1. *Sei $\varphi\colon V \times V \to \mathbb{R}$ eine symmetrische Bilinearform auf einem endlichdimensionalen euklidischen Vektorraum V. Sei $\dim V = n$. Dann existiert eine Orthonormalbasis $\{b_1, \ldots, b_n\}$ von V, bezüglich der φ durch eine Diagonalmatrix dargestellt wird. Die Diagonalelemente sind gerade die Eigenwerte λ_i der Matrix von φ (als Endomorphismus betrachtet) bezüglich einer beliebigen Orthonormalbasis von V.*

Beweis. Wir benutzen die lineare Abbildung $f\colon V \to V$, so dass $\varphi(\xi, \eta) = \langle \xi, f(\eta)\rangle$ für alle $\xi, \eta \in V$ gilt. Da φ symmetrisch ist, ist f selbstadjungiert. Somit gibt es eine Orthonormalbasis $\{b_1, \ldots, b_n\}$ von V aus Eigenvektoren von f zu Eigenwerten $(\lambda_i)_i$. Es gilt

$$\varphi(b_i, b_j) = \langle b_i, f(b_j)\rangle = \langle b_i, \lambda_j b_j\rangle = \lambda_j \delta_{ij}.$$

Somit ist φ bezüglich dieser Basis durch eine Diagonalmatrix C dargestellt und die Diagonaleinträge sind die Eigenwerte von f.

Bezüglich einer weiteren Orthogonalbasis ist φ durch $C' = E^T C E$ dargestellt, wobei E eine Basiswechselmatrix ist. Da beide Basen Orthonormalbasen sind, ist E orthogonal und es gilt $E^T = E^{-1}$. Wegen $C' = E^{-1} C E$ sind C und C' ähnlich und besitzen daher die gleichen Eigenwerte. □

Beispiel 3.3.2. Betrachte die symmetrische Bilinearform φ auf $\mathbb{R}^3$, die bezüglich der Standardbasis durch die Matrix

$$C = \begin{pmatrix} 4 & -5 & -2 \\ -5 & 4 & -2 \\ -2 & -2 & -8 \end{pmatrix}$$

gegeben ist. Sei $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ die durch C bezüglich der Standardbasis dargestellte lineare Abbildung. Die Eigenwerte von f sind die Nullstellen des zugehörigen charakteristischen Polynoms (etwas Rechenarbeit)

$$\chi_f(X) = \chi_C(X) = -(X+9)(X-9)X.$$

Bezüglich einer Orthonormalbasis $\{b_1, b_2, b_3\}$ aus Eigenvektoren ist f also durch

$$C' = \begin{pmatrix} 9 & 0 & 0 \\ 0 & -9 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

dargestellt. Orthonormale Eigenvektoren sind

$$b_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad b_2 = \frac{1}{3\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 4 \end{pmatrix} \quad \text{und} \quad b_3 = \frac{1}{3} \begin{pmatrix} 2 \\ 2 \\ -1 \end{pmatrix}.$$

In diesem Fall hat die Gleichung $C' = E^T C E$ die Gestalt

$$\begin{pmatrix} 9 & 0 & 0 \\ 0 & -9 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{3\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{4}{3\sqrt{2}} \\ \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \end{pmatrix} \begin{pmatrix} 4 & -5 & -2 \\ -5 & 4 & -2 \\ -2 & -2 & -8 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{2}{3} \\ -\frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{2}{3} \\ 0 & \frac{4}{3\sqrt{2}} & -\frac{1}{3} \end{pmatrix}.$$

Lemma 3.3.3. *Sei $\varphi\colon V \times V$ eine symmetrische Bilinearform auf einem endlichdimensionalen euklidischen Vektorraum. Dann gibt es eine Basis, so dass die darstellende Matrix die folgende Diagonalgestalt hat:*

$$\begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix}$$

Beweis. Benutze zunächst eine Basis $\{b_1, \ldots, b_n\}$ wie in Theorem 3.3.1. Nummeriere die Eigenwerte ohne Einschränkung so, dass $\lambda_1, \ldots, \lambda_s$ positiv, $\lambda_{s+1}, \ldots, \lambda_t$ negativ und $\lambda_{t+1}, \ldots, \lambda_n$ Null sind. Definiere die Vektoren

$$\begin{aligned} a_i &:= \frac{1}{\sqrt{\lambda_i}} b_i, && 1 \le i \le s, \\ a_i &:= \frac{1}{\sqrt{-\lambda_i}} b_i, && s+1 \le i \le t, \\ a_i &:= b_i, && t+1 \le i \le n. \end{aligned}$$

Die Menge $\{a_1, \ldots, a_n\}$ besteht aus paarweise zueinander orthogonalen Vektoren, von denen keiner der Nullvektor ist. Somit handelt es sich um eine Basis von V. Bezüglich dieser Basis hat φ aufgrund des Transformationsverhaltens von Bilinearformen darstellenden Matrizen die gewünschte Darstellung. □

Die Zahlen s und t hängen nicht von der speziellen Basiswahl ab.

Theorem 3.3.4 (Sylvesterscher Trägheitssatz)**.** *Sei Q eine quadratische Form in einem reellen endlich dimensionalen Vektorraum V. Dann gibt es eine Basis $\{a_1, \ldots, a_n\}$ von V, so dass Q in dieser Basis durch*

$$Q(\xi) = \sum_{i=1}^{s} \left(\xi^i\right)^2 - \sum_{i=s+1}^{t} \left(\xi^i\right)^2$$

mit $\xi = \sum_{i=1}^{n} \xi^i a_i$ gegeben ist. s und t hängen dabei nur von Q ab.

Beweis. Die Existenz einer solchen Darstellung haben wir gerade in Skalarproduktträumen gezeigt. Wir können hier ein beliebiges Skalarprodukt auf V benutzen um das letzte Resultat anzuwenden. Es bleibt also, die letzte Aussage zu zeigen. Der Nullraum einer Bilinearform ist von der Basis unabhängig. Somit hängt $t = \dim V - \dim N$ nur von Q und nicht von der Basis ab.

Sei $\alpha_1, \dots, \alpha_n$ eine weitere Basis von V zu der es σ mit

$$Q(\eta) = \sum_{i=1}^{\sigma} \left(\eta^i\right)^2 - \sum_{i=\sigma+1}^{t} \left(\eta^i\right)^2$$

für $\eta = \sum_{i=1}^{n} \eta^i \alpha_i$ gibt. Wir behaupten, dass $s = \sigma$ gilt. Setze

$$\begin{aligned} U &:= \langle a_{s+1}, \dots, a_t \rangle, \\ W &:= \langle \alpha_1, \dots, \alpha_\sigma, \alpha_{t+1}, \dots, \alpha_n \rangle. \end{aligned}$$

Es gilt

$$\begin{aligned} Q(\xi) &< 0 && \text{für } \xi \in U \setminus \{0\}, \\ Q(\eta) &\geq 0 && \text{für } \eta \in W. \end{aligned}$$

Somit ist $U \cap W = \{0\}$. Wir erhalten

$$n \geq \dim(U + W) = \dim U + \dim W = (t - s) + n - (t - \sigma) = n - s + \sigma.$$

Es folgt $s \geq \sigma$. Analog folgt $s \leq \sigma$. Daraus erhält man die Behauptung. □

Definition 3.3.5. In der Notation des Sylvesterschen Trägheitssatzes nennen wir eine quadratische Form Q bzw. die (nach Theorem 3.1.7) zugehörige symmetrische Bilinearform φ mit $Q(\xi) = \varphi(\xi, \xi)$

(i) positiv definit, wenn $s = t = n$ gilt. Alternativ definiert man Q als positiv definit, wenn $Q(\xi) > 0$ für alle $\xi \in V$ mit $\xi \neq 0$ gilt.
(ii) negativ definit, wenn $s = 0$ und $t = n$ gelten. Alternativ: Wenn $-Q$ positiv definit ist.
(iii) positiv semidefinit, wenn $s = t$ gilt. Alternativ: Wenn $Q(\xi) \geq 0$ für alle $\xi \in V$ gilt.
(iv) negativ semidefinit, wenn $s = 0$ gilt. Alternativ: Wenn $-Q$ positiv semidefinit ist.
(v) indefinit, wenn $0 < s < t$ gilt. Alternativ, wenn es $\xi, \eta \in V$ mit $Q(\xi) < 0 < Q(\eta)$ gibt.

Beispiele 3.3.6.

(i) Eine symmetrische Bilinearform φ ist genau dann ein Skalarprodukt, wenn φ positiv definit ist.
(ii) Betrachte die symmetrische Bilinearform auf $\mathbb{R}^4$ mit zugehöriger quadratischer Form

$$Q(\xi) = 2\xi^1\xi^2 + 2\xi^1\xi^3 + 2\xi^1\xi^4 + 2\xi^2\xi^3 + 2\xi^2\xi^4 + 2\xi^3\xi^4.$$

Die zugehörige Matrix ist

$$\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}.$$

Das zugehörige charakteristische Polynom lautet (eine Seite Rechnung)

$$\chi(X) = X^4 - 6X^2 - 8X - 3 = (X + 1)^3(X - 3).$$

Somit ist die Matrix der Bilinearform bezüglich einer wie im Sylvesterschen Trägheitssatz geeignet normierten Basis durch

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

gegeben. Q ist also indefinit.

Theorem 3.3.7. *Seien Q_1 und Q_2 zwei quadratische Formen auf einem reellen Vektorraum V. Sei Q_1 positiv definit. Dann gibt es eine Basis $a_1, \ldots, a_n$ von V und $c_1, \ldots, c_n \in \mathbb{R}$, so dass*

$$Q_1(\xi) = \sum_{i=1}^{n} \left(\xi^i\right)^2, \quad \text{und} \quad Q_2(\xi) = \sum_{i=1}^{n} c_i \left(\xi^i\right)^2$$

für $\xi = \sum_{i=1}^{n} \xi^i a_i$ gilt.

Beweis. Die zu Q_1 gehörige symmetrische Bilinearform ist positiv definit und daher ein Skalarprodukt auf V. Nach Theorem 3.3.1 gibt es daher eine bezüglich dieses Skalarproduktes orthonormale Basis $\{a_1, \ldots, a_n\}$, in der die symmetrische Bilinearform zu Q_2 durch eine Diagonalmatrix dargestellt wird. □

3.4. **Extremaleigenschaften der Eigenwerte.** Vermöge

$$\varphi(\xi, \xi) = \langle \xi, f(\xi) \rangle$$

übertragen sich die folgenden Ergebnisse auch auf symmetrische Bilinearformen.

Wir setzen $\mathbb{S}^{n-1} := \{x \in \mathbb{R}^n \colon \|x\| = 1\}$ bezüglich der vom Standardskalarprodukt auf $\mathbb{R}^n$ induzierten Norm.

Für das folgende Theorem benötigen wir einen kleinen Hilfssatz:

Lemma 3.4.1. *Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch mit Eigenwerten $\lambda_1 \leq \ldots \leq \lambda_n \in \mathbb{R}$. Sei U ein k-dimensionaler Unterraum von $\mathbb{R}^n$. Wir verwenden das Standardskalarprodukt auf $\mathbb{R}^n$. Dann gibt es ein $x \in U$ mit $\langle x, x \rangle = \|x\|^2 = 1$ und $\langle x, Ax \rangle \geq \lambda_k$.*

Beweis. Wähle eine Orthonormalbasis $\{a_i\}_{1 \leq i \leq n}$ aus Eigenvektoren von A mit $Aa_i = \lambda_i a_i$. Setze $W := \langle a_k, \ldots, a_n \rangle$. Aus Dimensionsgründen ist $U \cap W \neq \{0\}$. Sei $x \in U \cap W$ mit $\|x\| = 1$. Aus $x \in W$ mit $x = \sum_{i=k}^{n} x^i a_i$ folgt

$$\langle Ax, x \rangle = \left\langle \sum_{i=k}^{n} x^i \underbrace{Aa_i}_{=\lambda_i a_i}, \sum_{j=k}^{n} x^j a_j \right\rangle = \sum_{i=k}^{n} \left(x^i\right)^2 \underbrace{\lambda_i}_{\geq \lambda_k} \geq \lambda_k \cdot \sum_{i=k}^{n} \left(x^i\right)^2 = \lambda_k.$$

Somit ergibt sich die Behauptung. □

Theorem 3.4.2. *Sei f ein selbstadjungierter Endomorphismus auf $\mathbb{R}^n$. Sei $\mathcal{U}_k$ die Menge aller k-dimensionalen Unterräume von $\mathbb{R}^n$. Dann gilt für die nach Größe geordneten und entsprechend ihrer Vielfachheit aufgeführten Eigenwerte $\lambda_1 \leq \ldots \leq \lambda_n$*

$$\lambda_k = \inf_{U \in \mathcal{U}_k} \sup_{x \in \mathbb{S}^{n-1} \cap U} \langle x, f(x) \rangle.$$

Beweis. Das Supremum wird stets angenommen, da $\mathbb{S}^{n-1} \cap U$ kompakt (= beschränkt und abgeschlossen) ist und da $x \mapsto \langle x, f(x) \rangle$ stetig ist.

Sei $\{a_i\}_{1\leq i\leq n}$ eine Orthogonalbasis von $\mathbb{R}^n$ mit $f(a_i) = \lambda_i a_i$. Setze $U_1 := \langle a_1, \ldots, a_k\rangle$. Dann gilt für $x = \sum_{i=1}^{k} x^i a_i \in U_1$ mit $\|x\| = 1$, also $x \in \mathbb{S}^{n-1}$

$$\langle x, f(x)\rangle = \sum_{i,j=1}^{k} x^i x^j \langle a_i, f(a_j)\rangle = \sum_{i,j=1}^{k} x^i x^j \lambda_j \underbrace{\langle a_i, a_j\rangle}_{=\delta_{ij}} = \sum_{i=1}^{k} \left(x^i\right)^2 \lambda_i$$

$$\leq \lambda_k \sum_{i=1}^{k} \left(x^i\right)^2 = \lambda_k \|x\|^2 = \lambda_k.$$

Insbesondere folgt also

$$\langle a_k, f(a_k)\rangle = \lambda_k = \sup_{x\in\mathbb{S}^{n-1}\cap U_1} \langle x, f(x)\rangle.$$

Lemma 3.4.1 zeigt, dass

$$\inf_{U\in\mathcal{U}_k} \sup_{x\in\mathbb{S}^{n-1}\cap U} \langle x, f(x)\rangle \geq \lambda_k = \sup_{x\in\mathbb{S}^{n-1}\cap U_1} \langle x, f(x)\rangle$$

gilt. Somit folgt Gleichheit in der obigen Ungleichung und damit die Behauptung. □

Bemerkung 3.4.3. Sei U ein Unterraum von $\mathbb{R}^n$. Sei f ein symmetrischer Endomorphismus von $\mathbb{R}^n$ mit Standardskalarprodukt. Dann gilt

$$\sup_{x\in\mathbb{S}^{n-1}\cap U} \langle x, f(x)\rangle = \sup_{\substack{x\in U\\ x\neq 0}} \frac{\langle x, f(x)\rangle}{\|x\|^2}.$$

Dieser letzte Quotient heißt Rayleigh-Quotient. Wir hätten das obige Theorem auch mit Hilfe des Rayleigh-Quotienten formulieren können.

Wir wollen noch eine weitere Möglichkeit kennen lernen, Eigenwerte als Extrema wiederzufinden. Das folgende Theorem kann auch aus Theorem 3.4.2 gefolgert werden. Wir geben trotzdem einen unabhängigen Beweis.

Theorem 3.4.4. *Sei $f\colon \mathbb{R}^n \to \mathbb{R}^n$ bezüglich des Standardskalarproduktes selbstadjungiert. Seien $\lambda_1 \leq \ldots \leq \lambda_n$ die Eigenwerte von f. Dann gilt*

$$\lambda_n = \sup_{x\in\mathbb{S}^{n-1}} \langle x, f(x)\rangle = \sup_{\substack{x\in\mathbb{R}^n\\ x\neq 0}} \frac{\langle x, f(x)\rangle}{\|x\|^2}.$$

Beweis. Wir zeigen nur die erste Gleichheit. Die zweite Gleichheit folgt aus Homogenitätsgründen.

Aufgrund der Kompaktheit von $\mathbb{S}^{n-1}$ wird das Supremum in einem Punkt $x_0 \in \mathbb{S}^{n-1}$ angenommen. Sei $\varepsilon > 0$ und $\gamma\colon (-\varepsilon, \varepsilon) \to \mathbb{S}^{n-1}$ eine C^1-Kurve mit $\gamma(0) = x_0$. Dann nimmt die Funktion $t \mapsto \langle \gamma(t), f(\gamma(t))\rangle$ in $t = 0$ ein Maximum an. Somit gilt $0 = \langle \gamma'(0), f(x_0)\rangle + \langle x_0, f(\gamma'(0))\rangle$. Sei $x_1 \in \mathbb{S}^{n-1}$ mit $\langle x_0, x_1\rangle = 0$. Dann ist $\gamma(t) := \cos t \cdot x_0 + \sin t \cdot x_1$ eine Kurve in $\mathbb{S}^{n-1}$ wie oben betrachtet mit $\gamma'(0) = x_1$. Da f selbstadjungiert ist erhalten wir

$$0 = \langle x_1, f(x_0)\rangle + \langle x_0, f(x_1)\rangle = 2\langle x_1, f(x_0)\rangle.$$

Somit ist $f(x_0)$ ein Vielfaches von x_0 und daher ist x_0 ein Eigenvektor. Da für jeden Eigenvektor $y \in \mathbb{S}^{n-1}$ von f zum Eigenwert λ

$$\lambda = \langle y, f(y)\rangle$$

gilt, liefert die obige Formel auch den größten Eigenwert und das Maximum wird an einem zugehörigen Eigenvektor angenommen. □

Auch diese Methode läßt sich auf die anderen Eigenwerte verallgemeinern.

Theorem 3.4.5. *Sei $f\colon \mathbb{R}^n \to \mathbb{R}^n$ bezüglich des Standardskalarproduktes selbstadjungiert. Seien $\lambda_1 \leq \ldots \leq \lambda_n$ die Eigenwerte von f. Seien $a_{k+1}, \ldots, a_n$ Eigenvektoren zu den Eigenwerten $\lambda_{k+1}, \ldots, \lambda_n$. Setze $U := \langle a_{k+1}, \ldots, a_n \rangle$. Dann gilt*

$$\lambda_k = \sup_{x \in \mathbb{S}^{n-1} \cap U^\perp} \langle x, f(x) \rangle = \sup_{\substack{x \in U^\perp \\ x \neq 0}} \frac{\langle x, f(x) \rangle}{\|x\|^2}.$$

Beweis. Werde das Supremum in $a_k \in \mathbb{S}^{n-1} \cap U^\perp$ angenommen. Sei $b \in U^\perp$ mit $\langle b, a_k \rangle = 0$ und $\|b\| = 1$ beliebig. Betrachte eine Kurve wie oben, hier $\mathbb{R} \ni t \mapsto \cos t \cdot a_k + \sin t \cdot b$. Wiederum folgt $0 = \langle b, f(a_k) \rangle$ für alle Vektoren b wie angegeben. Sei $v \in U$ ein beliebiger Eigenvektor zum Eigenwert λ. Dann gilt $\langle v, f(a_k) \rangle = \langle f(v), a_k \rangle = \lambda \langle v, a_k \rangle = 0$. Zusammengenommen folgt $\langle w, f(a_k) \rangle = 0$ für alle $w \in \mathbb{S}^{n-1} \cap \langle a_k \rangle^\perp$. Daher ist a_k ein Eigenvektor von f. Nach Konstruktion kann a_k nicht im Erzeugnis der Eigenvektoren $a_{k+1}, \ldots, a_n$ liegen. Für jeden Eigenvektor $x \in \mathbb{S}^{n-1}$ liefert $\langle x, f(x) \rangle$ gerade den zugehörigen Eigenwert. Somit folgt die obige Formel. □

Theorem 3.4.6. *Sei $a > 0$. Sei $(-a, a) \ni t \mapsto A(t) \in R^{n \times n}$ stetig und sei $A(t)$ für alle $t \in (-a, a)$ symmetrisch. (Dabei identifizieren wir $\mathbb{R}^{n \times n}$ mit $\mathbb{R}^{n \cdot n}$ und verwenden auf $\mathbb{R}^{n \cdot n}$ eine beliebige Norm.) Seien $\lambda_i(t)$, $1 \leq i \leq n$, $t \in (-a, a)$ die angeordneten Eigenwerte von $A(t)$ mit Vielfachheit, gelte also $\lambda_1(t) \leq \lambda_2(t) \leq \ldots \leq \lambda_n(t)$ für alle $t \in (-a, a)$. Dann hängen die Eigenwerte $\lambda_i(t)$ für festes $i \in \{1, \ldots, n\}$ stetig von t ab.*

Beweis. Wir wollen Theorem 3.4.2 benutzen. Seien $t_0 \in (-a, a)$ und $\varepsilon > 0$ beliebig. Gelte $A(t) = \left(a^i_j(t)\right)_{1 \leq i,j \leq n}$. Dann gibt es $\delta > 0$, so dass $(t_0 - \delta, t_0 + \delta) \subset (-a, a)$ und $\left|a^i_j(t) - a^i_j(t_0)\right| \leq \frac{\varepsilon}{n^2}$ für alle $1 \leq i, j \leq n$ und $t \in (t_0 - \delta, t_0 + \delta)$ (da $t \mapsto A(t)$ stetig ist) gelten. Sei $x \in \mathbb{S}^{n-1}$ beliebig. Wegen $1 = \langle x, x \rangle = \sum_{i=1}^{n} \left(x^i\right)^2$ gilt $\left|x^i\right| \leq 1$ für alle $i \in \{1, \ldots, n\}$. Wir erhalten also

$$\begin{aligned}
|\langle x, A(t)x \rangle - \langle x, A(t_0)x \rangle| &= |\langle x, (A(t) - A(t_0))x \rangle| \\
&= \left| \sum_{i,j,k=1}^{n} x^i \delta_{ij} \left(a^j_k(t) - a^j_k(t_0)\right) x^k \right| \\
&\leq n^2 \underbrace{\left(\sup_{1 \leq i \leq n} \left|x^i\right| \right)^2}_{\leq 1} \frac{\varepsilon}{n^2} \leq \varepsilon.
\end{aligned}$$

Es folgt nach Theorem 3.4.2 (wobei $\mathcal{U}_k$ wieder die Menge aller k-dimensionalen Unterräume von $\mathbb{R}^n$ bezeichnet)

$$\begin{aligned}
\lambda_k(t) &= \inf_{U \in \mathcal{U}_k} \sup_{x \in \mathbb{S}^{n-1} \cap U} \langle x, A(t)x \rangle \\
&\leq \inf_{U \in \mathcal{U}_k} \sup_{x \in \mathbb{S}^{n-1} \cap U} \langle x, A(t_0)x \rangle + \varepsilon \\
&= \lambda_k(t_0) + \varepsilon.
\end{aligned}$$

Analog folgt die umgekehrte Ungleichung und daher die Stetigkeit. □

Das folgende Theorem gilt auch für hermitesche Matrizen, siehe [5, Theorem 7.2.5]. Für positiv semidefinite Matrizen gilt es nicht mit „≥ 0“. Wir schreiben $\mathbb{K}$ statt $\mathbb{R}$ oder $\mathbb{C}$.

Theorem 3.4.7. *Sei $A = \left(a^i_j\right)_{1 \leq i,j \leq n} \in \mathbb{K}^{n \times n}$ hermitesch. Dann sind die folgenden beiden Aussagen äquivalent:*

(i) A ist positiv definit,
(ii) alle Unterdeterminanten $\det\left(a^i_j\right)_{1\le i,j\le k} \equiv \det A_k$, $k = 1,\dots,n$, *erfüllen* $\det A_k > 0$ *(und sind im hermiteschen Fall reell).*

Beweis.
„(i) $\implies$ (ii)“: Sei A positiv definit. Betrachte $A_k = \left(a^i_k\right)_{1\le i,j\le k}$ für ein k mit $1 \le k \le n$. Dann ist auch A_k positiv definit. A_k ist diagonalisierbar und hat positive Eigenwerte. Somit ist auch das Produkt dieser Eigenwerte, $\det A_k$, positiv.
„(ii) $\implies$ (i)“: Per Induktion. Sei die Aussage bereits für $(n-1)\times(n-1)$-Matrizen gezeigt, A_{n-1} also positiv definit. Nach Theorem 3.4.2, mit Infimum und Supremum vertauscht um die größeren Eigenwerte zuerst zu bekommen, besitzt A mindestens $(n-1)$ positive Eigenwerte, da wir schon für die spezielle Wahl $U = \langle e_1,\dots,e_k\rangle$ mit $1 \le k \le n-1$ auch k positive Werte erhalten die dann untere Schranken für die Eigenwerte von A sind. Das Produkt der Eigenwerte ist aber ebenfalls positiv. Somit sind alle Eigenwerte positiv. □

3.5. **Flächen im $\mathbb{R}^3$.** Wir geben nur die wichtigste Definition an.

Definition 3.5.1. Sei $\varphi\colon \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ symmetrisch und nicht ausgeartet. Seien $\lambda_1 \le \lambda_2 \le \lambda_3$ die Eigenwerte der zugehörigen selbstadjungierten linearen Abbildung. Dann heißt die Menge

$$\left\{x \in \mathbb{R}^3 \colon \varphi(x,x) = 1\right\}$$

(i) Ellipsoid, falls $\lambda_i > 0$ für $i = 1,2,3$ gilt. Die Werte $\frac{1}{\sqrt{\lambda_i}}$ heißen Halbachsen. Ein Ellipsoid heißt genau dann Rotationsellipsoid, wenn $\lambda_1 = \lambda_2$ oder $\lambda_2 = \lambda_3$ gilt. Ein Ellipsoid heißt Sphäre, falls $\lambda_1 = \lambda_2 = \lambda_3$ gilt.
(ii) einschaliges Hyperboloid, falls $\lambda_1 < 0 < \lambda_2 \le \lambda_3$ gilt. Es ist genau dann um den Eigenraum zu λ_1 rotationssymmetrisch, falls $\lambda_2 = \lambda_3$ gilt.
(iii) zweischaliges Hyperboloid, falls $\lambda_1 \le \lambda_2 < 0 < \lambda_3$ gilt. Es ist genau dann um den Eigenraum zu λ_3 rotationssymmetrisch, falls $\lambda_1 = \lambda_2$ gilt.

4. Ringe

4.1. **Ringe.**

Definition 4.1.1 (Ring)**.** Ein Ring A ist eine additive abelsche Gruppe, d. h. wir schreiben die Verknüpfung als „+“, mit einer Multiplikation „$\cdot$“, so dass folgende Axiome für alle $a,b,c \in A$ gelten

(R1) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ (Assoziativität)
(R2) $a \cdot (b + c) = a \cdot b + a \cdot c$ (Distributivität)
(R3) $(a + b) \cdot c = a \cdot c + b \cdot c$ (Distributivität)

Hierbei verwenden wir eine Klammersetzung nach der Regel „Punkt vor Strich“. Wir schreiben später ab statt $a \cdot b$.

(R4) Ein Ring mit Einselement, also einem Element $1 = 1_A \in A$, das $a \cdot 1 = a = 1 \cdot a$ für alle $a \in A$ erfüllt, heißt Ring mit 1. Häufig wird die Existenz eines Einselementes bereits in der Definition eines Ringes verlangt.
(R5) Ein Ring heißt kommutativ, falls $a \cdot b = b \cdot a$ für alle $a,b \in A$ gilt.

Beispiele 4.1.2.
(i) $\mathbb{Z}$ ist ein kommutativer Ring mit Eins.
(ii) Die geraden Zahlen sind ein Ring ohne Eins.
(iii) Ein Körper K ist ein kommutativer Ring mit Eins.
(iv) $K^{n\times n}$ ist ein Ring. Für $n \ge 2$ ist er nicht kommutativ.

Definition 4.1.3 (Ringhomomorphismus)**.**

(i) Seien A, B Ringe. Dann heißt eine Abbildung $\varphi\colon A \to B$ Ringhomomorphismus, wenn $\varphi(a+b) = \varphi(a) + \varphi(b)$ und $\varphi(a \cdot b) = \varphi(a) \cdot \varphi(b)$ für alle $a, b \in A$ gelten. Bei Ringen mit Eins verlangt man zusätzlich $\varphi(1) = 1$ (dies folgt nicht; betrachte die Nullabbildung).

(ii) Der Deutlichkeit halber könnte man auch $\varphi(a +_A b) = \varphi(a) +_B \varphi(b)$ oder $\varphi(1_A) = 1_B$ schreiben.

(iii) Ist φ bijektiv, so heißt φ Ringisomorphismus oder Isomorphismus. Dann ist φ^{-1} ebenfalls ein Ringisomorphismus (Übung).

(iv) A und B heißen isomorph, falls es einen Ringisomorphismus $\varphi\colon A \to B$ gibt.

Bemerkung 4.1.4.

(i) Seien $\varphi\colon A \to B$ und $\psi\colon B \to C$ Ringhomomorphismen. Dann ist auch die Verknüpfung $\psi \circ \varphi\colon A \to C$ ein Ringhomomorphismus.

(ii) Sei A ein Ring, $B \subset A$ eine Teilmenge, die unter Addition und Multiplikation abgeschlossen ist. Dann ist die Inklusionsabbildung $B \to A$ ein Ringhomomorphismus.

(iii) Sei A ein Ring mit Eins. Dann ist $\varphi\colon \mathbb{Z} \to A$ mit $0 \mapsto 0$, $n \mapsto \underbrace{1 + \ldots + 1}_{n \text{ Stück}}$ sowie $-n \mapsto -\varphi(n)$ für $n > 0$ ein Ringhomomorphismus.

Für einen Ring A mit Eins gibt es genau einen Ringhomomorphismus $\varphi\colon \mathbb{Z} \to A$. Wir schreiben dementsprechend in diesem Falle auch $n = \sum_{i=1}^{n} 1$ und na für $n \in \mathbb{N}$ und $a \in A$.

(iv) Sei $K[t]$ der Polynomring über einem Körper. Dann ist $\varphi_a\colon K[t] \to K$ mit $\varphi_a(f) := f(a)$ für $f \in K[t]$ ein Ringhomomorphismus. Es gilt

$$\ker \varphi_a = \{(t-a)h\colon h \in K[t]\}.$$

Beweis: Polynomdivision.

(v) Seien $A_1, \ldots, A_n$ Ringe. Dann ist $A := A_1 \times \ldots \times A_n$ mit komponentenweise definierter Addition und Multiplikation ein Ring. Sind alle Ringe $A_1, \ldots, A_n$ kommutativ bzw. Ringe mit Eins, so ist auch A kommutativ bzw. ein Ring mit Eins $(1, \ldots, 1)$. Die Projektionsabbildungen

$$\pi_i\colon A \to A_i \quad \text{mit} \quad \pi_i(a_1, \ldots, a_n) = a_i$$

für $i = 1, \ldots, n$ sind Ringhomomorphismen.

Bemerkung 4.1.5 (Konvention)**.** Seien ab jetzt alle Ringe kommutativ mit Eins.

Wir wiederholen:

Definition 4.1.6. Sei A ein kommutativer Ring mit Eins. Eine Teilmenge $I \subset A$ heißt Ideal von A, falls I eine Untergruppe von A ist und für alle $a \in I$ und $b \in A$ auch $ab \in I$ gilt.

Wir sagen auch, dass $I \subset A$ ein Ideal sei.

Bemerkung 4.1.7. Ein Ideal $I \subset A$ in einem Ring (mit Eins) ist nur dann ein Teilring (mit Eins), falls $I = A$ gilt.

Die Angaben in Klammern hier und später sind als Erinnerungen zu verstehen. Ohne sie wird das jeweilige Resultat möglicherweise falsch.

Lemma 4.1.8. *Sei A ein (kommutativer) Ring und $I \subset A$ ein Ideal. Dann ist*

$$A/I := \{a + I\colon a \in A\}$$

mit $\bar{a} \equiv a + I := \{a + b\colon b \in I\}$ mit vertreterweise definierter Addition und Multiplikation ein Ring, der Quotientenring von A nach I.

Die Projektionsabbildung $\pi\colon A \to A/I$ mit $a \mapsto a + I$ ist ein (surjektiver) Ringhomomorphismus.

Beweis. Wir zeigen nur die Wohldefiniertheit von Addition und Multiplikation. Seien $a, b \in A$ und $c, d \in I$. Dann gilt

$$(a+c)+(b+d) = (a+b) + \underbrace{(c+d)}_{\in I}$$

sowie

$$(a+c)\cdot(b+d) = ab + \underbrace{ad+cb+cd}_{\in I}.$$

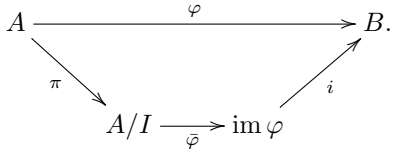
Wir bemerken, dass die Mengen $a+I$ für $a \in A$ die Äquivalenzklassen von A unter der Äquivalenzrelation $b \equiv c \pmod{I}$ sind, wobei $b \equiv c \pmod{I}$ genau dann gilt, wenn $b - c \in I$ ist. □

Theorem 4.1.9. *Sei A ein kommutativer Ring mit Eins. Dann ist $I \subset A$ genau dann ein Ideal, wenn es einen kommutativen Ring B mit Eins und einen Ringhomomorphismus $\varphi\colon A \to B$ mit $\ker\varphi = I$ gibt.*

Beweis.
„$\Longleftarrow$“: Seien $a \in \ker\varphi$ und $b \in A$. Dann ist $\varphi(ab) = \varphi(a)\cdot\varphi(b) = 0\cdot\varphi(b) = 0$. Somit ist $ab \in \ker\varphi$.

„$\Longrightarrow$“: Betrachte die Projektion $\pi\colon A \to A/I$. Dann gelten nach Definition $I \subset \ker\pi$ sowie $a + I \neq 0$ für $a \notin I$. □

Theorem 4.1.10 (Homomorphiesatz für Ringe)**.** *Sei $\varphi\colon A \to B$ ein Ringhomomorphismus. Sei $I \subset A$ ein Ideal. Sei $\pi\colon A \to A/I$ die Quotientenabbildung. Sei $i\colon \operatorname{im}\varphi \to B$ die Inklusionsabbildung. Dann gibt es genau dann einen Ringhomomorphismus $\bar\varphi\colon A/I \to \operatorname{im}\varphi \subset B$ mit $\varphi = i\circ\bar\varphi\circ\pi$, wenn $I \subset \ker\varphi$ gilt. Existiert $\bar\varphi$, so ist diese Abbildung eindeutig bestimmt und es gelten $\ker\bar\varphi = (\ker\varphi)/I$ sowie $i(\operatorname{im}\bar\varphi) = \operatorname{im}\varphi$. Das folgende Diagramm kommutiert*



Beweis.

(i) Sei $I \not\subset \ker\varphi$. Sei $a \in I \setminus \ker\varphi$. Nehme an, es gäbe eine solche Abbildung $\bar\varphi$. Dann folgte $0 \neq \varphi(a) = i\circ\bar\varphi\circ\pi(a) = i\circ\bar\varphi(0) = 0$. Widerspruch.

(ii) Definiere $\bar\varphi\colon A/I \to \operatorname{im}\varphi$ durch $a + I \mapsto \varphi(a)$. Wegen $I \subset \ker\varphi$ ist $\bar\varphi$ wohldefiniert. Da φ ein Ringhomomorphismus ist, gilt dies auch für $\bar\varphi$.

(iii) Die Kommutativität des Diagrammes bzw. $\varphi = i\circ\bar\varphi\circ\pi$ ist nach Definition klar.

(iv) Da $\bar\varphi$ wohldefiniert ist, folgt $\ker\bar\varphi = (\ker\varphi)/I$. $i(\operatorname{im}\bar\varphi) = \operatorname{im}\varphi$ ist nach Definition klar.

(v) Sei $\tilde\varphi\colon A/I \to \operatorname{im}\varphi$ eine weitere solche Abbildung. Da π surjektiv ist, folgt aus $i\circ\bar\varphi\circ\pi = i\circ\tilde\varphi\circ\pi$ die Gleichheit $i\circ\bar\varphi = i\circ\tilde\varphi$. Da schließlich i injektiv ist, folgt aus $i\circ\bar\varphi = i\circ\tilde\varphi$ bereits $\bar\varphi = \tilde\varphi$. Somit ist $\bar\varphi$ eindeutig bestimmt. □

Beispiele 4.1.11.

(i) Sei $n \in \mathbb{N} \equiv \{0, 1, 2, \ldots\}$. Dann ist $n\mathbb{Z}$ ein Ideal von $\mathbb{Z}$. Wir erhalten als Quotienten- oder Restklassenring $\mathbb{Z}/n\mathbb{Z} = \{\bar 0, \bar 1, \ldots, \overline{n-1}\}$. Wir werden sehen, dass es keine weiteren Ideale in $\mathbb{Z}$ gibt (Division mit Rest).

(ii) Ist K ein Körper den wir als Ring auffassen, so sind $\{0\}$ und K Ideale von K.
Wir werden noch sehen, dass ein Ring, in dem es nur diese beiden trivialen Ideale gibt, stets ein Körper ist (Faktorisiere ein maximales Ideal heraus).

(iii) Sei K ein Körper und $A = K[t]$. Sei $p \in K[t]$. Dann ist $(p) := \{p \cdot q \colon q \in K[t]\}$ ein Ideal in $K[t]$. Wir werden sehen, dass es keine weiteren Ideale in $K[t]$ gibt (Division mit Rest).

Bemerkung 4.1.12 (Konstruktion von Idealen)**.**

(i) Sei $X \subset A$ beliebig. Dann gibt es ein eindeutig bestimmtes kleinstes Ideal von A, das X enthält nämlich

$$(X) := \left\{ \sum_{i=1}^{n} a_i x_i \colon n \in \mathbb{N},\, x_i \in X,\, a_i \in A,\, i = 1, \ldots, n \right\}.$$

Jedes Ideal, welches X enthält muss nämlich auch diese Menge enthalten und der Ausdruck auf der rechten Seite ist ein Ideal.

Beachte, dass $(\emptyset) = \{0\}$ gilt.

Ist $X = \{x_1, \ldots, x_n\}$ endlich, so schreiben wir auch $(X) = (x_1, \ldots, x_n) = Ax_1 + \ldots + Ax_n$.

(ii) Ein Ideal $I \subset A$ heißt Hauptideal, falls es ein $a \in I$ mit $I = (a)$ gibt.

Für $n \in \mathbb{Z}$ und $p \in K[t]$ haben wir die Hauptideale $n\mathbb{Z}$ und (p) bereits gesehen.

(iii) Ist $(I_\lambda)_{\lambda \in \Lambda}$ eine Familie von Idealen in A, so ist auch $\bigcap\limits_{\lambda \in \Lambda} I_\lambda$ ein Ideal von A.

Dies folgt direkt nach der Definition eines Ideals.

(iv) Seien $I, J \subset A$ Ideale. Dann ist $I \cup J$ im allgemeinen kein Ideal mehr (z. B. $2\mathbb{Z} \cup 3\mathbb{Z}$). Wir definieren daher

$$I + J := (I \cup J) = \{a + b \colon a \in I,\, b \in J\}.$$

(v) Seien $I, J \subset A$ Ideale. Dann definieren wir das Idealprodukt durch

$$IJ := (\{ab \colon a \in I,\, b \in J\}).$$

Es gilt

$$IJ = \{a_1 b_1 + \ldots + a_n b_n \colon n \in \mathbb{N},\, a_i \in I,\, b_i \in J,\, i = 1, \ldots, n\},$$

da IJ sämtliche Summen auf der rechten Seite enthalten muss. Da die rechte Seite bereits ein Ideal ist, folgt aufgrund der geforderten Minimalität beim von einer Menge erzeugten Ideal die Gleichheit.

Es gilt $IJ \subset I \cap J$ (siehe unten).

Definition 4.1.13. Sei A ein Ring (kommutativ und mit Eins).

(i) Seien $I, J \subset A$ zwei Ideale. Dann heißen I und J relativ prim oder teilerfremd, wenn $I + J = (1)$ gilt. Es gibt also $a \in I$ und $b \in J$ mit $a + b = 1$.

(ii) Seien $I_1, \ldots, I_n$ Ideale von A. Dann heißen die Ideale paarweise relativ prim, falls für beliebige $i \neq j \in \{1, \ldots, n\}$ bereits $I_i + I_j = A$ gilt.

Theorem 4.1.14. *Seien $I, J \subset A$ zwei Ideale. Dann gilt $IJ \subset I \cap J$. Sind I und J relativ prim, so gilt die Gleichheit $IJ = I \cap J$.*

Beweis.

(i) Seien $a \in I$ und $b \in J$. Dann ist $ab \in I \cap J$. Also folgt $IJ \subset I \cap J$.

(ii) Seien I und J relativ prim. Dann gibt es $a \in I$ und $b \in J$ mit $a + b = 1$. Sei nun $c \in I \cap J$. Dann gilt $c = (a + b)c = ac + bc \in IJ$. Also ist $I \cap J \subset IJ$. □

Beispiele 4.1.15 (Ideale von $\mathbb{Z}$)**.** Seien $a, b \in \mathbb{Z}$. Sei $\mathbb{P}$ die Menge aller Primzahlen. Seien $\alpha_i, \beta_i \in \mathbb{N}$, $i \in \mathbb{P}$, fast alle Null, $\varepsilon, \delta \in \{\pm 1\}$ und gelte

$$a = \varepsilon \cdot \prod_{p \in \mathbb{P}} p^{\alpha_p} \quad \text{sowie} \quad b = \delta \cdot \prod_{p \in \mathbb{P}} p^{\beta_p}.$$

Aufgrund der Existenz und Eindeutigkeit der Primfaktorzerlegung existiert solch eine Darstellung. Die unendlichen Produkte sind wohldefiniert, da fast alle Faktoren gleich 1 sind. Wir definieren den größten gemeinsamen Teiler (ggT) und das kleinste gemeinsame Vielfache (kgV) durch

$$\mathrm{ggT}(a,b) := \prod_{p\in\mathbb{P}} p^{\min\{\alpha_p,\beta_p\}} \quad \text{sowie} \quad \mathrm{kgV}(a,b) := \prod_{p\in\mathbb{P}} p^{\max\{\alpha_p,\beta_p\}}.$$

ggT und kgV sind aufgrund der Eindeutigkeit der Primfaktorzerlegung wohldefiniert. Es gelten

$$\begin{aligned}(a)+(b) &= (a,b) = (\mathrm{ggT}(a,b)),\\ (a)\cap(b) &= (\mathrm{kgV}(a,b))\end{aligned}$$

sowie

$$(a)\cdot(b) = (ab).$$

Beweis.

(i) $(a)+(b)=(a,b)$ gilt nach Definition. „$\subset$“ ist klar, da jedes Element in (a) und (b) durch $\mathrm{ggT}(a,b)$ teilbar ist. Die Umkehrung gilt, da es $c,d\in\mathbb{Z}$ mit $ac+bd=\mathrm{ggT}(a,b)$ gibt. Dies werden wir später als Konsequenz aus dem Euklidischen Algorithmus zeigen.

(ii) Dies folgt aus der Definition von $\mathrm{kgV}(a,b)$: Für jeden Primfaktor ist die mindestens auftretende Potenz in den Mengen von Zahlen auf beiden Seiten gleich.

(iii) In der Beschreibung des Idealproduktes können wir sämtliche Summanden a_ib_i durch ab oder Null ersetzen und erhalten $(a)\cdot(b)=\{nab\colon n\in\mathbb{Z}\}=(ab)$.

□

Definition 4.1.16. Sei $I\subset A$ ein Ideal. Dann heißt I maximales Ideal von A, falls $I\neq A$ gilt und für jedes Ideal J mit $I\subset J$ entweder $I=J$ oder $J=A$ gilt.

Theorem 4.1.17. *Sei $I\subset A$ ein Ideal. Dann ist I genau dann maximal, wenn A/I ein Körper ist.*

Beweis.

„$\Longrightarrow$“: Sei $a\in A\setminus I$ beliebig. Wir wollen nachweisen, dass $a+I$ in A/I invertierbar ist. Da I maximal ist, gilt $A=I+(a)$. Es existieren daher insbesondere $b\in A$ und $c\in I$ mit $1=c+ba$. Dies impliziert $1+I=(b+I)(a+I)$ in A/I.

„$\Longleftarrow$“: Ist I nicht maximal, so gibt es ein Ideal J mit $I\subsetneq J\subsetneq A$. Sei $a\in J$. Gäbe es $b+I\in A/I$ mit $1\equiv ab \pmod I$, so wäre $1\in I+Aa\subset J$. Also wäre $A=A1\subset J$. Widerspruch. □

Um zu zeigen, dass es in jedem Ring ein maximales Ideal gibt, benötigen wir das Zornsche Lemma.

Lemma 4.1.18 (Zornsches Lemma)**.** *Sei $\mathcal{M}$ nichtleer und $\leq$ eine Halbordnung auf $\mathcal{M}$. Besitzt jede total geordnete Teilmenge $\mathcal{N}\subset\mathcal{M}$ eine obere Schranke in $\mathcal{M}$, also ein $a\in\mathcal{M}$ mit $b\leq a$ für alle $b\in\mathcal{N}$, so besitzt $\mathcal{M}$ mindestens ein maximales Element m. (m ist maximales Element von $\mathcal{M}$, falls $m\in\mathcal{M}$ gilt und aus $m\leq a$, $a\in\mathcal{M}$ stets $a=m$ folgt.)*

Beweis. Vorlesung über Mengenlehre. □

Wir zeigen damit zunächst als Nachtrag, dass jeder Vektorraum eine Basis besitzt.

Theorem 4.1.19. *Sei V ein K-Vektorraum. Dann besitzt V eine Basis.*

Beweis. Wir wollen das Zornsche Lemma benutzen und zeigen, dass V eine maximale linear unabhängige Teilmenge besitzt. Sei $\mathcal{M}$ die Menge aller linear unabhängigen Teilmengen von V. Dann ist $\emptyset \in \mathcal{M}$, die Menge also nicht leer. Als Halbordnung auf $\mathcal{M}$ verwenden wir die Mengeninklusion $\subset$. Sei $\mathcal{N} \subset \mathcal{M}$ eine total geordnete Teilmenge von $\mathcal{M}$. Definiere $S := \bigcup_{N \in \mathcal{N}} N$. Dann ist S eine Teilmenge von V. Wir wollen zeigen, dass S eine obere Schranke für $\mathcal{N}$ ist. $N \subset S$ für beliebiges $N \in \mathcal{N}$ ist nach Definition klar. Also müssen wir noch nachweisen, dass S linear unabhängig ist. Seien dazu $n \in \mathbb{N}$ und $a_1, \dots, a_n \in S$ beliebig. Dann gibt es $N_1, \dots, N_n \in \mathcal{N}$ mit $a_i \in N_i$ für $i = 1, \dots, n$. Da $\mathcal{N}$ total geordnet ist finden wir induktiv ein i_0 mit $N_i \subset N_{i_0}$ für $i = 1, \dots, n$. Also gilt $a_1, \dots, a_n \in N_{i_0}$. Daraus lässt sich wegen der linearen Unabhängigkeit von N_{i_0} die Null nur auf triviale Art und Weise linear kombinieren. Also ist S linear unabhängig. Aufgrund des Zornschen Lemmas existiert nun ein maximales Element $B \in \mathcal{M}$. Dies ist auch eine maximale linear unabhängige Teilmenge von V. □

Theorem 4.1.20. *Sei A ein Ring und $I \subsetneq A$ ein Ideal. Dann gibt es ein maximales Ideal J mit $I \subset J \subset A$.*

Beweis. Wir wollen das Zornsche Lemma benutzen. Sei $\mathcal{M}$ die Menge aller Ideale in A, die I enthalten und ungleich A sind. Dann ist $I \in \mathcal{M}$ und somit gilt $\mathcal{M} \neq \emptyset$. Auf $\mathcal{M}$ verwenden wir $\subset$ als partielle Ordnung. Sei $\mathcal{N} \subset \mathcal{M}$ eine beliebige total geordnete Teilmenge. Wir wollen zeigen, dass $\mathcal{N}$ eine obere Schranke besitzt. Definiere dazu $S := \bigcup_{N \in \mathcal{N}} N$. Dann gilt nach Definition $N \subset S$ für alle $N \in \mathcal{N}$. $I \subset S$ ist ebenfalls klar. Weiterhin ist S ein Ideal: Je endlich viele Elemente $a_i \in S$ liegen in einem gemeinsamen $N \in \mathcal{N}$, da $\mathcal{N}$ total geordnet ist. Damit rechnet man direkt nach, dass S ein Ideal ist.

Wir müssen ausschließen, dass bereits $S = A$ gilt. Wäre $S = A$, so erhalten wir $1 \in S$ und damit $1 \in N$ für ein $N \in \mathcal{N}$. Dies widerspricht aber der Definition von $\mathcal{M}$. Somit können wir das Zornsche Lemma anwenden und erhalten ein maximales Element J. J ist bereits ein maximales Ideal: Sei $a \in A \setminus J$ beliebig. Dann ist $J + (a) \supsetneq J$ ein weiteres Ideal. Da J bezüglich der Halbordnung maximal ist, gilt $J + (a) \notin \mathcal{M}$, also $J + (a) = A$. Damit ist J ein maximales Ideal. □

4.2. **Teilbarkeit in Ringen.** Wir wollen wieder stets kommutative Ringe mit Eins betrachten.

Definition 4.2.1. Sei A ein kommutativer Ring mit Eins. Dann heißt A nullteilerfrei (oder Integritätsring), wenn für $a, b \in A$ mit $ab = 0$ folgt, dass $a = 0$ oder $b = 0$ gilt.

Bemerkung 4.2.2.

(i) In einem nullteilerfreien Ring folgt aus $ac = bc$ mit $c \neq 0$, dass $(a - b)c = 0$ und somit $a = c$ ist.
(ii) Körper sind (als Ringe aufgefasst) nullteilerfrei, ebenso Teilringe nullteilerfreier Ringe.
(iii) Sei A ein nullteilerfreier Ring. Dann ist auch der Polynomring $A[t]$ nullteilerfrei: Betrachte die jeweils höchsten von Null verschiedenen Koeffizienten.
(iv) Sei $n \in \mathbb{N} \setminus \{0\}$. Dann ist $\mathbb{Z}/n\mathbb{Z}$ genau dann nullteilerfrei, wenn n eine Primzahl ist.

Beweis. Wir zeigen nur die letzte Aussage:

„$\Longrightarrow$“: Sei $\mathbb{Z}/n\mathbb{Z}$ nullteilerfrei. Wäre n keine Primzahl, so gäbe es $k, l \in \mathbb{N}$ mit $k, l \in \{1, 2, \dots, n-1\}$ und $kl = n$. Daraus folgt $(k + n\mathbb{Z})(l + n\mathbb{Z}) = 0 + n\mathbb{Z}$. Dies widerspricht der Annahme, dass $\mathbb{Z}/n\mathbb{Z}$ nullteilerfrei ist.

„$\Longleftarrow$“: Sei n eine Primzahl. Gelte $(k+n\mathbb{Z})(l+n\mathbb{Z}) = 0+n\mathbb{Z}$ für $k, l \in \mathbb{Z}$. Somit gibt es ein $r \in \mathbb{Z}$ mit $kl = rn$. Also folgt $n|kl$. Da n eine Primzahl ist folgt $n|k$ oder $n|l$. Also gilt $k+n\mathbb{Z} = 0+n\mathbb{Z}$ oder $l+n\mathbb{Z} = 0+n\mathbb{Z}$. Somit ist $\mathbb{Z}/n\mathbb{Z}$ nullteilerfrei. □

Bemerkung 4.2.3. Sei A ein kommutativer Ring mit Eins. Dann ist $A^* := \{a \in A\colon \exists\, b \in A \text{ mit } ab = 1\}$ eine abelsche Gruppe (Direktes Nachrechnen). Die Elemente von A^* heißen Einheiten.

Sei $a \in A^*$. Für das Element $b \in A$ mit $ab = 1$ schreiben wir $b = a^{-1}$. Es gilt $a^{-1} \in A^*$. Das Element $b \in A$ ist eindeutig bestimmt: $b = b\left(a\tilde{b}\right) = \tilde{b}$.

Es gilt $\mathbb{Z}^* = \{\pm 1\}$. Sei K ein Körper. Dann ist $(K[t])^* = K^* = K \setminus \{0\}$.

Wir nehmen ab jetzt an, dass sämtliche betrachteten Ringe nicht nur kommutativ sind und eine Eins enthalten sondern auch nullteilerfrei sind.

Definition 4.2.4. Sei A ein Ring (kommutativ, nullteilerfrei, mit Eins). Seien $a, b \in A$.

(i) Wir schreiben $a|b$, d. h. a teilt b, falls es ein $c \in A$ mit $b = ac$ gibt. a heißt dann ein Teiler von b.
(ii) Gilt $a|b$ und $b|a$, so heißen a und b assoziiert: $a \sim b$.

Lemma 4.2.5. *Sei A ein Ring (kommutativ, nullteilerfrei, mit Eins). Seien a, b, c, $a', b' \in A$. Dann gelten:*

(i) $a|b \iff (b) \subset (a)$
(ii) Gilt $a|b$ und $a'|b'$, so gilt auch $aa'|bb'$.
(iii) Für $c \neq 0$ sind $a|b$ und $ac|bc$ äquivalent.
(iv) Gelten $a|b$ und $a|c$, so folgt $a|(b+c)$.
(v) Die Aussagen $a \sim b$, $(a) = (b)$ und die Existenz eines $u \in A^*$ mit $b = au$ sind äquivalent

Beweis.

(i) „$\Longrightarrow$“: Gelte $a|b$, also ist $b = ac$ für ein $c \in A$. Daraus folgt $(b) = \{db\colon d \in A\} = \{(dc)a\colon d \in A\} \subset \{fa\colon f \in A\} = (a)$.
„$\Longleftarrow$“: Aus $(b) \subset (a)$ folgt insbesondere, dass es ein $c \in A$ mit $ca = b$ gibt. Somit folgt $a|b$.
(ii) Klar.
(iii) „$\Longrightarrow$“: Ist Klar.
„$\Longleftarrow$“: Gelte $ac|bc$. Somit gibt es ein $d \in A$ mit $acd = bc$ oder $(ad-b)c = 0$. Da A nullteilerfrei ist folgt $ad = b$ und somit $a|b$.
(iv) Klar.
(v) Nach (i) sind die ersten beiden Aussagen äquivalent.
„$\Longrightarrow$“: Ist $a = 0$, so folgt aus $a|b$ bereits $b = 0$ und die Aussage ist trivial. Sei also ohne Einschränkung $a \neq 0$. Gelte $a \sim b$. Nach Definition gibt es somit $c, d \in A$ mit $a = bc$ und $b = ad$. Wir erhalten $a = bc = acd$. Wegen $a \neq 0$ folgt $cd = 1$. Somit sind c und d Einheiten.
„$\Longleftarrow$“: Klar. □

Definition 4.2.6. Sei A ein Ring. Dann heißt $a \in A$ irreduzibel (unzerlegbar), falls $a \notin A^* \cup \{0\}$ gilt und aus $a = bc$, $b, c \in A$, bereits $b \in A^*$ oder $c \in A^*$ folgt.

Ist $a \in A \setminus (A^* \cup \{0\})$ nicht irreduzibel, so heißt a reduzibel.

Bemerkung 4.2.7.

(i) Ein Element $a \notin A^* \cup \{0\}$ ist genau dann irreduzibel, wenn jeder Teiler von a zu 1 oder a assoziiert ist. Wir sagen auch, dass a keine echten Teiler besitzt.
(ii) In $A = \mathbb{Z}$ sind die irreduziblen Elemente genau $\{\pm p\colon p \text{ ist Primzahl}\}$.

(iii) Sei K ein Körper. Dann ist $f \in K[t]$ genau dann irreduzibel, wenn $\deg f \geq 1$ gilt und f keinen Teiler $g \in K[t]$ mit $1 \leq \deg g < \deg f$ besitzt.

Polynome vom Grad eins sind also stets irreduzibel.

(iv) Ist insbesondere $K = \mathbb{C}$, so sind genau die affin linearen Polynome irreduzibel. Dies folgt aus dem Fundamentalsatz der Algebra (jedes nichtkonstante komplexe Polynom besitzt eine Nullstelle), vergleiche eine Vorlesung über Funktionentheorie.

(v) Ist $K = \mathbb{R}$, so sind auch die zu $f = t^2 + at + b$ mit $a^2 - 4b < 0$ assoziierten Polynome irreduzibel.

Beweis. „$\Longrightarrow$“: Sei f wie angegeben. Ist $\deg f = 1$, so ist klar, dass f irreduzibel ist. Ist $\deg f = 2$, leiten wir mit quadratischer Ergänzung zwei Nullstellen her (Mitternachtsformel): Wir betrachten zunächst f vermöge $\mathbb{R} \subset \mathbb{C}$ als komplexes Polynom. Schreibe $t^2 + at + b = 0$ als $\left(t + \frac{a}{2}\right)^2 - \frac{a^2}{4} + b = 0$ und weiter als $\left(t + \frac{a}{2}\right)^2 = -\underbrace{\left(b - \frac{a^2}{4}\right)}_{>0}$. Wir erhalten $t + \frac{a}{2} = \pm i\sqrt{b - \frac{a^2}{4}}$ oder $t = -\frac{a}{2} \pm i\sqrt{b - \frac{a^2}{4}}$. Man überprüft direkt, dass dies Nullstellen sind. Wegen der Polynomdivision (demnächst) wissen wir, dass ein Polynom n-ten Grades maximal n Nullstellen besitzen kann. Somit besitzt f keine reellen Nullstellen. Wegen $\deg f = 2$ ist f damit irreduzibel.

„$\Longleftarrow$“: Sei $f \in \mathbb{R}[t]$ mit $\deg f > 1$ irreduzibel. Vermöge $\mathbb{R} \subset \mathbb{C}$ können wir f als komplexes Polynom auffassen. Dann zerfällt f über $\mathbb{C}$ in Linearfaktoren, d. h. f ist ein Produkt affin linearer Funktionen vom Grad eins. Jede dieser affin linearen Funktionen besitzt genau eine Nullstelle. Ist eine dieser Nullstellen reell, so sind wir fertig. Sei c eine Nullstelle. Dann ist auch $\bar{c}$ eine Nullstelle, denn es gilt

$$f(\bar{c}) = \bar{c}^2 + a\bar{c} + b = \overline{c^2 + ac + b} = \bar{0} = 0.$$

Nun ist $(t-c)(t-\bar{c}) = t^2 - (c+\bar{c})t - c\bar{c}$ ein reelles Polynom, das f teilt. Somit ist $\deg f \geq 3$ für ein irreduzibles Polynom ausgeschlossen. Mit quadratischer Ergänzung wie oben sehen wir, dass ein Polynom zweiten Grades mit $a^2 - 4b \geq 0$ reduzibel ist. Man erhält direkt die Nullstellen $t = -\frac{a}{2} \pm \sqrt{\frac{a^2}{4} - b}$. □

(vi) In $\mathbb{Z}$ können wir eine Nicht-Primzahl als Produkt von zwei Nicht-Einheiten schreiben. Ist eine dieser Zahlen noch keine Primzahl, so können wir sie wiederum als Produkt von zwei Nicht-Einheiten schreiben. Beispiel: $12 = 2{\cdot}6 = 2{\cdot}2{\cdot}3$. Nach endlich vielen Schritten bricht dies ab, da jeder der Faktoren einer Zahl einen kleineren Betrag als das Produkt besitzt. In allgemeinen Ringen braucht solch ein Vorgehen nicht abzubrechen, jedoch in Hauptidealringen (Definition folgt nach). In Hauptidealringen wird nämlich jede aufsteigende Idealkette stationär, vergleiche den Beweis von Theorem 4.2.14.

Definition 4.2.8. Sei A ein Ring (kommutativ, nullteilerfrei, mit Eins). Dann heißt A Hauptidealring, falls jedes Ideal in A ein Hauptideal ist.

Bemerkung 4.2.9.

(i) In einem Körper gibt es nur Hauptideale, nämlich (0) und (1).

(ii) Jeder euklidische Ring ist ein Hauptidealring (Definition und Beweis folgen). Somit sind $\mathbb{Z}$ und $K[t]$, K ein Körper, Hauptidealringe.

Definition 4.2.10. Sei A ein Hauptidealring. Seien $a_1, \ldots, a_n \in A$. Da A ein Hauptidealring ist, gibt es somit $b, c \in A$ mit

$$(a_1, \ldots, a_n) = (b) \quad \text{sowie} \quad (a_1) \cap \ldots \cap (a_n) = (c).$$

Wir nennen b einen größten gemeinsamen Teiler von $a_1, \ldots, a_n$, $b = \mathrm{ggT}(a_1, \ldots, a_n)$, und c ein kleinstes gemeinsames Vielfaches von $a_1, \ldots, a_n$, $c = \mathrm{kgV}(a_1, \ldots, a_n)$.

Lemma 4.2.11. *Sie A ein Hauptidealring. Seien $a_1, \ldots, a_n, b, c \in A$. Gelte $b = \mathrm{ggT}(a_1, \ldots, a_n)$ sowie $c = \mathrm{kgV}(a_1, \ldots, a_n)$. Dann gelten*

(i) $b|a_i$ für alle $i = 1, \ldots, n$.
(ii) Ist $b' \in A$ und gilt $b'|a_i$ für alle $i = 1, \ldots, n$, so folgt $b'|b$.
(iii) $a_i|c$ für alle $i = 1, \ldots, n$.
(iv) Ist $c' \in A$ und gilt $a_i|c'$ für alle $i = 1, \ldots, n$, so folgt $c|c'$.

Beweis.

(i) Es gilt $(a_i) \subset (a_1, \ldots, a_n) = (b)$ für alle $i = 1, \ldots, n$. Also folgt $b|a_i$ für $i = 1, \ldots, n$.
(ii) Gilt $b'|a_i$, $i = 1, \ldots, n$, so folgt $(a_i) \subset (b')$ für $i = 1, \ldots, n$ und daher auch $(a_1, \ldots, a_n) \subset (b')$. Wegen $(b) = (a_1, \ldots, a_n) \subset (b')$ erhalten wir $b'|b$.
(iii) Nach Definition des kgV folgt $(c) \subset (a_i)$, also auch $a_i|c$ für alle $i = 1, \ldots, n$.
(iv) Gelte $a_i|c'$ für alle $i = 1, \ldots, n$. Dann gilt $(c') \subset (a_i)$ für alle $i = 1, \ldots, n$. Es folgt $(c') \subset (a_1) \cap \ldots \cap (a_n) = (c)$. Also gilt $c|c'$. □

Bemerkung 4.2.12.

(i) Das Lemma rechtfertigt die Bezeichnungen „größter gemeinsamer Teiler“ bzw. „kleinstes gemeinsames Vielfaches“.
(ii) Beachte, dass ggT und kgV nur bis auf Einheiten bestimmt sind. Um dies zu betonen schreiben wir auch $\mathrm{ggT} \sim \ldots$.
(iii) Ist A kein Hauptidealring, so brauchen ggT oder kgV nicht zu existieren.

Als Vorbereitung für die Zerlegung in irreduzible Elemente zeigen wir

Lemma 4.2.13. *Sei A ein Hauptidealring und $p \in A$ irreduzibel. Seien $a_1, \ldots, a_n \in A$ und gelte $p|a_1 \cdot \ldots \cdot a_n$. Dann gibt es ein $i \in \{1, \ldots, n\}$ mit $p|a_i$.*

Beweis. Wir zeigen die Aussage nur für den Fall $n = 2$, also $p|ab$. Der allgemeine Fall folgt dann per Induktion. Wir unterscheiden zwei Fälle:

(i) Gilt $(a, p) = A = (1)$, so gibt es $r, s \in A$ mit $ra + sp = 1$. Wir multiplizieren diese Gleichung mit b und erhalten $r\underline{ab} + bs\underline{p} = b$. Da p die unterstrichenen Terme teilt folgt auch $p|b$.
(ii) Gelte nun $(a, p) \neq (1)$. Da A ein Hauptidealring ist, gibt es ein $c \in A$ mit $(a, p) = (c)$. Wegen $(c) \neq (1)$ ist c keine Einheit. Nach Definition ist $c = \mathrm{ggT}(a, p)$. Also folgt $c|p$. Da p irreduzibel ist, erhalten wir $c \sim p$. Aus $c|a$ folgt also auch $p|a$. □

Theorem 4.2.14. *Sei A ein Hauptidealring. Sei $a \in A \setminus (A^* \cup \{0\})$. Dann gibt es $n \in \mathbb{N}$ und irreduzible Elemente $p_1, \ldots, p_n \in A$ mit*

$$a = p_1 \cdot \ldots \cdot p_n.$$

Bis auf die Anordnung und die Wahl anderer assoziierter irreduzibler Elemente sind die Faktoren eindeutig bestimmt.

Beweis. **Existenz:** Wir behaupten, dass sich jedes Element $a \in A \setminus (A^* \cup \{0\})$ als ein endliches Produkt von irreduziblen Elementen schreiben lässt. Widerspruchsbeweis: Angenommen, a wäre ein Gegenbeispiel. Dann ist a selbst nicht irreduzibel, besitzt also eine Darstellung $a = a_1 \cdot a_1'$ mit $a_1, a_1' \in A \setminus A^*$. Nach Annahme lässt sich mindestens einer der beiden Faktoren nicht als endliches Produkt von irreduziblen Elementen schreiben, ohne Einschränkung a_1. Wir können daher $a_1 = a_2 \cdot a_2'$

mit $a_2, a_2' \in A \setminus A^*$ schreiben. Nach Annahme lässt sich mindestens einer der beiden Faktoren wiederum nicht als endliches Produkt von irreduziblen Elementen schreiben, ohne Einschränkung a_2. Induktiv erhalten wir somit eine Folge

$$a = a_0, a_1, a_2, \ldots \in A$$

mit der Eigenschaft $a_{i+1}|a_i$ für alle $i \in \mathbb{N}$ aber $a_i \not\sim a_{i+1}$. Dies können wir auch als Idealkette

$$(a_0) \subsetneq (a_1) \subsetneq (a_2) \subsetneq \ldots$$

schreiben. Wir definieren $I := \bigcup_{i\in\mathbb{N}} (a_i)$. Da die Vereinigung einer Folge von aufsteigenden Idealen wieder ein Ideal ist (leicht nachzurechnen), ist I ein Ideal. Nun ist A ein Hauptidealring, also $I = (b)$ für ein $b \in A$. Nach Definition von I gibt es ein i_0 mit $b \in (a_{i_0})$. Wir erhalten

$$(b) \subset (a_{i_0}) \subsetneq (a_{i_0+1}) \subset (b).$$

Widerspruch.

Eindeutigkeit: Gelte $a = p_1 \cdot \ldots \cdot p_r = q_1 \cdot \ldots \cdot q_s$ mit $r, s \in \mathbb{N}$. Es folgt $p_1 \cdot \ldots \cdot p_r \sim q_1 \cdot \ldots \cdot q_s$. Es gilt $p_r | q_1 \cdot \ldots \cdot q_s$. Nach Lemma 4.2.13 teilt p_r einen der Faktoren auf der rechten Seite, ohne Einschränkung q_s. Da q_s selbst irreduzibel ist, folgt $p_r \sim q_s$. Schreiben wir die Relation $\sim$ als Gleichheit mit Hilfe einer zusätzlichen Einheit, so können wir kürzen und erhalten $p_1 \cdot \ldots \cdot p_{r-1} \sim q_1 \cdot \ldots \cdot q_{s-1}$. Nach endlich vielen Schritten sind die Faktoren auf einer Seite aufgebraucht und wir erhalten beispielsweise $1 \sim q_1 \cdot \ldots \cdot q_{s-r}$. Für $s \neq r$ erhalten wir einen Widerspruch, da ein irreduzibles Element keine Einheit sein kann, für $r = s$ folgt die Behauptung. □

Definition 4.2.15. Sei A ein Hauptidealring. $P \subset A$ heißt ein Vertretersystem der irreduziblen Elemente von A, wenn P aus jeder Äquivalenzklasse dieser Elemente unter der Äquivalenzrelation $\sim$ genau ein Element enthält.

Beispiele 4.2.16.

(i) Ist $A = \mathbb{Z}$, so sind die Primzahlen ein solches Vertretersystem irreduzibler Elemente in A.

(ii) Ist $A = K[t]$, K ein Körper, so wählen wir unter den jeweils assoziierten Polynomen das normierte aus, d. h. das mit führendem Koeffizienten Eins (den Nachweis, dass $K[t]$ ein Hauptidealring ist, liefern wir noch nach). Dies ist ein solches Vertretersystem irreduzibler Elemente in $K[t]$.

Aus Theorem 4.2.14 erhalten wir

Korollar 4.2.17. *Sei A ein Hauptidealring und P ein Vertretersystem der irreduziblen Elemente in A. Sei $f \in A \setminus \{0\}$. Dann gibt es $u \in A^*$ und $v_p(f) \in \mathbb{N}$, $p \in P$, wobei fast alle $v_p(f)$ verschwinden, so dass*

$$f = u \cdot \prod_{p\in P} p^{v_p(f)}$$

gilt.

Korollar 4.2.18. *Sei A ein Hauptidealring und seien $a, b \in A \setminus \{0\}$. Sei P ein Vertretersystem der irreduziblen Elemente von A. Dann sind folgende Aussagen äquivalent:*

(i) $a|b$

(ii) $v_p(a) \leq v_p(b)$ *für alle* $p \in P$.

Beweis. „(ii) $\Longrightarrow$ (i)“: Klar.

„(i) $\Longrightarrow$ (ii)“: Gelte $a|b$, also $c \cdot u_a \cdot \prod_{p\in P} p^{v_p(a)} = u_b \cdot \prod_{p\in P} p^{v_p(b)}$ für ein $c \in A$. Gelte $v_p(a) \geq 1$. Dann teilt p die linke Seite und daher auch die rechte Seite. Da ein

irreduzibles Element kein nicht assoziiertes irreduzibles Element teilen kann folgt $v_p(b) \geq 1$. Wir teilen nun beide Seiten durch p. Per Induktion nach $v_p(a)$ folgt daraus die Behauptung. □

Aus der Darstellung in Korollar 4.2.17 können wir auch ggT und kgV ablesen:

Korollar 4.2.19. *Sei A ein Hauptidealring. Seien $a, b \in A \setminus \{0\}$. Dann gelten*

$$\mathrm{ggT}(a,b) \sim \prod_{p \in P} p^{\min\{v_p(a), v_p(b)\}} \quad \textit{sowie} \quad \mathrm{kgV}(a,b) \sim \prod_{p \in P} p^{\max\{v_p(a), v_p(b)\}}.$$

Korollar 4.2.20. *Sei A ein Hauptidealring und $p \in A \setminus \{0\}$. Dann sind die folgenden Aussagen äquivalent:*

(i) p ist irreduzibel.
(ii) (p) ist ein maximales Ideal von A.
(iii) $A/(p)$ ist ein Körper.

Beweis. Nach Theorem 4.1.17 müssen wir nur noch die Äquivalenz von (i) und (ii) zeigen.

„(i) $\Longrightarrow$ (ii)“: Da p irreduzibel ist gilt $(p) \neq (1)$. Wäre (p) nicht maximal, so gäbe es ein $b \in A$ (da A ein Hauptidealring ist) mit $(p) \subsetneq (b) \subsetneq (1)$. Daher gilt $b|p$ und $p \not\sim b$. Dies widerspricht der Irreduzibilität von p.

„(ii) $\Longrightarrow$ (i)“: Wäre p reduzibel, so gäbe es Nicht-Einheiten $a, b \in A$ mit $p = ab$. Somit wäre $(p) \subsetneq (a) \subsetneq (1)$. Dies widerspricht aber der Maximalität von (p). □

4.3. **Euklidische Ringe.** Euklidische Ringe verallgemeinern die (vermutlich bereits aus der Schule bekannte) Division mit Rest für $\mathbb{Z}$ und $K[t]$.

Definition 4.3.1. Sei A ein kommutativer nullteilerfreier Ring mit Eins. Eine euklidische Wertefunktion (Gradfunktion) auf A ist eine Abbildungen $\delta\colon A \setminus \{0\} \to \mathbb{N}$ mit der folgenden Eigenschaft: Seien $a \in A$, $b \in A \setminus \{0\}$. Dann gibt es $q, r \in A$ mit $a = qb + r$ und ($r = 0$ oder $\delta(r) < \delta(b)$). Wir nennen dies Division mit Rest und r den Rest.

Ein kommutativer nullteilerfreier Ring mit Eins und einer euklidischen Wertefunktion heißt euklidischer Ring.

Bemerkung 4.3.2.

(i) Auf $A = \mathbb{Z}$ definiert $\delta(a) = |a|$, $a \neq 0$, eine Gradfunktion. Beispiel für eine Division mit Rest ist im Falle $a = 13$ und $b = 7$ die Gleichheit $13 = 1 \cdot 7 + 6 = 2 \cdot 7 - 1$, q und r sind also im Allgemeinen nicht eindeutig bestimmt.

(ii) Ist K ein Körper, so ist $A = K[t]$ ein euklidischer Ring mit Gradfunktion $\delta(f) := \deg(f)$ für $f \neq 0$.

Beweis: Um $a = qb + r$ mit den gewünschten Eigenschaften zu erzielen sei ohne Einschränkung $\deg a > \deg b$. Wir subtrahieren ein Vielfaches von b von a, so dass der Grad strikt abnimmt. Nach endlich vielen Schritten erhalten wir damit das gewünschte Ergebnis.

Beispiel: In $\mathbb{F}_2[t]$ wollen wir $t^4 + t^3 + 1$ mit Rest durch $t^2 + 1$ dividieren. Wir erhalten

$$\begin{aligned}
\left(t^4 + t^3 + 1\right) - t^2 \cdot \left(t^2 + 1\right) &= t^4 + t^3 + 1 + t^4 + t^2 = t^3 + t^2 + 1, \\
\left(t^3 + t^2 + 1\right) - t \cdot \left(t^2 + 1\right) &= t^3 + t^2 + 1 + t^3 + t = t^2 + t + 1, \\
\left(t^2 + t + 1\right) - 1 \cdot \left(t^2 + 1\right) &= t^2 + t + 1 + t^2 + 1 = t,
\end{aligned}$$

also

$$\left(t^4 + t^3 + 1\right) = \left(t^2 + t + 1\right) \cdot \left(t^2 + 1\right) + t.$$

Im Fall $A = K[t]$ sind q, r eindeutig bestimmt: Gelte $a = pb + q = sb + r$ wie in der Division mit Rest gefordert. Gelte $\deg a \geq \deg b$. Sonst ist aus Gradgründen $p = s = 0$. Nun ergibt sich der führende Koeffizient von a als Produkt der führenden Koeffizienten von pb und sb. Somit stimmen die führenden Koeffizienten von p und s und die Grade überein. Seien diese führenden Terme durch ct^k gegeben. Wir subtrahieren überall $ct^k \cdot b$. Somit haben wir den Grad von a um Eins erniedrigt. Die Behauptung folgt nun per Induktion.

(iii) $\mathbb{Z}[i]$ mit $i^2 = -1$ ist ein euklidischer Ring (Übung).

Theorem 4.3.3. *Sei (A, δ) ein euklidischer Ring. Dann ist A ein Hauptidealring.*

Beweis. Sei $I \subset A$ ein beliebiges Ideal von A. Gelte ohne Einschränkung $A \neq (0)$. Da δ Werte in $\mathbb{N}$ annimmt, gibt es ein $a \in I$ mit $\delta(a) = \min\limits_{I \setminus \{0\}} \delta$. Wir behaupten, dass $(a) = I$ gilt. Sei $b \in I$ beliebig, so erhalten wir nach Division mit Rest $b = ca + r$ für $c \in A$ und $r \in A$ mit $r = 0$ oder $\delta(r) < \delta(a)$. Wegen $a, b \in I$ folgt $r \in I$. Ist $r = 0$, so sind wir fertig. Sonst ist $\delta(r) < \delta(a)$ was aber wegen $\delta(a) = \min\limits_{I} \delta$ nicht sein kann. Es gilt also $b = ca$. Da b beliebig war erhalten wir $I = (a)$. □

Daraus folgen die bereits angekündigten Beispiele für Hauptidealringe.

Korollar 4.3.4. *$\mathbb{Z}$ und $K[t]$, K ein Körper, sind Hauptidealringe.*

Aus Korollar 4.2.17 erhalten wir damit die folgenden beiden Resultate.

Theorem 4.3.5. *Sei $n \in \mathbb{N} \setminus \{0, 1\}$. Dann lässt sich n als Produkt von Primzahlen darstellen. Die Darstellung ist bis auf die Reihenfolge eindeutig.*

Theorem 4.3.6. *Sei K ein Körper. Sei $f \in K[t] \setminus \{0\}$. Dann gibt es eine Darstellung $f = c \cdot f_1 \cdot \ldots \cdot f_r$, wobei $c \in K^*$ ist und f_i normierte irreduzible Polynome in $K[t]$ sind. Bis auf die Reihenfolge ist diese Darstellung eindeutig.*

Euklidischer Algorithmus 4.3.7. Sei A ein euklidischer Ring. Seien $a, b \in A$ mit $b \neq 0$. Dann können wir iterativ wie folgt eine Division mit Rest durchführen:

$$\begin{aligned}
a &= q_0 b + r_1, && \delta(r_1) < \delta(b), && r_1 \neq 0,\\
b &= q_1 r_1 + r_2, && \delta(r_2) < \delta(r_1), && r_2 \neq 0,\\
r_1 &= q_2 r_2 + r_3, && \delta(r_3) < \delta(r_2), && r_3 \neq 0,\\
&\vdots && \vdots && \vdots\\
r_i &= q_{i+1} r_{i+1} + r_{i+2}, && \delta(r_{i+2}) < \delta(r_{i+1}), && r_{i+2} \neq 0,\\
&\vdots && \vdots && \vdots\\
r_{m-2} &= q_{m-1} r_{m-1} + r_m, && \delta(r_m) < \delta(r_{m-1}), && r_m \neq 0,\\
r_{m-1} &= q_m r_m + 0.
\end{aligned}$$

Da $i \mapsto \delta(r_i)$ strikt monoton fallend ist, bricht diese iterierte Division mit Rest nach endlich vielen Schritten ab.

Theorem 4.3.8. *Mit den Bezeichnungen aus dem euklidischen Algorithmus gilt $(a, b) = (r_m)$, also $r_m \sim \mathrm{ggT}(a, b)$.*

Beweis. Von oben beginnend erhalten wir induktiv $r_1 \in (a, b)$, $r_2 \in (a, b)$, …, $r_m \in (a, b)$. Somit gilt $(r_m) \subset (a, b)$.

Von unten beginnend erhalten wir induktiv $r_{m-1} \in (r_m)$, $r_{m-2} \in (r_m)$, $r_{m-3} \in (r_m)$, …, $r_1 \in (r_m)$, $b \in (r_m)$ und $a \in (r_m)$. Wir erhalten $(a, b) \subset (r_m)$. Die Behauptung folgt. □

Bemerkung 4.3.9.

(i) Wir können den euklidischen Algorithmus auch verwenden um den größten gemeinsamen Teiler $d = \mathrm{ggT}(a, b)$ als $d = ra+sb$ darzustellen. Dazu setzen wir iterativ die drittletzte, viertletzte, ..., zweite, erste Gleichung in die vorletzte Gleichung ein.

(ii) Beispiel: In $\mathbb{Z}$ wollen wir $\mathrm{ggT}(91, 17)$ bestimmen und als Linearkombination von 91 und 17 darstellen. Es gilt

$$\begin{aligned} 91 &= 5 \cdot 17 + 6, \\ 17 &= 3 \cdot 6 - 1, \\ 6 &= (-6) \cdot (-1). \end{aligned}$$

Durch Einsetzen erhalten wir von unten her

$$-1 = 17 - 3 \cdot 6 = 17 - 3 \cdot (91 - 5 \cdot 17) = -3 \cdot 91 + 16 \cdot 17$$

oder $1 = 3 \cdot 91 - 16 \cdot 17$.

(iii) Sei nun $A = \mathbb{F}_2[t]$, $f = t^7 + t^6 + t^4 + t^3 + t^2 + t + 1$, $g = t^3 + 1$. Wir erhalten

$$\begin{aligned} t^7 + t^6 + t^4 + t^3 + t^2 + t + 1 &= \left(t^4 + t^3\right) \cdot \left(t^3 + 1\right) + t^2 + t + 1, \\ t^3 + 1 &= (t + 1) \cdot \left(t^2 + t + 1\right) + 0 \end{aligned}$$

und damit

$$\mathrm{ggT}(f, g) = t^2 + t + 1 = f + \left(t^4 + t^3\right) g.$$

(iv) Will man $\mathrm{ggT}(a_1, a_2, a_3)$ bestimmen und ihn aus den a_i's linear kombinieren, so benutzt man

$$(a_1, a_2, a_3) = ((a_1, a_2), a_3)$$

und wendet den euklidischen Algorithmus mehrfach an. Dies funktioniert analog auch für mehr als drei Faktoren.

Bemerkung 4.3.10 (Simultane Kongruenzen)**.** Beschreibung des Problems: Sei A ein Ring (kommutativ, nullteilerfrei und mit Eins) und seien $I_1, \ldots, I_n$ Ideale von A. Seien $a_1, \ldots, a_n \in A$ gegeben. Wir wollen das System

$$\begin{aligned} x &\equiv a_1 && (\mathrm{mod}\, I_1), \\ x &\equiv a_2 && (\mathrm{mod}\, I_2), \\ &\vdots \\ x &\equiv a_n && (\mathrm{mod}\, I_n) \end{aligned}$$

auf Lösbarkeit untersuchen. Dies lässt sich wie folgt umformulieren: Definiere

$$\begin{aligned} \varphi\colon A &\to A/I_1 \times \ldots \times A/I_n, \\ \varphi(x) &= (x + I_1, \ldots, x + I_n). \end{aligned}$$

Die simultane Lösbarkeit der obigen Kongruenzen ist äquivalent zu $(a_1+I_1, \ldots, a_n+I_n) \in \mathrm{im}\, \varphi$. Es gilt $\ker \varphi = I_1 \cap \ldots \cap I_n$. Gibt es überhaupt eine Lösung $x_0 \in A$, so ist eine beliebige Lösung aus $x_0 + \ker \varphi$.

Bevor wir die simultanen Kongruenzen in Bemerkung 4.3.10 lösen können benötigen wir

Proposition 4.3.11. *Sei A ein Ring (kommutativ, nullteilerfrei und mit Eins). Seien $I, J_1, \ldots, J_r$ Ideale von A. Gilt $I + J_i = A$ für alle $i = 1, \ldots, r$, so folgt $I + J_1 \cdot \ldots \cdot J_r = A$ und insbesondere $I + J_1 \cap \ldots \cap J_r = A$.*

Beweis. Nach Voraussetzung gibt es $u_i \in I$ und $v_i \in J_i$ mit $u_i + v_i = 1$ für $i = 1, \ldots, r$. Wir multiplizieren diese Gleichungen und erhalten

$$1 = (u_1 + v_1) \cdot \ldots \cdot (u_r + v_r) \in v_1 \cdot \ldots \cdot v_r + I.$$

Alle weiteren Summanden beim Ausmultiplizieren enthalten mindestens einen Faktor $u_i \in I$ und können daher wie angegeben zusammengefasst werden. Es folgt $1 \in J_1 \cdot \ldots J_r + I$.

Bei der Definition des Idealproduktes haben wir bereits bemerkt, dass für Ideale K, L stets $KL \subset K \cap L$ gilt. Da wir sowohl beim Idealprodukt als auch beim Schnitt beliebig klammern dürfen folgt die letzte Behauptung per Induktion. □

Theorem 4.3.12 (Chinesischer Restsatz)**.** *Sei A ein Ring (kommutativ, nullteilerfrei und mit Eins). Seien $I_1, \ldots, I_n$ paarweise relativ prime Ideale von A. Dann ist*

$$\varphi\colon A \to A/I_1 \times \ldots \times A/I_n,$$
$$\varphi(x) = (x + I_1, \ldots, x + I_n)$$

surjektiv. Für beliebige $a_1, \ldots, a_n \in A$ gibt es also eine simultane Lösung der Kongruenzen in Bemerkung 4.3.10.

Beweis. Sei zunächst $n = 2$. Seien $a_1, a_2 \in A$ beliebig vorgegeben. Wir suchen nun $x \in A$ mit $x \equiv a_i \pmod{I_i}$ für $i = 1, 2$. Nach Voraussetzung gibt es $b_1 \in I_1$ und $b_2 \in I_2$ mit $b_1 + b_2 = 1$. Setze $x := a_1 b_2 + a_2 b_1$. Da $b_1 \in I_1$ ist folgt $x \equiv a_1 b_2 \equiv a_1(b_2 + b_1) \equiv a_1 \pmod{I_1}$. Ebenso erhalten wir $x \equiv a_2 \pmod{I_2}$.

Für $n > 2$ führen wir den Beweis per Induktion. Seien $a_1, \ldots, a_n \in A$ beliebig. Nach Induktionsannahme dürfen wir annehmen, dass wir $n-1$ Kongruenzen bereits simultan lösen können, dass also ein $y \in A$ mit $y \equiv a_i \pmod{I_i}$ für $i = 2, \ldots, n$ existiert. Nach Proposition 4.3.11 dürfen wir den Fall $n = 2$ auf die Ideale I_1 und $I_2 \cap \ldots \cap I_n$ anwenden. Somit gibt es ein $x \in A$ mit $x \equiv a_1 \pmod{I_1}$ und $x \equiv y \pmod{I_2 \cap \ldots \cap I_n}$. Daraus folgt $x \equiv y \equiv a_i \pmod{I_i}$ für $i = 2, \ldots, n$ und damit die Behauptung. □

Korollar 4.3.13. *Sei A ein Ring (kommutativ, nullteilerfrei und mit Eins). Seien $I_1, \ldots, I_n$ paarweise teilerfremde Ideale von A. Dann ist die Abbildung*

$$A/(I_1 \cap \ldots \cap I_n) \xrightarrow{\sim} (A/I_1) \times \ldots \times (A/I_n)$$

mit

$$a + I_1 \cap \ldots \cap I_n \mapsto (a + I_1, \ldots, a + I_n)$$

ein Ringisomorphismus.

Beweis. Aufgrund des chinesischen Restsatzes ist die Abbildung surjektiv. Nach Herausdividieren des Kernes erhalten wir aufgrund des Homomorphiesatzes einen Isomorphismus. □

Bemerkung 4.3.14 (Konstruktive Lösung)**.** Sei A ein euklidischer Ring. Seien $a_i, b_i \in A$, $i = 1, \ldots, n$, und seien die Ideale (b_i) paarweise relativ prim. Wir wollen die simultanen Kongruenzen

$$x \equiv a_i \pmod{b_i} \quad \text{für } i = 1, \ldots, n \tag{4.1}$$

lösen. Definiere

$$c_i := b_1 \cdot \ldots \cdot b_{i-1} \cdot b_{i+1} \cdot \ldots \cdot b_n, \quad i = 1, \ldots, n.$$

Nach Proposition 4.3.11 gilt $I_i + I_1 \cdot \ldots \cdot I_{i-1} \cdot I_{i+1} \cdot \ldots \cdot I_n = A$. Wegen $I_1 \cdot \ldots \cdot I_{i-1} \cdot I_{i+1} \cdot \ldots \cdot I_n = (b_1 \cdot \ldots \cdot b_{i-1} \cdot b_{i+1} \cdot \ldots \cdot b_n) = (c_i)$ folgt also $(b_i, c_i) = A$. Somit gibt es $u_i, v_i \in A$, die man mit dem euklidischen Algorithmus bestimmen kann, so dass $u_i b_i + v_i c_i = 1$ für $i = 1, \ldots, n$ gilt. Insbesondere folgt für $i \neq j$ nach Definition von c_j

$$c_i v_i \equiv 1 \pmod{b_i} \quad \text{und} \quad c_i v_i \equiv 0 \pmod{b_j}.$$

Somit ist

$$x_0 := \sum_{i=1}^{n} a_i c_i v_i$$

eine Lösung der simultanen Kongruenzen (4.1).

Nach Voraussetzung sind die Ideale (b_i) paarweise relativ prim zueinander. Somit ist $\mathrm{ggT}(b_i, b_j) \sim 1$ für $i \neq j$, nach Wahl eines Vertretersystems von irreduziblen Elementen von A taucht also kein irreduzibler Faktor in der Produktdarstellung wie in Korollar 4.2.17 für $i \neq j$ sowohl bei b_i als auch bei b_j auf: $v_p(b_i) \cdot v_p(b_j) = 0$ für alle Vertreter p und alle $i \neq j$. Somit ist der Kern der zu den simultanen Kongruenzen (4.1) gehörigen Abbildung durch $(b_1 \cdot \ldots \cdot b_n)$ gegeben. Eine beliebige Lösung von (4.1) liegt also in der Menge $x_0 + (b_1 \cdot \ldots \cdot b_n)$.

Beispiel 4.3.15. Sei $A = \mathbb{Z}$. Wir hatten bereits gesehen, dass $\mathrm{ggT}(91, 17) = 1$ und $1 = 3 \cdot 91 - 16 \cdot 17$ gelten. Seien $a, b \in \mathbb{Z}$ beliebig. Wir wollen die simultanen Kongruenzen

$$\begin{aligned} x &\equiv a \pmod{91}, \\ x &\equiv b \pmod{17} \end{aligned}$$

lösen. Da wir nur zwei Kongruenzen haben, brauchen wir die c_i's nicht extra zu berechnen. Aus der Linearkombination der Eins folgt

$$\begin{aligned} 3 \cdot 91 &\equiv 1 \pmod{17}, & -16 \cdot 17 &\equiv 1 \pmod{91}, \\ 3 \cdot 91 &\equiv 0 \pmod{91}, & -16 \cdot 17 &\equiv 0 \pmod{17}. \end{aligned}$$

Daraus lesen wir direkt ab, dass $x \in b \cdot 3 \cdot 91 - a \cdot 16 \cdot 17 + 17 \cdot 91\,\mathbb{Z} = -a \cdot 272 + b \cdot 273 + 1547\,\mathbb{Z}$ die allgemeine Lösung ist.

4.4. **Elementarteilersatz für Matrizen.**

Bemerkung 4.4.1. Sei A ein Hauptidealring. Seien $a_j^i \in A$ sowie $u^i \in A$ für $i = 1, \ldots, m$, $j = 1, \ldots, n$. Betrachte das lineare Gleichungssystem

$$\sum_{j=1}^{n} a_j^i x^j = u^i, \quad i = 1, \ldots, m.$$

Wir suchen Lösungen $x = \left(x^j\right)_{1 \leq i \leq n}$ aller m Gleichungen mit $x^j \in A$, $1 \leq j \leq n$.

Im Spezialfall $m = 1$ haben wir die Gleichung $\sum\limits_{j=1}^{n} a_j x^j = u$. Hier haben wir eine genaue Bedingung für die Lösbarkeit: Gilt $\mathrm{ggT}(a_1, \ldots, a_n) | u$, so ist die Gleichung lösbar.

Beweis.
„$\Longrightarrow$“: Gelte $\mathrm{ggT}(a_1, \ldots, a_n) | u$. Dann stellen wir den größten gemeinsamen Teiler (z. B. mit Hilfe des euklidischen Algorithmusses) als Linearkombination der a_i, $i = 1, \ldots, n$, dar. Nach Multiplikation c, so dass $c \cdot \mathrm{ggT}(a_1, \ldots, a_n) = u$ gilt erhalten wir die gesuchte Lösung.

„$\Longleftarrow$“: Da A ein Hauptidealring ist, können wir jedes Element als Produkt einer Einheit mit irreduziblen Elementen darstellen. Gilt $\mathrm{ggT}(a_1, \ldots, a_n) \nmid u$, so enthält $\mathrm{ggT}(a_1, \ldots, a_n)$ einen irreduziblen Faktor zu einer Potenz, der nur zu einer strikt kleineren Potenz in u auftaucht. Daher können wir die Gleichung nicht lösen. □

Wie bei linearen Gleichungssystemen über Körpern schreiben wir auch lineare Gleichungssysteme über Ringen in Matrixform: $Tx = u$ mit $T = \left(a_i^j\right)_{\substack{1 \leq j \leq m \\ 1 \leq i \leq n}}$, $a_i^j \in A$, $x = \left(x^i\right)_{1 \leq i \leq n}$, $x^i \in A$, $u = \left(u^j\right)_{1 \leq j \leq m}$. Dabei verwenden wir dieselbe

Matrixmultiplikation wie für Matrizen über Körpern und schreiben $T \in A^{m\times n} \equiv M_{m\times n}(A)$, $x \in A^n$, $u \in A^m$.

Bemerkung 4.4.2 (Erinnerung)**.** Sei A ein Ring, $T = \left(a_i^j\right)_{1\le i,j\le n} \in A^{n\times n}$. Definiere $b_j^i := (-1)^{i+j} \cdot \det A_i^j$ und $B = \left(b_j^i\right)_{1\le i,j\le n}$, wobei $\left(A_i^j\right)_{1\le i,j\le n}$ die Matrix ist, die man aus $\left(a_i^j\right)_{i,j}$ erhält, wenn man die j-te Zeile und die i-te Spalte streicht. Dann gilt (Folgerung aus dem Entwicklungssatz von Laplace)

$$T \cdot B = \det T \cdot \mathbb{1}.$$

Der Beweis funktioniert genauso wie über Körpern. Schreibweise: $B = T^{adj}$ oder $B = T^\sharp$. Somit ist eine Matrix $T \in A^{n\times n}$ genau dann invertierbar, wenn $\det T \in A^*$ gilt.

Beweis.
„$\Longrightarrow$“: Sei T invertierbar. Dann gibt es eine Matrix $C \in A^{n\times n}$ mit $TC = \mathbb{1}$. Somit gilt $\det T \cdot \det C = 1$. Also ist $\det T \in A^*$.

„$\Longleftarrow$“: Sei umgekehrt $\det T \in A^*$. Dann besitzt $\det T$ ein Inverses in A und wir definieren mit B wie oben $C := (\det T)^{-1} \cdot B$. Die Folgerung aus dem Entwicklungssatz von Laplace liefert, dass C die Inverse zu T ist. □

Aufgrund des Determinantenmultiplikationssatzes bilden die invertierbaren Matrizen in $A^{n\times n}$ eine Gruppe.

Bemerkung 4.4.3 (Elementare Umformungen)**.** Sei A ein Ring. Sei $T \in A^{m\times n}$. Betrachte die folgenden elementaren Umformungen für T:

(i) Addiere a-mal die j-te Zeile zur i-ten Zeile, $i \neq j$, $a \in A$.
(ii) Vertausche die Zeilen i und j, $i \neq j$.
(iii) Addiere a-mal die l-te Spalte zur k-ten Spalte, $k \neq l$, $a \in A$.
(iv) Vertausche die Spalten k und l, $k \neq l$.

Genauso wie bei Matrizen über Körpern können wir diese elementaren Umformungen wie folgt mit Hilfe von Matrizen darstellen:

(1) $T \mapsto (\mathbb{1} + aE_{ij})T$,
(2) $T \mapsto \Pi_{ij}T$,
(3) $T \mapsto T(\mathbb{1} + aE_{kl})$,
(4) $T \mapsto T\Pi_{kl}$,

wobei $A^{n\times n} \ni E_{ij} = \left(a_l^k\right)_{1\le k,l\le n}$ mit $a_l^k = \delta_i^k\delta_{lj}$ (eine Matrix mit nur einer Eins in Zeile i und Spalte j, sonst Nullen) und $A^{n\times n} \ni \Pi_{ij} = \left(a_l^k\right)_{1\le k,l\le n}$ mit

$$a_l^k = \begin{cases} 1, & \text{für } k,l \in \{i,j\} \text{ und } k \neq l, \\ 0, & \text{für } k,l \in \{i,j\} \text{ und } k = l, \\ \delta_l^k, & \text{sonst.} \end{cases}$$

(eine Diagonalmatrix, bei der die Einträge mit $k,l \in \{i,j\}$ jedoch zu $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ abgeändert sind) gelten. Für Matrizen auf der linken Seite verwenden wir dieselben Definitionen, erlauben jedoch Indices bis m.

Theorem 4.4.4 (Elementarteilersatz)**.**
Sei A ein euklidischer Ring. Sei $T \in A^{m\times n}$. Dann gibt es endlich viele elementare

Umformungen, die T in die Blockgestalt

$$\left(\begin{array}{ccc|c} d_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & d_r & \\ \hline & 0 & & 0 \\ & & & \end{array}\right)$$

mit $0 \leq r \leq \min\{m,n\}$ und $d_i \in A$ sowie $d_1|d_2|\ldots|d_r$ mit $d_i \neq 0$ bringen.

Sei $T = \left(a_i^j\right)_{\substack{1\leq j\leq m\\ 1\leq i\leq n}}$. Ist $T \neq 0$, so gilt

$$d_1 \sim \mathrm{ggT}\left(\left\{a_i^j \colon 1 \leq j \leq m,\, 1 \leq i \leq n\right\}\right) \equiv \mathrm{ggT}(T).$$

Beweis. Dies folgt per Induktion unmittelbar aus den folgenden beiden Aussagen:

(i) $\mathrm{ggT}(T)$ ändert sich nicht, wenn wir elementare Zeilen- und Spaltenumformungen auf T anwenden.

(ii) Durch elementare Umformungen können wir T in die Gestalt

$$\begin{pmatrix} d_1 & (0) \\ (0) & \tilde{T} \end{pmatrix}$$

bringen, so dass $d_1 \sim \mathrm{ggT}(T)$ gilt und d_1 sämtliche Einträge von $\tilde{T}$ teilt.

Wir zeigen diese beiden Behauptungen separat:

(i) Seien $U \in A^{m\times m}$ und $V \in A^{n\times n}$ invertierbar, also $U \in Gl_m(A)$ und $V \in Gl_n(A)$. Teilt ein $a \in A$ sämtliche Einträge in T, so auch in UTV: Beim Berechnen des Matrixproduktes UTV erhalten wir in jedem Eintrag eine Summe von Produkten. In jedem dieser Produkte steht aber ein Eintrag aus T. Also gilt $\mathrm{ggT}(T)|\,\mathrm{ggT}(UTV)$. Wenden wir dasselbe Argument mit U^{-1} statt U und V^{-1} statt V an, so folgt $\mathrm{ggT}(T) \sim \mathrm{ggT}(UTV)$. Insbesondere ändert sich $\mathrm{ggT}(T)$ also nicht, wenn wir T durch elementare Umformungen verändern.

(ii) Sei δ die Gradfunktion des euklidischen Ringes A. Durch Vertauschen von Zeilen und Spalten können wir ohne Einschränkung annehmen, dass $\delta\left(a_1^1\right) = \min\left\{\delta\left(a_j^i\right) : a_j^i \neq 0,\, 1 \leq i \leq m,\, 1 \leq j \leq n\right\} =: \delta(T)$ gilt. Wir argumentieren per Induktion nach diesem Minimalwert $\delta(T)$ und wollen $\delta(T)$ durch elementare Umformungen verkleinern.

Sei zunächst $\delta\left(a_1^1\right) = 0$. Damit ist jedes Element $a \in A$ ohne Rest durch a_1^1 teilbar. Insbesondere können wir also durch elementare Zeilenumformungen erreichen, dass die Einträge in der ersten Spalte (außer ganz oben) Null werden. Ebenso können wir durch elementare Spaltenumformungen erreichen, dass sämtliche Einträge der ersten Zeile (außer ganz links) Null werden.

Sei nun $\delta(T) \geq 1$.

(a) Gilt $a_1^1 \nmid a_1^i$ für ein $i \geq 2$. Division mit Rest liefert

$$a_1^i = q a_1^1 + r \quad \text{mit } r \neq 0 \text{ und } \delta(r) < \delta\left(a_1^1\right) = \delta(T).$$

Subtrahieren wir q-mal die erste Zeile von der i-ten Zeile, so erhalten wir eine Matrix $\tilde{T}$ mit $\delta\left(\tilde{a}_1^i\right) < \delta\left(a_1^1\right) = \delta(T)$. Per Induktion sind wir damit fertig.

Falls aber $a_1^1|a_1^i$ für alle $i \geq 2$ gilt, so können wir durch elementare Zeilenumformungen in der ersten Spalte (außer ganz oben) lauter Nullen erzeugen.

(b) Wie gerade eben wenden wir nun Spaltenumformungen an und erhalten (außer ganz links) lauter Nullen in der ersten Zeile.

(c) Klar ist, dass für die neue Matrix $\mathrm{ggT}(T)|a_1^1$ gilt, wobei wir wieder die Bezeichnungen T und a_j^i verwenden. Ist $a_1^1 \not\sim \mathrm{ggT}(T)$, so gibt es ein Element a_q^p, $2 \leq p, q$, mit $a_1^1 \nmid a_q^p$. Addieren wir nun Spalte q zur Spalte 1, so können wir mit dem ersten Schritt die Matrix mit elementaren Umformungen so ändern, dass $\delta(T)$ strikt kleiner wird.

Der Beweis der zweiten Behauptung vereinfacht sich, wenn bereits irgendwo in der Matrix ein Element steht, das ein größter gemeinsamer Teiler von T ist; ebenso wenn solch ein Element leicht durch elementare Zeilen- und Spaltenumformungen erzeugt werden kann. Dann bringt man dieses nach links oben und addiert Vielfache der ersten Zeile/Spalte, so dass in der ersten Zeile und Spalte außer ganz links oben nur Nullen stehen. □

Bemerkung 4.4.5. Da $T \in A^{n\times m}$ durch endlich viele elementare Umformungen in die angegebene Gestalt gebracht werden kann gibt es $U \in Gl_n(A)$ und $V \in Gl_m(A)$, so dass UTV wie angegeben aussieht.

Bemerkung 4.4.6. Für den Beweis in dem Fall, dass A lediglich ein Hauptidealring und kein euklidischer Ring ist siehe [3].

Definition 4.4.7. Die Elemente $d_1, \ldots, d_r$ aus dem Elementarteilersatz heißen Elementarteiler von T.

Beispiel 4.4.8. Sei nun $A = \mathbb{Z}$. Wir wollen die Matrix

$$T = \begin{pmatrix} -17 & 14 & 39 \\ -8 & 6 & 21 \end{pmatrix}$$

in die Form aus dem Elementarteilersatz, also in Elementarteilerform, bringen.

Wir sind in der glücklichen Lage, dass $\mathrm{ggT}(T) = 1$ gilt und wir eine Linearkombination aus den Elementen der ersten Spalte $-17 + (-2) \cdot (-8) = -1$ leicht sehen. Wir addieren also (-2)-mal die zweite Zeile mit Hilfe von $\left(\begin{smallmatrix} 1 & -2 \\ 0 & 1 \end{smallmatrix}\right)$ zur ersten und Multiplizieren die gesamte Matrix mit -1.

$$\begin{pmatrix} -1 & 2 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} -17 & 14 & 39 \\ -8 & 6 & 21 \end{pmatrix} = \begin{pmatrix} 1 & -2 & 3 \\ 8 & -6 & -21 \end{pmatrix}.$$

Nun addieren wir die erste Spalte, multipliziert mit 2 bzw. (-3), zur zweiten und dritten Spalte und erhalten

$$\begin{pmatrix} -1 & 2 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} -17 & 14 & 39 \\ -8 & 6 & 21 \end{pmatrix} \begin{pmatrix} 1 & 2 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} 1 & -2 & 3 \\ 8 & -6 & -21 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 8 & 10 & -45 \end{pmatrix}.$$

Nun subtrahieren wir die erste Zeile (mit 8 multipliziert) von der zweiten. Der Zwischenschritt in der zweiten Zeile dient nur dazu, die Transformationsmatrizen auf der linken Seite von T zusammenzufassen.

$$\begin{pmatrix} 1 & 0 \\ -8 & 1 \end{pmatrix} \begin{pmatrix} -1 & 2 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} -17 & 14 & 39 \\ -8 & 6 & 21 \end{pmatrix} \begin{pmatrix} 1 & 2 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} -1 & 2 \\ 8 & -17 \end{pmatrix} \begin{pmatrix} -17 & 14 & 39 \\ -8 & 6 & 21 \end{pmatrix} \begin{pmatrix} 1 & 2 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} 1 & 0 \\ -8 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 8 & 10 & -45 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 10 & -45 \end{pmatrix}.$$

Wegen $\mathrm{ggT}(10, -45) = 5$ kombinieren wir die 5 nun, indem wir zunächst die zweite Spalte (mit 5 multipliziert) zur dritten addieren, dann die dritte Spalte von der zweiten (mit 2 multipliziert) subtrahieren und schließlich die zweite und die dritte Spalte vertauschen. Mit der Zwischenrechnung

$$\begin{aligned}
&\begin{pmatrix} 1 & 2 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \\
= &\begin{pmatrix} 1 & 2 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & -2 \end{pmatrix} \\
= &\begin{pmatrix} 1 & 2 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & -9 \\ 0 & 1 & -2 \end{pmatrix} \\
= &\begin{pmatrix} 1 & 7 & -12 \\ 0 & 5 & -9 \\ 0 & 1 & -2 \end{pmatrix}
\end{aligned}$$

erhalten wir

$$\begin{aligned}
&\begin{pmatrix} -1 & 2 \\ 8 & -17 \end{pmatrix} \begin{pmatrix} -17 & 14 & 39 \\ -8 & 6 & 21 \end{pmatrix} \begin{pmatrix} 1 & 7 & -12 \\ 0 & 5 & -9 \\ 0 & 1 & -2 \end{pmatrix} \\
= &\begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \end{pmatrix}.
\end{aligned}$$

Beachte, dass wir nicht behaupten, dass die invertierbaren Matrizen in $\mathbb{Z}^{2\times 2}$ bzw. $\mathbb{Z}^{3\times 3}$ eindeutig bestimmt sind. Es gilt nämlich auch

$$\begin{aligned}
&\begin{pmatrix} -1 & 2 \\ 2 & -3 \end{pmatrix} \begin{pmatrix} -17 & 14 & 39 \\ -8 & 6 & 21 \end{pmatrix} \begin{pmatrix} -13 & 5 & 12 \\ -10 & 4 & 9 \\ -2 & 1 & 2 \end{pmatrix} \\
= &\begin{pmatrix} 1 & -2 & 3 \\ -10 & 10 & 15 \end{pmatrix} \begin{pmatrix} -13 & 5 & 12 \\ -10 & 4 & 9 \\ -2 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \end{pmatrix}.
\end{aligned}$$

Man überzeugt sich leicht anhand der Determinante, dass die jeweils „äußeren“ Matrizen invertierbar sind, da sie eine Determinante in $\mathbb{Z}^* = \{\pm 1\}$ besitzen.

Bemerkung 4.4.9. Sei A ein euklidischer Ring. Seien $T \in A^{n\times m}$ und $w \in A^n$ gegeben. Wenn wir $x \in A^m$ suchen, das das in Matrixform als $Tx = w$ geschriebene lineare Gleichungssystem löst, so können wir zunächst (wie im Elementarteilersatz) über A invertierbare Matrizen $U \in A^{n\times n}$ und $V \in A^{m\times m}$ bestimmen, so dass

$$D := UTV = \left(\begin{array}{ccc|c} d_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & d_r & \\ \hline & 0 & & 0 \\ & & & \end{array}\right)$$

in Elementarteilerform ist. Dann ist $Tx = w$ äquivalent zu

$$Uw = UTx = UTVV^{-1}x = DV^{-1}x.$$

Setze $\tilde{w} := Uw$ und $\tilde{x} := V^{-1}x$. Dann ist $\tilde{w} = D\tilde{x}$ ein Gleichungssystem, das man genau dann lösen kann, wenn

(i) $d_i|\tilde{w}^i$ für alle Komponenten $\tilde{w}^i$ von $\tilde{w}$ mit $1 \leq i \leq r$ und
(ii) $\tilde{w}^i = 0$ für alle $i > r$

gelten. Ist das Gleichungssystem lösbar, so ist die allgemeine Lösung durch

$$\tilde{x} = \left(\frac{\tilde{w}^1}{d_1}, \ldots, \frac{\tilde{w}^r}{d_r}, u^{r+1}, \ldots, u^m\right)^T$$

für beliebige $u^{r+1}, \ldots, u^m \in A$ gegeben. Im Falle $r = m$ ist die Lösung eindeutig. Beachte schließlich noch, dass $x = V\tilde{x}$ gilt.

Beispiel 4.4.10. Sei $A = \mathbb{Z}$. Betrachte das lineare Gleichungssystem

$$\begin{array}{rcrcrcl} -17x^1 & + & 14x^2 & + & 39x^3 & = & w^1, \\ -8x^1 & + & 6x^2 & + & 21x^3 & = & w^2. \end{array}$$

Ohne weitere Definition verwenden wir nun die obige Notation.

Erinnerung: Es gilt

$$UTV = \begin{pmatrix} -1 & 2 \\ 8 & -17 \end{pmatrix} \begin{pmatrix} -17 & 14 & 39 \\ -8 & 6 & 21 \end{pmatrix} \begin{pmatrix} 1 & 7 & -12 \\ 0 & 5 & -9 \\ 0 & 1 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \end{pmatrix} = D.$$

Wir schreiben das lineare Gleichungssystem als

$$\underbrace{UTV}_{=D}\underbrace{V^{-1}x}_{=\tilde{x}} = \underbrace{Uw}_{=\tilde{w}}.$$

Wir möchten also

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \end{pmatrix} \begin{pmatrix} \tilde{x}^1 \\ \tilde{x}^2 \\ \tilde{x}^3 \end{pmatrix} = \begin{pmatrix} -w^1 + 2w^2 \\ 8w^1 - 17w^2 \end{pmatrix}$$

lösen. Dies ist genau dann lösbar, falls $5|(8w^1 - 17w^2)$ oder $8w^1 \equiv 17w^2 \pmod 5$ gilt. Die allgemeine Lösung lautet in diesem Fall

$$\tilde{x} = \begin{pmatrix} -w^1 + 2w^2 \\ \frac{1}{5}\left(8w^1 - 17w^2\right) \\ n \end{pmatrix}$$

bzw.

$$\begin{aligned} x = V\tilde{x} &= \begin{pmatrix} 1 & 7 & -12 \\ 0 & 5 & 9 \\ 0 & 1 & -2 \end{pmatrix} \begin{pmatrix} -w^1 + 2w^2 \\ \frac{1}{5}\left(8w^1 - 17w^2\right) \\ n \end{pmatrix} \\ &= \begin{pmatrix} -w^1 + 2w^2 + \frac{7}{5}\left(8w^1 - 17w^2\right) - 12n \\ 8w^1 - 17w^2 + 9n \\ \frac{1}{5}\left(8w^1 - 17w^2\right) - 2n \end{pmatrix}, \end{aligned}$$

wobei $n \in \mathbb{Z}$ beliebig ist und $\frac{1}{5}a$ als die Zahl in $\mathbb{Z}$ zu verstehen ist, die nach Multiplikation mit 5 gerade a ergibt.

Nachdem wir Moduln betrachtet haben werden werden wir noch einen Elementarteilersatz für Moduln kennen lernen.

5. Moduln

5.1. **Moduln.** Moduln über Ringen verallgemeinern den Begriff eines Vektorraumes über einem Körper.

Sei A in diesem Kapitel stets ein kommutativer Ring mit Eins.

Definition 5.1.1. Sei A ein kommutativer Ring mit Eins. Ein A-Modul ist eine abelsche Gruppe $(M,+)$ mit einer Abbildung $A \times M \to M$, $(a,x) \mapsto a \cdot x \equiv ax$ (Skalarmultiplikation), so dass die folgenden Eigenschaften für alle $a,b \in A$ und alle $x,y \in M$ gelten:

(i) $a(bx) = (ab)x$, (Assoziativität)
(ii) $(a+b)x = ax + bx$, (Distributivgesetze)
(iii) $a(x+y) = ax + ay$ und
(iv) $1x = x$.

Bemerkung 5.1.2.

(i) Sei A ein kommutativer Ring mit Eins. Dann ist A ein A-Modul.
(ii) Wie für Vektorräume zeigt man, dass für alle $a \in A$ und alle $x \in M$
   (a) $0 \cdot x = 0$,
   (b) $a \cdot 0 = 0$ und
   (c) $(-a)x = -(ax) = a(-x)$ gelten

   Beweis.
   (a) Es gilt $0 = 0 + 0$, somit folgt aus $0x = (0+0)x = 0x + 0x$ auch $0x = 0$.
   (b) Analog.
   (c) Es gelten $0 = ax - (ax)$, $0 = 0x = (a-a)x = ax + (-a)x$ sowie $0 = a(x - x) = ax + a(-x)$. Die rechten Seiten stimmen überein. Wir subtrahieren überall ax und erhalten die Behauptung. □

(iii) Sei $n \in \mathbb{N}$. Dann ist $M = A^n$ mit den komponentenweisen Verknüpfungen

$$(x^1, \ldots, x^n) + (y^1, \ldots, y^n) := (x^1 + y^1, \ldots, x^n + y^n)$$

und

$$a \cdot (x^1, \ldots, x^n) := (ax^1, \ldots, ax^n)$$

für $a, x^i, y^i \in A$, $1 \leq i \leq n$, ein A-Modul.
(iv) Ist K ein Körper, so stimmen die Definition eines K-Moduls und eines K-Vektorraumes überein.
(v) Sei M eine abelsche Gruppe. Dann wird M durch die folgende Definition zu einem $\mathbb{Z}$-Modul: Für $\mathbb{Z} \ni n \geq 0$ und $x \in M$ setzen wir $nx = \underbrace{x + \ldots + x}_{n \text{ Summanden}}$. Ist $\mathbb{Z} \ni n < 0$, so setzen wir $nx := -((-n)x)$.

   Es gibt nur eine Möglichkeit, eine abelsche Gruppe zu einem $\mathbb{Z}$-Modul zu machen: Da $1 \cdot x = x$ für $x \in M$ nach Definition eines Moduls eindeutig bestimmt ist erhalten wir induktiv, dass $nx = (1 + \ldots + 1)x = 1x + \ldots + 1x = x + \ldots + x$ ebenfalls eindeutig bestimmt ist. Wegen $nx + (-n)x = 0$ ist auch $(-n)x$ für $n \geq 0$ eindeutig bestimmt. Mit dieser Definition wird M tatsächlich ein $\mathbb{Z}$-Modul.

   Indem wir die Operation von $\mathbb{Z}$ auf der abelschen Gruppe eines Moduls vergessen ist jedes $\mathbb{Z}$-Modul natürlich auch eine abelsche Gruppe.

   Somit stimmen abelsche Gruppen und $\mathbb{Z}$-Moduln überein.

Beispiel 5.1.3. Sei K ein Körper, V ein K-Vektorraum und $f \in \mathrm{End}(V) \equiv \mathrm{End}_K(V)$ fixiert. Wir wollen V damit zu einem $K[t]$-Modul machen: Definiere für

$p \in K[t]$, $v \in V$ die Skalarmultiplikation durch $p \cdot v := p(f)(v)$, d. h. für $p = \sum_{i=0}^{d} a_i t^i$, $a_i \in K$, durch

$$\left(\sum_{i=0}^{d} a_i t^i\right) \cdot v := \sum_{i=0}^{d} a_i f^i(v).$$

Bezeichnung: Mit dieser Skalarmultiplikation schreiben wir für diesen $K[t]$-Modul V_f.

Wir behaupten, dass jeder $K[t]$-Modul von dieser Form ist: Sei nämlich M ein beliebiger $K[t]$-Modul. Wir können die Skalarmultiplikation $K[t] \times M \to M$ vermöge $K \subset K[t]$ auf $K \times M$ einschränken. Dann ist M mit der eingeschränkten Skalarmultiplikation ein Vektorraum. Definiere eine Abbildung $f\colon M \to M$ durch $f(v) := tv$ für beliebiges $v \in M$. Dies ist eine K-lineare Abbildung, also gilt $f \in \mathrm{End}_K(M)$. Betrachte nun M_f, den wie oben im Falle V_f definierten Modul. Dann ist M_f gerade der $K[t]$-Modul M.

Daher können wir alternativ $K[t]$-Moduln oder Paare (V, f) mit einem K-Vektorraum V und $f \in \mathrm{End}_K(V)$ betrachten. Wir untersuchen damit letztlich das Gleiche.

Definition 5.1.4. Seien M, N zwei A-Moduln. Dann heißt eine Abbildung

$$f\colon M \to N$$

eine $(A-)$lineare Abbildung oder ein Homomorphismus von A-Moduln, falls

(i) $f(x+y) = f(x) + f(y)$ und
(ii) $f(ax) = af(x)$

für alle $a \in A$ und alle $x, y \in M$ gelten.

Ein Homomorphismus f von A-Moduln heißt Isomorphismus, falls f bijektiv ist. (Ist f ein Isomorphismus, so ist f^{-1} ebenfalls linear.)

Zwei A-Moduln M und N heißen isomorph, falls es einen Isomorphismus $f\colon M \to N$ gibt.

Beispiel 5.1.5. Sei $T \in A^{m \times n}$. Dann ist die Abbildung $f\colon A^n \to A^m$ mit $x \mapsto Tx$ linear.

Ist umgekehrt $f\colon A^n \to A^m$ linear, so gibt es ein $T \in A^{m \times n}$ mit $f(x) = Tx$. Dies zeigt man analog zur Aussage für Vektorräume ausgehend von den Bildern $f(0, \ldots, 0, 1, 0, \ldots, 0) \in A^m$ für alle Positionen der Eins.

Der folgende Satz charakterisiert die $K[t]$-linearen Abbildungen zwischen $K[t]$-Moduln.

Theorem 5.1.6. *Sei K ein Körper. Seien (V, f), (W, g) Paare von K-Vektorräumen und Endomorphismen $f \in \mathrm{End}_K(V)$ und $g \in \mathrm{End}_K(W)$. Dann sind die $K[t]$-linearen Abbildungen $\varphi\colon V_f \to W_g$ genau die K-linearen Abbildungen $\varphi\colon V \to W$ mit $\varphi \circ f = g \circ \varphi$.*

Beweis.
„$\Longrightarrow$“: Sei $\varphi\colon V_f \to W_g$ eine $K[t]$-lineare Abbildung. Dann ist $\varphi\colon V \to W$ insbesondere eine K-lineare Abbildung. (Beachte, dass die dem Vektorraum V und dem Modul V_f zugrunde liegenden Mengen übereinstimmen.) Für $v \in V$ gilt

$$\varphi(f(v)) = \varphi(tv) = t\varphi(v) = g(\varphi(v))$$

aufgrund der Definition von V_f bzw. W_g und der $K[t]$-Linearität. Somit folgt $\varphi \circ f = g \circ \varphi$.

„$\Longleftarrow$“: Sei nun $\varphi\colon V \to W$ eine K-lineare Abbildung mit $\varphi \circ f = g \circ \varphi$. Somit folgt nach Definition von V_f und W_g, dass $\varphi(tv) = t\varphi(v)$ für alle $v \in V$ gilt. Wir

erhalten per Induktion $\varphi \circ f^n = g^n \circ \varphi$ und somit $\varphi(t^n v) = t^n \varphi(v)$ für alle $n \geq 0$. Somit ist φ eine $K[t]$-lineare Abbildung. □

Statt Matrizen in Ähnlichkeitsklassen einzuteilen können wir auch $K[t]$-Moduln in Isomorphieklassen einteilen. Dies ist im folgenden Sinne äquivalent:

Theorem 5.1.7. *Sei K ein Körper, $V = K^n$. Seien $A, B \in K^{n\times n}$ und V_A, V_B die $K[t]$-Moduln zu den Paaren $(V, x \mapsto Ax)$ und $(V, x \mapsto Bx)$. Dann sind die folgenden beiden Aussagen äquivalent:*

(i) A und B sind ähnliche Matrizen.
(ii) V_A und V_B sind isomorphe $K[t]$-Moduln.

Beweis. In Theorem 5.1.6 haben wir gesehen, dass $K[t]$-lineare Abbildungen durch K-lineare Abbildungen gegeben sind, so dass $\varphi \circ f = g \circ \varphi$ gilt. Stellen wir K-lineare Abbildungen über V als Matrizen dar, so sind $K[t]$-lineare Abbildungen $V_A \to V_B$ durch Matrizen $S \in \mathrm{End}_K(V)$ mit $BS = SA$ gegeben. Somit entsprechen $K[t]$-lineare Isomorphismen $V_A \to V_B$ gerade Matrizen $S \in Gl_n(K)$ mit $BS = SA$. Also gilt $B = SAS^{-1}$. Somit sind A und B ähnliche Matrizen. Die Argumentation funktioniert in beide Richtungen. Somit folgt die Behauptung. □

Definition 5.1.8 (Untermodul)**.** Sei M ein A-Modul. Dann heißt eine Teilmenge $U \subset M$ ein ($A-$)Untermodul von M, wenn U eine Untergruppe der abelschen Gruppe $(M, +)$ ist, die unter Skalarmultiplikation abgeschlossen ist: Für alle $a \in A$ und $u \in U$ gilt $au \in U$.

Beispiel 5.1.9.

(i) Sei $M = A$ ein kommutativer Ring, aufgefasst als A-Modul. Dann sind die Untermodule von M genau die Ideale von A.
(ii) Sei M eine abelsche Gruppe. Betrachte M als $\mathbb{Z}$-Modul. Dann sind die Untermoduln von M genau die Untergruppen von M als abelsche Gruppe.
(iii) Sei V ein K-Vektorraum und $f \in \mathrm{End}_K(V)$. Dann sind die Untermoduln des $K[t]$-Moduls V_f genau die f-invarianten Unterräume U von V, also die Untervektorräume $U \subset V$ mit $f(U) \subset U$. Solch ein Untermodul ist (isomorph zu) $U_{f|_U}$. (Leichte Übung)
(iv) Sei U ein Untermodul eines A-Moduls M. Dann ist U mit den auf U induzierten Verknüpfungen selbst wieder ein A-Modul.
(v) Seien M, N zwei A-Moduln. Sei $f\colon M \to N$ eine A-lineare Abbildung. Dann ist

$$\ker f = \{x \in M\colon f(x) = 0\}$$

ein Untermodul von M und

$$\mathrm{im}\, f = \{f(x)\colon x \in M\}$$

ein Untermodul von N.
Beweis: Wie bei Vektorräumen.
(vi) Seien $(U_\lambda)_{\lambda\in\Lambda}$, Λ eine beliebige Indexmenge, Untermoduln von M. Dann sind auch $\bigcap_{\lambda\in\Lambda} U_\lambda$ und

$$\sum_{\lambda\in\Lambda} U_\lambda := \left\{\sum_{\lambda\in\Lambda} u_\lambda\colon u_\lambda \in U_\lambda \text{ mit } u_\lambda = 0 \text{ für fast alle } \lambda \in \Lambda\right\}$$

Untermoduln von M.
Beweis: Standard.

Definition 5.1.10 (Quotientenmodul)**.** Sei M ein A-Modul und $U \subset M$ ein Untermodul von M. Dann definieren wir den Quotientenmodul (oder Faktormodul) M/U durch die Äquivalenzrelation

$$x \equiv y \pmod{U} \quad :\Longleftrightarrow \quad x - y \in U.$$

(Da U ein Untermodul ist sieht man leicht, dass dies eine Äquivalenzrelation ist.) Wir bezeichnen die Äquivalenzklasse von $x \in M$ mit $x+U = \{x+u \colon u \in U\}$. Setze nun $M/U := \{x + U \colon x \in M\}$. M/U besteht also aus allen Äquivalenzklassen der obigen Äquivalenzrelation.

Auf M/U definieren wir Verknüpfungen durch

$$(x + U) + (y + U) := (x + y) + U$$

und

$$a(x + U) := (ax) + U$$

für $x, y \in M$ und $a \in A$. Da U ein Untermodul von M ist, sind diese Verknüpfungen wohldefiniert.

Bemerkung 5.1.11. Sei M ein A-Modul und $U \subset M$ ein Untermodul von M. Dann ist die Abbildung

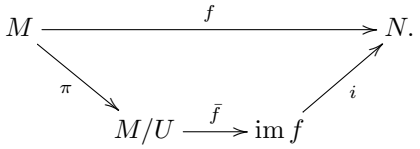
$$\begin{aligned} \pi \colon M &\to M/U, \\ x &\mapsto x + U \end{aligned}$$

A-linear und surjektiv.

Beweis: Einfach.

Es gilt der folgende Homomorphiesatz:

Theorem 5.1.12 (Homomorphiesatz für Moduln)**.** *Sei A ein kommutativer Ring. Seien M, N zwei A-Moduln. Sei $f \colon M \to N$ eine A-lineare Abbildung. Sei $U \subset M$ ein A-Untermodul. Genau dann, wenn $U \subset \ker f$ gilt, gibt es eine A-lineare Abbildung $\bar{f} \colon M/U \to \operatorname{im} f$ mit $f = i \circ \bar{f} \circ \pi$, wobei $\pi \colon M \to M/U$ die Projektion $x \mapsto x + U$ und $i \colon \operatorname{im} f \to N$ die Inklusionsabbildung ist. Existiert solch eine Abbildung $\bar{f}$, so ist sie eindeutig bestimmt und es gilt $\ker \bar{f} = (\ker f)/U$ sowie $i(\operatorname{im} \bar{f}) = \operatorname{im} f$. Existiert $\bar{f}$, so kommutiert das Diagramm*



Ist $U = \ker f$, so ist $\bar{f}$ ein Isomorphismus. Ist f surjektiv, so gilt $N \cong M/\ker f$.

Beweis. Analog zum Homomorphiesatz für Vektorräume oder Ringe. □

Korollar 5.1.13. *Sei A ein kommutativer Ring. Sei M ein A-Modul. Seien U, V Untermoduln von M. Dann gilt*

$$U/(U \cap V) \cong (U + V)/V.$$

Der Isomorphismus ist durch $U/(U \cap V) \ni u + (U \cap V) \mapsto u + V \in (U + V)/V$ und die Inverse durch $(U + V)/V \ni u + v + V \mapsto u + (U \cap V) \in U/(U \cap V)$ für $u \in U$ und $v \in V$ gegeben.

Beweis. Definiere $\varphi\colon U+V \to U/(U\cap V)$ durch $u+v \mapsto u+(U\cap V)$ für $u \in U$ und $v \in V$. Die Abbildung φ ist wohldefiniert, denn gelte $u_1+v_1 = u_2+v_2$ mit $u_i \in U$ und $v_i \in V$, so folgt $u_1-u_2 = v_2-v_1 \in U\cap V$. φ ist offensichtlicherweise surjektiv und es gilt $V \subset \ker\varphi$. Sei umgekehrt $u+v \in \ker\varphi$. Dann folgt $0+(U\cap V) = u+(U\cap V)$, also $u \in U\cap V \subset V$ und somit $u+v \in V$. Es gilt also $V = \ker\varphi$. Der Homomorphiesatz für Moduln liefert also

$$(U+V)/V \cong U/(U\cap V)$$

und die Gestalt des Isomorphismus. Die zweite angegebene Abbildung ist wohldefiniert und eine Umkehrabbildung. □

Definition 5.1.14. Sei M ein A-Modul und seien $x_1,\dots,x_n \in M$.

(i) Dann heißt $(x_1,\dots,x_n)$ ein Erzeugendensystem von M, falls $M = Ax_1+\ldots+Ax_n$ gilt.

(ii) $(x_1,\dots,x_n)$ heißt linear abhängig, falls Ringelemente $a^1,\dots,a^n \in A$ mit $\sum_{i=1}^{n} a^i x_i = 0$ und $a^i \neq 0$ für ein $i \in \{1,\dots,n\}$ existieren.
Andernfalls lässt sich $0 \in M$ nur auf triviale Art und Weise kombinieren, d. h. aus $\sum_{i=1}^{n} a^i x_i = 0$ für $a^i \in A$, $i \in \{1,\dots,n\}$, folgt bereits $a^i = 0$ für alle $i \in \{1,\dots,n\}$. Dann heißt $(x_1,\dots,x_n)$ linear unabhängig.

(iii) $(x_1,\dots,x_n)$ heißt Basis von M, falls $(x_1,\dots,x_n)$ linear unabhängig und ein Erzeugendensystem von M ist.

(iv) Sei $\mathcal{F} := (x_\lambda)_{\lambda\in\Lambda}$ eine beliebige Familie mit $x_\lambda \in M$ für alle $\lambda \in \Lambda$. Ist Λ eine endliche Menge, so stimmen die folgenden Definitionen mit den obigen Definitionen überein:
   (a) $\mathcal{F}$ heißt Erzeugendensystem, falls für jedes $x \in M$ Ringelemente $a^\lambda \in A$ mit $\sum_{\lambda\in\Lambda} a^\lambda x_\lambda = x$ existieren, wobei $a^\lambda = 0$ für fast alle $\lambda \in \Lambda$ gilt.
   (b) $\mathcal{F}$ heißt linear unabhängig, falls jede endliche Teilfamilie linear unabhängig ist. Sonst heißt $\mathcal{F}$ linear abhängig.
   (c) $\mathcal{F}$ heißt Basis von M, falls $\mathcal{F}$ ein linear unabhängiges Erzeugendensystem von M ist.

Bemerkung 5.1.15.

(i) Im allgemeinen gibt es keine nichtleeren linear unabhängigen Familien in einem Modul. Beispiel: Seien $A = \mathbb{Z}$ und $M = \mathbb{Z}/n\mathbb{Z}$ für ein $n \in \mathbb{N}\setminus\{0\}$. Dann gibt es keine linear unabhängige Teilmenge von M, da für beliebige $a \in M$ stets $na = 0$ gilt.

(ii) Anders als in Vektorräumen braucht ein minimales Erzeugendensystem keine Basis zu sein. Beispiel: Seien wieder $A = \mathbb{Z}$ und $M = \mathbb{Z}/n\mathbb{Z}$ für ein $n \in \mathbb{N}\setminus\{0,\pm1\}$. Ein Erzeugendensystem enthält mindestens ein Element ungleich Null. Damit ist es aber nicht linear unabhängig wie wir gerade gesehen haben.

(iii) Anders als in Vektorräumen braucht auch eine maximal linear unabhängige Menge keine Basis zu sein. Beispiel: Seien A und M wieder wie oben. Dann ist nur die leere Menge linear unabhängig. Somit ist keine linear unabhängige Menge ein Erzeugendensystem.

(iv) Sei $M = A^n$, $n \in \mathbb{N}$. Dann ist $(e_1,\dots,e_n)$ eine Basis des A-Moduls A^n, wobei $e_i = (0,\dots,0,1,0,\dots,0)$ mit dem Einser an der i-ten Stelle gilt.

Lemma 5.1.16. *Sei A ein kommutativer Ring. Sei M ein A-Modul mit Basis $(x_1,\dots,x_n)$. Dann ist die Abbildung $\varphi\colon A^n \to M$ mit $(a^1,\dots,a^n) \mapsto \sum_{i=1}^{n} a^i x_i$ ein Isomorphismus von A-Moduln.*

Beweis. Klar ist, dass φ eine A-lineare Abbildung ist. φ ist surjektiv, da $(x_1, \ldots, x_n)$ ein Erzeugendensystem ist. φ ist injektiv, da $(x_1, \ldots, x_n)$ eine linear unabhängige Familie ist. □

Definition 5.1.17. Sei A ein kommutativer Ring und M ein A-Modul.

(i) Der Modul M heißt endlich erzeugt, falls M ein endliches Erzeugendensystem besitzt.
(ii) Der Modul M heißt frei, falls M eine Basis besitzt.

Theorem 5.1.18. *Sei A ein kommutativer Ring und M ein endlich erzeugter A-Modul.*

(i) M ist genau dann frei, wenn $M \cong A^n$ für ein $n \in \mathbb{N}$ ist.
(ii) Ist M ein freier Modul, so haben alle Basen dieselbe Mächtigkeit (oder Länge). Wir definieren den Rang von M, $\operatorname{rk}(M)$, als die Anzahl der Elemente einer Basis. Die obige Aussage zeigt, dass $\operatorname{rk}(M)$ für endlich erzeugte A-Moduln wohldefiniert ist.

Beweis.

(i) Eine Richtung folgt aus Lemma 5.1.16, die andere Richtung ist klar.
(ii) Im Spezialfall, dass A ein Körper ist, ist M ein Vektorraum und die Aussage ist bekannt. Ist $A = \{0\}$, so ist $M = \{0\}$ und die Aussage trivial. Wir nehmen daher ab jetzt an, dass $A \neq \{0\}$ gilt.

Sei sei $I \subsetneq A$ ein beliebiges maximales Ideal. Dann ist A/I ein Körper. Wir dürfen aufgrund des ersten Teiles annehmen, dass $M = A^n$ für ein $n \in \mathbb{N}$ gilt. Da I ein Ideal von A ist, ist $U := IM \subset AM = M$ ein Untermodul von $M = A^n$. Es folgt $M/U = A^n/(IM) = A^n/I^n = (A/I)^n$. Wir erhalten $n = \dim_{A/I} M/U$. Wäre $M \cong A^m \cong A^n$ für ein $m \neq n$ so folgte, dass $A^n/(I(A^n))$ und $A^m/(I(A^m))$ über A/I isomorphe Vektorräume wären. Dies ist jedoch unmöglich, da sie verschiedene Dimensionen haben. □

Wie für Vektorräume definieren wir auch für Moduln direkte Summen.

Definition 5.1.19 (Externe direkte Summe von Moduln)**.** Seien $M_1, \ldots, M_n$ jeweils A-Moduln über einem kommutativen Ring A. Wir definieren die äußere oder externe direkte Summe $M_1 \oplus \ldots \oplus M_n$ als das kartesische Produkt $M_1 \times \ldots \times M_n$ mit komponentenweiser Addition und Skalarmultiplikation.

Ist die gegebene Familie von A-Moduln unendlich, so unterscheiden wir zwischen der direkten Summe und dem direkten Produkt.

Definition 5.1.20. Sei A ein kommutativer Ring mit Eins. Sei $(M_\lambda)_{\lambda \in \Lambda}$, Λ eine beliebige Indexmenge, eine Familie von A-Moduln. Dann sind die direkte Summe $\bigoplus_{\lambda \in \Lambda} M_\lambda$ und das direkte Produkt $\prod_{\lambda \in \Lambda} M_\lambda$ der Familie $(M_\lambda)_{\lambda \in \Lambda}$ Teilmengen des kartesischen Produktes $\prod_{\lambda \in \Lambda} M_\lambda$ mit komponentenweiser Addition und Skalarmultiplikation. Die Elemente haben die Form $(x_\lambda)_{\lambda \in \Lambda}$ mit $x_\lambda \in M_\lambda$ für alle $\lambda \in \Lambda$. Für das direkte Produkt nehmen wir als Menge das gesamte kartesische Produkt, für die äußere direkte Summe nur die Elemente $(x_\lambda)_{\lambda \in \Lambda}$ mit $x_\lambda \in M_\lambda$ und $x_\lambda = 0$ für fast alle $\lambda \in \Lambda$.

Definition 5.1.21 (Interne direkte Summe von Moduln)**.** Sei A ein kommutativer Ring und $U_1, \ldots, U_n \subset M$ Untermoduln eines A-Moduls M. Definiere $U := \sum_{i=1}^{n} U_i = \{u_1 + \ldots + u_n \colon u_i \in U_1 \text{ für } i \in \{1, \ldots, n\}\}$.

Wir definieren die kanonische A-lineare Abbildung

$$\varphi\colon \bigoplus_{i=1}^{n} U_i \to U \quad \text{durch} \quad (u_1, \ldots, u_n) \mapsto \sum_{i=1}^{n} u_i.$$

φ ist genau dann ein Isomorphismus, wenn aus $\sum_{i=1}^{n} u_i = 0$ mit $u_i \in U_i$ für $i \in \{1, \ldots, n\}$ bereits $u_1 = u_2 = \ldots = u_n = 0$ folgt. Ist φ ein Isomorphismus, so sagen wir, dass die interne Summe $U = \sum_{i=1}^{n} U_i$ direkt ist. (Betrachten wir Moduln über Körpern, also Vektorräume, so ist eine interne Summe von Vektorräumen genau dann direkt, wenn dies für die interne Summe als Summe von Untermoduln gilt.)

Definition 5.1.22. Sei A ein Ring (kommutativ, mit Eins). Dann heißt ein A-Modul M zyklisch, wenn $M = Ax$ für ein $x \in M$ gilt.

Zyklische Moduln sind isomorph zu Quotienten von A.

Theorem 5.1.23. *Sei A ein kommutativer Ring mit Eins. Sei M ein A-Modul. Dann ist M genau dann zyklisch, wenn es ein Ideal I von A mit $M \cong A/I$ gibt. Hier bezeichnet „$\cong$“ Isomorphie von A-Moduln.*

Beweis.
„$\Longrightarrow$“: Sei M zyklisch, gelte also $M = Ax$ für ein $x \in M$. Dann ist $\varphi\colon A \to M$ mit $\varphi(a) := ax$ für $a \in A$ surjektiv und A-linear. Definiere $I := \ker \varphi$. Dann ist I ein Ideal von A. Aufgrund des Homomorphiesatzes erhalten wir einen Isomorphismus von A-Moduln:

$$\bar{\varphi}\colon A/I \xrightarrow{\sim} M \quad \text{mit} \quad \bar{\varphi}(a + I) := ax \quad \text{für } a \in A.$$

Also ist $M \cong A/I$.

„$\Longleftarrow$“: Gelte $M \cong A/I$. Dann ist $A \ni a \mapsto a \cdot (1 + I)$ surjektiv. Somit ist M zyklisch. □

Bemerkung 5.1.24. Das Ideal I im Beweis von Theorem 5.1.23 hängt nicht von der speziellen Wahl von x ab. Gelte nämlich $M = Ax = Ay$ für ein $y \in M$. Dann gilt $\alpha x = y$ für ein $\alpha \in A$. Definiere $\psi\colon A \to M$ durch $\psi(a) := ay$ für $a \in A$ und $J := \ker \psi$. Dann gilt $I = J$: Sei $a \in \ker \varphi$, also $ax = 0$. Dann folgt $0 = \alpha ax = a(\alpha x) = ay$. Also ist $a \in \ker \psi$. Aus Symmetriegründen erhalten wir $\ker \varphi = \ker \psi$.

Wie in diesem Beweis erhält man auch $I = \{a \in A\colon ax = 0 \text{ für alle } x \in M\}$. Somit ist I eindeutig bestimmt. Wir definieren den Annulator von M durch $\mathrm{Ann}(M) := \{a \in A\colon ax = 0 \text{ für alle } x \in M\}$.

Lemma 5.1.25. *Sei A ein Ring (kommutativ, mit Eins). Sei M ein zyklischer A-Modul. Dann gelten:*

(i) Sei U ein Untermodul von M. Dann ist der Faktormodul M/U von M selbst wieder zyklisch.
(ii) Ist A ein Hauptidealring, so ist auch jeder Untermodul von M zyklisch.

Beweis.
(i) Ist $M = Ax$, so folgt $M/U = A(x + U)$.
(ii) Nach Theorem 5.1.23 dürfen wir annehmen, dass $M = A/I$ für ein Ideal I von A gilt. Nach Definition hat jeder Untermodul von M die Form $U = J/I$ für ein Ideal J von A mit $I \subset J$ (U besitzt die Darstellung J/I für eine Menge J mit $I \subset J$. Man rechnet mit Hilfe der Definition eines Untermoduls nach, dass J ein Ideal ist.). Da A ein Hauptidealring ist, ist J ein Hauptideal, also von einem Element erzeugt. Somit sind J und daher auch U zyklisch. □

Theorem 5.1.26. *Sei A ein Hauptidealring und M ein endlich erzeugter A-Modul. Sei $U \subset M$ ein Untermodul. Dann ist U selbst ebenfalls endlich erzeugt.*

Beweis. Gelte $M = Ax_1 + \ldots + Ax_n$ für $x_i \in M$. Wir zeigen per Induktion nach n, dass auch U von n Elementen erzeugt wird.

Ist $n = 1$, so ist M zyklisch und die Behauptung folgt aus Lemma 5.1.25.

„$n-1 \rightsquigarrow n$“: Definiere $N := Ax_1 + \ldots + Ax_{n-1}$. Dann ist $U \cap N$ ein Untermodul von N. Daher wird $U \cap N$ nach Induktionsvoraussetzung von $n-1$ Elementen, $y_1, \ldots, y_{n-1}$, erzeugt. Nach dem Homomorphiesatz für Moduln, genauer: Korollar 5.1.13, folgt $U/(U \cap N) \cong (U+N)/N$. $(U+N)/N$ ist ein Untermodul von M/N. Es gilt $M/N = A(x_n + N)$. Also ist M/N zyklisch. Da A ein Hauptidealring ist, ist auch $(U+N)/N \cong U/(U \cap N)$ nach Lemma 5.1.25 zyklisch. Es gibt also ein $z \in U$ mit $U/(U \cap N) = A(z + (U \cap N))$. Wir erhalten insgesamt $U = Ay_1 + \ldots + Ay_n + Az$. □

5.2. **Der Elementarteilersatz für Moduln.** Sei A in diesem Abschnitt stets ein euklidischer Ring (da wir den Elementarteilersatz für Matrizen für Hauptidealringe nicht bewiesen haben).

Den Elementarteilersatz für Matrizen können wir für freie Moduln wie folgt umformulieren:

Theorem 5.2.1. *Seien M, N endlich erzeugte freie A-Moduln. Sei $f\colon M \to N$ eine A-lineare Abbildung. Dann gibt es Basen $(x_1, \ldots, x_m)$ von M und $(y_1, \ldots, y_n)$ von N und r mit $0 \le r \le \min\{m, n\}$ sowie $d_i \in A$, $1 \le i \le r$, mit $d_1 | d_2 | \ldots | d_r$ und $d_i \neq 0$, so dass*

$$f(x_i) = \begin{cases} d_i y_i & \text{für } i = 1, \ldots, r, \\ 0 & \text{sonst.} \end{cases}$$

Beweis. Da M, N endlich erzeugt und frei sind, dürfen wir ohne Einschränkung $M = A^m$ und $N = A^n$ annehmen. Gelte $f(x) = Tx$ für eine Matrix $T \in A^{n \times m}$. Aufgrund des Elementarteilersatzes für Matrizen, Theorem 4.4.4, gibt es Matrizen $U \in Gl_n(A)$ und $V \in Gl_m(A)$ mit

$$UTV = \left(\begin{array}{ccc|c} d_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & d_r & \\ \hline & 0 & & 0 \\ & & & \end{array}\right)$$

und $d_i \neq 0$ sowie $d_1 | d_2 | \ldots | d_r$.

Sei e_j, $1 \le j \le m$, die Standardbasis von A^m und sei e'_i, $1 \le i \le n$, die Standardbasis von A^n. Dann folgt aus der Gestalt von UTV

$$UTVe_j = \begin{cases} d_j e'_j & \text{für } j = 1, \ldots, r, \\ 0 & \text{sonst.} \end{cases}$$

Definiere nun $x_j := Ve_j$, $j = 1, \ldots, m$ sowie $y_i := U^{-1} e'_i$ für $i = 1, \ldots, n$. Für $j = 1, \ldots, r$ gilt $f(x_j) = TVe_j = U^{-1}UTVe_j = U^{-1} d_j e'_j = d_j y_j$ und für $j = r+1, \ldots, m$ erhalten wir $f(x_j) = TVe_j = 0$. □

Korollar 5.2.2. *Sei M ein endlich erzeugter freier A-Modul. Sei U ein Untermodul von M. Dann gilt:*

(i) Es gibt eine Basis $(x_1, \ldots, x_m)$ von M sowie r mit $0 \le r \le m$ und $d_i \in A$ mit $d_i \neq 0$ und $d_1 | d_2 | \ldots | d_r$, so dass $(d_1 x_1, \ldots, d_r x_r)$ eine Basis von U ist.

(ii) Insbesondere ist also U als A-Modul selbst frei und es gilt $\mathrm{rk}(U) \le \mathrm{rk}(M)$.

Beweis. Es genügt offensichtlicherweise, den ersten Teil zu zeigen.

Nach Theorem 5.1.26 ist auch U endlich erzeugt. Sei $u_1, \dots, u_m$ ein Erzeugendensystem von U. Definiere die A-lineare Abbildung $f\colon A^m \to U$ durch

$$f(a^1, \dots, a^m) := \sum_{i=1}^{m} a^i u_i.$$

Nach Voraussetzung gilt $\operatorname{im} f = U$. Nach Theorem 5.2.1 gibt es Basen $(z_1, \dots, z_m)$ von A^m und $(x_1, \dots, x_n)$ von M mit

$$f(z_i) = \begin{cases} d_i x_i & \text{für } 1 \le i \le r, \\ 0 & \text{sonst} \end{cases}$$

mit $d_1, \dots, d_r \neq 0$. Beschreiben wir f mit diesen Basen, so ist klar, dass auch $d_1 x_1, \dots, d_r x_r$ ein Erzeugendensystem für $\operatorname{im} f = U$ ist. Nach Definition der linearen Unabhängigkeit ist $\{d_1 x_1, \dots, d_r x_r\}$ linear unabhängig, da dies für $x_1, \dots, x_r$ in M gilt. Somit ist $(d_1 x_1, \dots, d_r x_r)$ eine Basis von U. □

Beispiel 5.2.3. Sei $A = \mathbb{Z}$. Identifiziere $T \in \mathbb{Z}^{2\times 3}$ mit einer Abbildung $T\colon \mathbb{Z}^3 \to \mathbb{Z}^2$ vermöge $T(x) := Tx$. Betrachte wieder das Beispiel

$$UTV = \begin{pmatrix} -1 & 2 \\ 8 & -17 \end{pmatrix} \begin{pmatrix} -17 & 14 & 39 \\ -8 & 6 & 21 \end{pmatrix} \begin{pmatrix} 1 & 7 & -12 \\ 0 & 5 & -9 \\ 0 & 1 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \end{pmatrix} = D.$$

Der Untermodul $\operatorname{im} T \subset \mathbb{Z}^2$ ist nach Korollar 5.2.2 ein freier Modul. Wegen $V \in Gl_3(\mathbb{Z})$ gilt $\operatorname{im} D = \operatorname{im} UT$. Eine Basis von $\operatorname{im} D$ ist $(1,0)^T$, $(0,5)^T$. Eine Basis von $\operatorname{im} T$ erhalten wir also, wenn wir $U^{-1} = \left(\begin{smallmatrix} -17 & -2 \\ -8 & -1 \end{smallmatrix}\right)$ auf die Basis anwenden. Wir erhalten also als eine Basis von $\operatorname{im} T$ die Vektoren $(17,8)^T$ und $5 \cdot (2,1)^T$. $\operatorname{im} T$ ist ein freier Untermodul von $\mathbb{Z}^2$ vom Rang 2.

Beachte, dass $\operatorname{im} T \subsetneq \mathbb{Z}^2$ gilt, aber $\operatorname{rk} \operatorname{im} T = \operatorname{rk} \mathbb{Z}^2 = 2$ gilt. Dies wäre über einem Körper, d. h. für Vektorräume, nicht möglich.

Wir wollen nun die Struktur von endlich erzeugten Moduln genauer untersuchen.

Theorem 5.2.4 (Elementarteilersatz für Moduln)**.**
Sei M ein endlich erzeugter A-Modul. Dann gibt es $r, n \ge 0$ und $d_i \in A \setminus (\{0\} \cup A^*)$ mit $d_1 | d_2 | \dots | d_r$ und $d_i \neq 0$ so dass

$$M \cong A^n \oplus A/(d_1) \oplus \dots \oplus A/(d_r).$$

Insbesondere ist also M die direkte Summe von endlich vielen zyklischen A-Moduln. Die Elemente $d_1, \dots, d_r$ heißen Elementarteiler von M, n der (torsionsfreie) Rang von M.

Wir werden später sehen, dass die Elementarteiler bis auf Einheiten eindeutig bestimmt sind.

Beweis. Sei $x_1, \dots, x_m$ ein Erzeugendensystem von M. Definiere $f\colon A^m \to M$ durch

$$f(a) := \sum_{i=1}^{m} a^i x_i.$$

Setze $U := \ker f$. Dann gilt nach Theorem 5.1.12 (Homomorphiesatz für Moduln) $M \cong A^m/U$. Nach Korollar 5.2.2 gibt es eine Basis $(y_1, \dots, y_m)$ von A^m und $0 \le r \le m$ mit $d_1 | \dots | d_r$, $d_i \neq 0$, so dass $(d_1 y_1, \dots, d_r y_r)$ eine Basis von U ist. Somit gilt

$$U = A d_1 y_1 \oplus \dots \oplus A d_r y_r.$$

Wir erhalten

$M \cong A^m/U$

$$\cong (Ay_1 \oplus \ldots \oplus Ay_r \oplus \ldots \oplus Ay_m)/(Ad_1y_1 \oplus \ldots \oplus Ad_ry_r \oplus \{0\} \oplus \ldots \oplus \{0\})$$
$$\cong A/(d_1) \oplus \ldots \oplus A/(d_r) \oplus A^{m-r}.$$

(Wenn keine Verwechslungsgefahr besteht schreibt man solche Quotienten auch manchmal mit waagrechtem „Bruchstrich“.) Sämtliche Summanden $A/(d_i)$ in denen d_i eine Einheit ist lassen wir weg und erhalten die Behauptung. □

Bemerkung 5.2.5. Elementarteiler von Matrizen und Elementarteiler von Moduln hängen wie folgt zusammen: Sei $T \in A^{m\times n}$ und seien $d_1|\ldots|d_r$ mit $d_i \neq 0$ die Elementarteiler von T. Sei $1 \leq s \leq r$ so gewählt, dass $d_s \sim 1$, aber $d_{s+1} \not\sim 1$ gelten. Vermöge $x \mapsto Tx$ induziert T eine Abbildung $A^n \to A^m$, die wir wieder mit T bezeichnen. Definiere einen A-Modul durch $M := A^m/\operatorname{im} T$. Dann liest man aus der Elementarteilerform von T ab, dass

$$M \cong A^{m-r} \oplus A/(d_{s+1}) \oplus \ldots \oplus A/(d_r)$$

gilt.

Damit sind die Elementarteiler von M gerade die Elementarteiler von T, die keine Einheiten sind. Sei umgekehrt M wie oben definiert gegeben. Sei t der torsionsfreie Rang von M. Setze $r := m - t$. Dann besitzt T genau $r = m - t$ Elementarteiler, nämlich die Elementarteiler von M und Einheiten. Diese Elementarteiler sind aufgrund der Bemerkung nach dem Elementarteilersatz für Moduln eindeutig bestimmt.

Somit sind die Elementarteiler von T durch die von M eindeutig bestimmt und umgekehrt.

Wir arbeiten nun auf den Eindeutigkeitsbeweis für die Elementarteiler hin.

Definition 5.2.6. Sei M ein A-Modul.

(i) Dann heißt $x \in M$ ein Torsionselement, falls es ein $0 \neq a \in A$ mit $ax = 0$ gibt. Für die Menge aller Torsionselemente von M schreiben wir M_{tors}.
(ii) Der Modul M heißt torsionsfrei, wenn $M_{\text{tors}} = \{0\}$ gilt. M heißt Torsionsmodul, falls $M_{\text{tors}} = M$ ist.

Bemerkung 5.2.7.

(i) M_{tors} ist ein Untermodul von M: Klar ist, dass M_{tors} unter Skalarmultiplikation mit Elementen aus dem Ring A abgeschlossen ist. Seien $x, y \in M_{\text{tors}}$ und $a \neq 0 \neq b$ mit $ax = by = 0$. Dann folgt $ab(x + y) = 0$ und da A nach Generalvoraussetzung in diesem Kapitel nullteilerfrei ist gilt $ab \neq 0$.
(ii) Sei M ein A-Modul. Dann ist M/M_{tors} ein torsionsfreier Modul: Sei $x+M_{\text{tors}} \in M/M_{\text{tors}}$ beliebig. Sei $a \in A$ mit $a \neq 0$ und $a(x + M_{\text{tors}}) = 0$ in M/M_{tors}. Dann ist $ax \in M_{\text{tors}}$, es gibt also ein $b \in A$ mit $b \neq 0$ und $0 = b(ax) = (ba)x$ in M. Somit ist $x \in M_{\text{tors}}$ und daher gilt $x + M_{\text{tors}} = 0$ in M/M_{tors}.
(iii) Sei $M = A^n \oplus A/(a_1) \oplus \ldots \oplus A/(a_r)$ mit $0 \neq a_i$, $1 \leq i \leq r$. Dann gelten $M_{\text{tors}} = A/(a_1)\oplus\ldots\oplus A/(a_r)$ und somit $M/M_{\text{tors}} = A^n$. n ist der torsionsfreie Rang von M. Da M_{tors} unabhängig von der Darstellung von M als äußere direkte Summe definiert ist, und $A^m \ncong A^n$ für $m \neq n$ gilt, ist n wohldefiniert.

Definition 5.2.8. Sei M ein A-Modul. Sei $p \in A$ irreduzibel. Definiere

$$M(p) := \{x \in M \colon p^n x = 0 \text{ für ein } n \geq 1\}.$$

$M(p)$ ist ein Torsionsuntermodul von M und heißt die p-primäre Komponente von M. (Zur Untermoduleigenschaft: Beachte insbesondere, dass aus $p^m x = 0$ und $p^n y = 0$ auch $p^{\max\{m,n\}}(x + y) = 0$ folgt.)

Der Modul M heißt p-primär, falls $M(p) = M$ gilt.

Theorem 5.2.9. *Sei M ein endlich erzeugter A-Modul und gleichzeitig ein Torsionsmodul. Dann gibt es (abgesehen von assoziierten Wahlen) nur endlich viele irreduzible Elemente $p \in A$ mit $M(p) \neq \{0\}$. Wir bezeichnen diese mit $p_1, \ldots, p_m$. Es gelte $p_i \not\sim p_j$ für $i \neq j$. Dann gilt*

$$M = M(p_1) \oplus \ldots \oplus M(p_m).$$

Beweis. Nach dem Elementarteilersatz für Moduln, Theorem 5.2.4, gibt es Ringelemente $d_1, \ldots, d_r$ mit $d_i \neq 0$, $1 \leq i \leq r$, und $M \cong \bigoplus_{i=1}^{r} A/(d_i)$. Betrachte $d_1 \cdot \ldots \cdot d_r$. Seien $p_1, \ldots, p_m$ die irreduziblen Teiler dieses Produktes wobei wir assoziierte Teiler nur einmal aufführen. Dann gibt es $e_{i,j} \geq 0$, $1 \leq i \leq r$, mit $d_i \sim \prod_{j=1}^{m} p_j^{e_{i,j}}$ für alle $1 \leq i \leq r$. Nach dem Korollar zum Chinesischen Restsatz, Korollar 4.3.13, gilt somit

$$A/(d_i) \cong A/\left(p_1^{e_{i,1}}\right) \oplus \ldots \oplus A/\left(p_m^{e_{i,m}}\right)$$

für alle $1 \leq i \leq r$, da $(d_i) = (p_1^{e_{i,1}}) \cap \ldots \cap (p_m^{e_{i,m}})$ gilt. Damit folgt

$$M \cong \bigoplus_{i=1}^{r} A/(d_i) \cong \bigoplus_{j=1}^{m} \bigoplus_{i=1}^{r} A/\left(p_j^{e_{i,j}}\right) \equiv \bigoplus_{j=1}^{m} M(p_j),$$

wobei wir in der letzten Gleichheit verwendet haben, dass genau die Komponenten der äußeren direkten Summe zu $M(p_j)$ beitragen, die die Form $A/(p_j^e)$ für ein $e \geq 1$ haben. □

Bemerkung 5.2.10. Im Beweis haben wir gesehen, dass $M(p) \neq \{0\}$ genau dann gilt, wenn p einen Elementarteiler von M teilt.

Wichtig für den Eindeutigkeitssatz der Elementarteiler ist

Theorem 5.2.11. *Sei $p \in A$ irreduzibel. Sei M ein endlich erzeugter p-primärer A-Modul. Dann gilt*

$$M \cong A/(p^{e_1}) \oplus \ldots \oplus A/(p^{e_r})$$

mit $r \geq 0$ und $e_1, \ldots, e_r \in \mathbb{N}\backslash\{0\}$. Dabei sind r und die e_i's (bis auf die Reihenfolge) eindeutig bestimmt.

Beweis. Nach dem Elementarteilersatz für Moduln, Theorem 5.2.4, und Theorem 5.2.9 samt Beweis gibt es r und $e_1, \ldots, e_r \geq 1$ wie angegeben. Summanden mit anderen nicht assoziierten irreduziblen Ringelementen treten nicht auf, da M ein p-primärer Modul ist. Ebenso ist $e_i = 0$ ausgeschlossen.

Sei $i \geq 0$ beliebig. Dann ist $p^i M := \{p^i x \colon x \in M\}$ ein endlich erzeugter Untermodul von M. Es gilt die Inklusionskette

$$M = p^0 M \supseteq pM \supseteq p^2 M \supseteq \ldots.$$

Für $i \geq 1$ definieren wir A-Moduln $V_i := V_i(M) := p^{i-1}M/p^i M$. Es gilt $pV_i = \{0\}$. Daher erhalten wir bei der Skalarmultiplikation von Elementen in V_i mit beliebigen Elementen in $a + (p)$ für $a \in A$ stets dasselbe Resultat. Somit ist V_i auch ein Modul über $A/(p) =: K$. Da p irreduzibel ist, ist K ein Körper. Somit ist V_i ein K-Vektorraum.

Wir wollen nun $\dim_K(V_i)$ bestimmen: Dazu betrachten wir einen festen Summanden $N := A/(p^e)$ von M. Ist $i \geq e$, so gilt $p^i N = \{0\}$. Ist $i \leq e$, so ist $p^i N \cong p^i A/(p^e)$. Für $i \geq 1$ erhalten wir also

$$V_i(N) = \frac{p^{i-1}N}{p^i N} \cong \begin{cases} (p^{i-1})/(p^i) \cong K & \text{für } i \leq e, \\ \{0\} & \text{für } i > e. \end{cases}$$

Dabei haben wir für $i \leq e$

$$\frac{p^{i-1}N}{p^iN} \cong \frac{p^{i-1}A/p^eA}{p^iA/p^eA} \cong \frac{p^{i-1}A}{p^iA}$$

benutzt. Der letzte Isomorphismus im Fall $i \leq e$ ist dabei durch

$$K \cong A/pA \to \frac{p^{i-1}A}{p^iA},$$
$$a + pA \mapsto p^{i-1}a + p^iA$$

gegeben. Die Definition von V_i können wir in der Darstellung von M als direkte Summe komponentenweise anwenden und erhalten

$$V_i(M) \cong \bigoplus_{j=1}^{r} V_i(A/(p^{e_j})).$$

Damit folgt

$$\dim_K V_i(M) = \#\{j \in \{1, \ldots, r\} \colon e_j \geq i\}. \tag{5.1}$$

Durch M sind die Zahlen $\dim_K V_i(M)$ (unabhängig von der Darstellung von M als direkte Summe) eindeutig bestimmt. Gleichung (5.1) legt damit aber auch die Zahlen e_j bis auf die Reihenfolge eindeutig fest. □

Beispiel 5.2.12. Aus (5.1) folgt

$$\dim_K V_1(M) \geq \dim_K V_2(M) \geq \ldots$$

und $\dim_K V_i(M) = 0$ für genügend große $i \in \mathbb{N}$. Sind

i	1	2	3	4	5	6	7	
$\dim_K V_i(M)$	4	4	3	1	1	0	0	$\ldots$

vorgegeben, so folgt mit (5.1)

j	1	2	3	4
e_j	5	3	3	2

sowie $r = 4$.

Nun erhalten wir daraus die Eindeutigkeit der Elementarteiler.

Theorem 5.2.13. *Die Elementarteiler eines endlich erzeugten A-Moduls oder einer Matrix über A sind bis auf Assoziiertheit eindeutig bestimmt.*

Beweis. Aufgrund der Bemerkung 5.2.5 genügt es zu zeigen, dass die Elementarteiler von Moduln eindeutig bestimmt sind. Nach dem Elementarteilersatz für Moduln, Theorem 5.2.4, gilt

$$M \cong A^n \oplus A/(d_1) \oplus \ldots \oplus A/(d_r)$$

mit $1 \not\sim d_1 | \ldots | d_r \neq 0$, $r \geq 0$. Dann können wir $d_i \sim p_1^{e_{i,1}} \cdot \ldots \cdot p_m^{e_{i,m}}$ für alle $1 \leq i \leq r$ mit irreduziblen und paarweise nicht assoziierten $p_1, \ldots, p_m \in A$ schreiben. Aus $d_i | d_j$ für $i \leq j$ folgt $e_{1,j} \leq \ldots \leq e_{r,j}$ für $1 \leq j \leq m$. Für die primären Komponenten folgt wie im Beweis von Theorem 5.2.9

$$M(p_j) \cong A/\left(p_j^{e_{1,j}}\right) \oplus \ldots \oplus A/\left(p_j^{e_{r,j}}\right).$$

Nach Theorem 5.2.11 sind die (angeordneten) Zahlen $e_{1,j} \leq \ldots \leq e_{r,j}$ durch $M(p_j)$ und damit letztlich durch M eindeutig bestimmt. Damit sind die d_i's bis auf Assoziiertheit ebenfalls eindeutig bestimmt. □

Korollar 5.2.14. *Sei M ein endlich erzeugter Torsionsmodul mit Elementarteilern $d_1 | \ldots | d_r$. Dann ist M genau dann zyklisch, wenn $r \leq 1$ ist.*

Beweis.
„$\Longrightarrow$“: Sei M zyklisch. Nach Theorem 5.1.23 und da A ein Hauptidealring ist gilt $M \cong A/(d)$ für ein $d \in A$. Somit ist d der einzige Elementarteiler von M und $r \leq 1$. Aufgrund der Eindeutigkeit der Elementarteiler gibt es keine Darstellung mit $r > 1$.
„$\Longleftarrow$“: Klar. □

Wir wollen die Situation für den Ring $A = \mathbb{Z}$ genauer betrachten. Wir betrachten also $\mathbb{Z}$-Moduln, d. h. abelsche Gruppen. Diese schreiben wir additiv.

Korollar 5.2.15. *Sei G eine endlich erzeugte abelsche Gruppe. Dann gilt $G \cong \mathbb{Z}^n \times H$ mit $n \geq 0$ und einer endlichen abelschen Gruppe H. n und H sind bis auf Isomorphie eindeutig bestimmt.*

Beweis. Die Behauptung folgt aus dem Elementarteilersatz, wenn wir H als den Torsionsmodul des $\mathbb{Z}$-Moduls G definieren. □

Insbesondere folgt für endliche abelsche Gruppen

Korollar 5.2.16. *Sei G eine endliche abelsche Gruppe. Dann ist $G \cong G_1 \times \ldots \times G_r$ für endliche zyklische Gruppen G_i. Es gibt zwei Normalformen:*

(i) Elementarteilerform: Es gilt $\#G_1 | \#G_2 | \ldots | \#G_r$, $\#G_i \equiv |G_i|$. Dabei sind die Mächtigkeiten $|G_i|$ eindeutig bestimmt.
(ii) Primärzerlegungsform: $|G_i|$ ist eine Primzahlpotenz einer Primzahl p_i für alle $1 \leq i \leq r$ mit $p_i \not\sim p_j$ für $i \neq j$. Dabei sind die Mächtigkeiten $|G_i|$ bis auf Permutationen eindeutig bestimmt.

Beweis.
(i) Ist gerade der Elementarteilersatz.
(ii) Dies folgt aus Theorem 5.2.9 inklusive Beweis. □

Beispiel 5.2.17. Sei $G = \mathbb{Z}/(6) \oplus \mathbb{Z}/(56) \oplus \mathbb{Z}/(210)$. Wir gehen wie im Beweis des Elementarteilersatzes bzw. von Theorem 5.2.9 vor und erhalten:

(i) Primärzerlegungsform: Es gelten $6 = 2 \cdot 3$, $56 = 2^3 \cdot 7$ und $210 = 2 \cdot 3 \cdot 5 \cdot 7$. Somit sind $\mathbb{Z}/(6) \cong \mathbb{Z}/(2) \oplus \mathbb{Z}/(3)$, $\mathbb{Z}/(56) \cong \mathbb{Z}/\left(2^3\right) \oplus \mathbb{Z}/(7)$ und $\mathbb{Z}/(210) \cong \mathbb{Z}/(2) \oplus \mathbb{Z}/(3) \oplus \mathbb{Z}/(5) \oplus \mathbb{Z}/(7)$ nach Korollar 4.3.13 (Korollar zum Chinesischen Restsatz) und wir erhalten

$$G \cong \underbrace{\left(\mathbb{Z}/(2) \oplus \mathbb{Z}/(2) \oplus \mathbb{Z}/\left(2^3\right)\right)}_{=:G_1} \oplus \underbrace{\left(\mathbb{Z}/(3) \oplus \mathbb{Z}/(3)\right)}_{=:G_2} \oplus \underbrace{\mathbb{Z}/(5)}_{=:G_3} \oplus \underbrace{\left(\mathbb{Z}/(7) \oplus \mathbb{Z}/(7)\right)}_{=:G_4}.$$

(ii) Elementarteilerform: Wir ordnen die direkten Summanden um und erhalten

$$\begin{aligned} G \cong\ & \mathbb{Z}/(2) \oplus \mathbb{Z}/(2) \oplus \mathbb{Z}/(8) \\ & \oplus \mathbb{Z}/(3) \oplus \mathbb{Z}/(3) \\ & \oplus \mathbb{Z}/(5) \\ & \oplus \mathbb{Z}/(7) \oplus \mathbb{Z}/(7) \\ \cong & \mathbb{Z}/(2) \oplus \mathbb{Z}/(2 \cdot 3 \cdot 7) \oplus \mathbb{Z}/(8 \cdot 3 \cdot 5 \cdot 7) \\ = & \mathbb{Z}/(2) \oplus \mathbb{Z}/(42) \oplus \mathbb{Z}/(840). \end{aligned}$$

Somit sind die Elementarteiler 2, 42 und 840.

5.3. **Anwendungen auf Endomorphismen von Vektorräumen.**

Bemerkung 5.3.1 (Erinnerung)**.** Sei K ein Körper, V ein K-Vektorraum mit $\dim(V) < \infty$ und $f \in \mathrm{End}_K(V)$. Dann ist die Menge V mit der Operation

$$K[t] \times V \ni (p, v) \mapsto (p(f))(v) \in V$$

ein $K[t]$-Modul: V_f.

Wir wollen nun den $K[t]$-Modul V_f mit unter Zuhilfenahme des Elementarteilersatzes betrachten.

Bemerkung 5.3.2.

(i) V_f ist als $K[t]$-Modul endlich erzeugt: Verwende die Basis über K.

(ii) V_f ist ein $K[t]$-Torsionsmodul: Für das Minimalpolynom $\mu_f(t)$ gilt $\mu_f(f) = 0$. Somit ist $\mu_f \cdot v = (\mu_f(f))(v) = 0(v) = 0$. Dies gilt für beliebiges $v \in V$.

Ohne Verwendung von μ_f sieht man dies wie folgt: Gelte $K[t] \cdot v_1 + \ldots + K[t] \cdot v_r = V_p$. Dann gibt es Polynome $p_i \in K[t]$ mit $p_i v_i = 0$. Somit ist $(p_1 \cdot \ldots \cdot p_r)v = 0$ für alle $v \in V$.

(iii) Wir erinnern an die Definition des Annulatorideals (des Annulators)

$$\mathrm{Ann}(V_f) := \{p \in K[t] \colon p \cdot v = 0 \text{ für alle } v \in V\}.$$

Damit folgt $\mathrm{Ann}(V_f) = (\mu_f)$. (Es ist einfach zu sehen, dass es sich dabei um ein Ideal handelt.)

(iv) Nach dem Elementarteilersatz für Moduln, Theorem 5.2.4, ist V_f die direkte Summe von zyklischen Torsionsmoduln über $K[t]$.

Definition 5.3.3. $f \in \mathrm{End}_K(V)$ heißt zyklischer Endomorphismus von V, falls der $K[t]$-Modul V_f zyklisch ist.

Theorem 5.3.4. *Sei $n = \dim(V)$, $1 \leq n < \infty$. Sei $f \in End_K(V)$. Dann sind die folgenden Aussagen äquivalent:*

(i) f ist ein zyklischer Endomorphismus.

(ii) Es gibt einen Vektor $v \in V$, so dass $v, f(v), \ldots, f^{n-1}(v)$ eine Basis des K-Vektorraumes V ist.

(iii) Es gibt eine Basis von V, so dass f bezüglich dieser Basis durch die Matrix

$$A = \begin{pmatrix} 0 & & & & a_0 \\ 1 & 0 & & & a_1 \\ & \ddots & \ddots & & \vdots \\ & & 1 & 0 & a_{n-2} \\ & & & 1 & a_{n-1} \end{pmatrix}$$

mit $a_0, \ldots, a_{n-1} \in K$ dargestellt wird. Formal: $A = \left(a^i_j\right)_{1 \leq i,j \leq n}$ mit $a^i_j = \delta^i_{j+1}$ für $j < n$ und $a^i_j \in K$ beliebig für $j = n$. (Eine solche Matrix nennen wir eine Begleitmatrix.)

Beweis.

„(i) $\Longrightarrow$ (ii)“: Da f zyklisch ist, gibt es ein $v \in V$ mit $K[t] \cdot v = V_f$. Insbesondere gilt daher $V = \left\langle \left\{ f^i(v) \colon i \geq 0 \right\} \right\rangle$. Wir behaupten, dass die Vektoren $v, f(v), \ldots, f^{n-1}(v)$ linear unabhängig sind. (Dann bilden sie aus Dimensionsgründen bereits eine Basis von V.) Falls dies nicht der Fall ist, gilt $f^k(v) = \sum_{i=0}^{k-1} a_j f^j(v)$ mit $a_i \in K$ und $1 \leq k \leq n-1$. Per Induktion erhalten wir daraus $f^{k+r}(v) \in \left\langle \left\{ f^i(v) \colon 0 \leq i \leq k-1 \right\} \right\rangle =: U$ für alle $r \in \mathbb{N}$. Somit folgt $V = U$, da f zyklisch ist. Dies ist aber unmöglich, da $k = \dim U < n = \dim V$ gilt.

„(ii) $\Longrightarrow$ (iii)“: Benutze die Basis $v, f(v), \ldots, f^{n-1}(v)$.

„(iii) $\Longrightarrow$ (i)“: Sei $v_0, \ldots, v_{n-1}$ die Basis in (iii). Dann gilt $v_i = f^i(v_0)$ für alle $1 \leq i \leq n-1$. Somit ist f zyklisch. □

Lemma 5.3.5. *Seien f und A wie in Theorem 5.3.4. Dann tauchen die Koeffizienten a_i im charakteristischen Polynom von f auf, sind also eindeutig durch f bestimmt.*

Beweis. Wir entwickeln nach der ersten Zeile und erhalten

$$\det(t\mathbb{1}-A)=\det\begin{pmatrix} t & & & & -a_0\\ -1 & t & & & -a_1\\ & \ddots & \ddots & & \vdots\\ & & -1 & t & -a_{n-2}\\ & & & -1 & t-a_{n-1}\end{pmatrix}$$

$$=t\cdot\det\begin{pmatrix} t & & & & -a_1\\ -1 & t & & & -a_2\\ & \ddots & \ddots & & \vdots\\ & & -1 & t & -a_{n-2}\\ & & & -1 & t-a_{n-1}\end{pmatrix}+(-a_0)\cdot\underbrace{(-1)^{n+1+n-1}}_{=1}.$$

Also folgt per Induktion

$$\det(t\mathbb{1}-A)=t^n-a_{n-1}t^{n-1}-\ldots-a_1t-a_0$$

und somit

$$\chi_A(t)=\det(A-t\mathbb{1})=(-1)^n\det(t\mathbb{1}-a)=(-1)^n\left(t^n-\sum_{i=0}^{n-1}a_it^i\right).\qquad\square$$

Lemma 5.3.6. *Sei V ein K-Vektorraum mit $1\leq\dim V<\infty$. Sei $f\in\mathrm{End}_K(V)$ zyklisch. Sei $p\in K[t]$ das eindeutig bestimmte normierte Polynom so dass $V_f\cong K[t]/(p)$ mit einem $K[t]$-Modulisomorphismus gilt. (Solch ein p existiert nach Theorem 5.1.23 und ist aufgrund der Eindeutigkeit der Elementarteiler und der Normierung eindeutig bestimmt.) Dann gelten*

(i) $p=\mu_f=\chi_f$, d. h. p, das Minimalpolynom von f und das charakteristische Polynom von f stimmen überein.

(ii) Ist $p=t^n+\sum_{i=0}^{n-1}a_it^i$, so ist f bezüglich einer geeigneten Basis durch die Matrix

$$A=\begin{pmatrix} 0 & & & & -a_0\\ 1 & 0 & & & -a_1\\ & \ddots & \ddots & & \vdots\\ & & 1 & 0 & -a_{n-2}\\ & & & 1 & -a_{n-1}\end{pmatrix}$$

dargestellt.

Beweis.
(ii): Da f bzw. V_f zyklisch sind, gibt es $v\in V$ mit $V_f=K[t]\cdot v$. Betrachte $\varphi\colon K[t]\to V_f$ mit $q\mapsto qv$. Aufgrund des Homomorphiesatzes gilt $V_f\cong K[t]/\ker\varphi$. Andererseits ist $V_f\cong K[t]/(p)$. Aufgrund der Eindeutigkeit der Elementarteiler gilt $\ker\varphi=(p)$. Betrachten wir $K[t]/(p)$ als K-Vektorraum, so gilt $\dim_K(K[t]/(p))=\dim_K V_f=\dim_K V=n$. Somit gilt $\deg p=n$. Nun ist $1,t,\ldots,t^{n-1}$ eine Basis des K-Vektorraumes $K[t]/(p)$. Also ist auch $v,f(v),\ldots,f^{n-1}(v)$ eine Basis von V und mit den Koeffizienten $a_i\in K$ von p gilt auch $f^n(v)=-\sum_{i=0}^{n-1}a_if^i(v)$. Wegen $f\left(f^{n-1}(v)\right)=f^n(v)$ lässt sich f bezüglich dieser Basis genau durch A darstellen.

(i): Aus Bemerkung 5.3.2 folgt dass p der normierte Erzeuger von $\mathrm{Ann}(f)$ ist. Somit gilt $p=\mu_f$. In Lemma 5.3.5 haben wir bereits gesehen, dass $p=\chi_f$ gilt. $\square$

Theorem 5.3.7. *Sei V ein K-Vektorraum mit $1\leq\dim V<\infty$. Sei nun $f\in\mathrm{End}_K(V)$ beliebig. Seien $p_1|\ldots|p_r$ die normierten Elementarteiler des $K[t]$-Moduls V_f. Dann gelten $\chi_f=p_1\cdot\ldots\cdot p_r$ und $\mu_f=p_r$.*

Beweis. Da V_f zyklisch ist, gilt nach dem Elementarteilersatz $V_f \cong \bigoplus_{j=1}^{r} K[t]/(p_j)$ mit einem Isomorphismus von $K[t]$-Moduln. Das Annulatorideal ergibt sich als Schnitt über die Annulatorideale der einzelnen Summanden. Also ist $\mathrm{Ann}(V_f) = \bigcap_{j=1}^{r} (p_j) = (p_r)$ da $p_1 | \dots | p_r$. Somit ist $\mu_f = p_r$.

Der Isomorphismus von $K[t]$-Moduln $V_f \cong \bigoplus_{j=1}^{r} K[t]/(p_j)$ ist zugleich auch ein Vektorraumisomorphismus (da wir die Multiplikation auf K einschränken können und beide Seiten Vektorräume sind). Er induziert eine Zerlegung $V = V_1 \oplus \dots \oplus V_r$ in f-invariante Unterräume V_j, so dass $f|_{V_j}$ zyklisch ist und p_j das Minimalpolynom ist, siehe Lemma 5.3.6. Wir können also f mit Hilfe einer Blockmatrix darstellen. Die Kästchen sehen dabei bezüglich einer geeigneten Basis so wie in Lemma 5.3.6 aus und ergeben jeweils als charakteristische Polynome die Minimalpolynome p_j. Somit ist

$$\chi_f = \prod_{i=1}^{r} \chi_{f|_{V_i}} = \prod_{i=1}^{r} p_i. \qquad \square$$

Korollar 5.3.8 (Satz von Cayley-Hamilton)**.** *Sei V ein K-Vektorraum mit $1 \leq \dim V < \infty$. Sei $f \in \mathrm{End}(V)$ beliebig. Dann gilt $\chi_f(f) = 0$.*

Beweis. Nach Theorem 5.3.7 gilt $\mu_f | \chi_f$ wie man direkt an den expliziten Darstellungen abliest. Für das Minimalpolynom gilt natürlich $\mu_f(f) = 0$. $\square$

Dieser Beweis ist unabhängig vom bisherigen Beweis des Satzes von Cayley-Hamilton.

Korollar 5.3.9. *Seien V, f wie in Theorem 5.3.7. Dann stimmen die irreduziblen Faktoren von χ_f und μ_f überein.*

Beweis. Jeder irreduzible Faktor von χ_f teilt ein p_i in der Notation von Theorem 5.3.7, also wegen $p_1 | \dots | p_r$ auch $\mu_f = p_r$. $\square$

Korollar 5.3.10. *Seien V, f wie in Theorem 5.3.7. Dann ist f genau dann zyklisch, wenn $\chi_f = \mu_f$ gilt.*

Zerfällt χ_f in Linearfaktoren (z. B. für $K = \mathbb{C}$, siehe Funktionentheorie), so sind die folgenden Bedingungen äquivalent.

(i) Es gilt $\chi_f = \mu_f$.
(ii) Alle Eigenräume von f sind eindimensional.
(iii) Jede f darstellende Jordanmatrix hat zu jedem Eigenwert nur einen Jordanblock.

Beweis.
„$\Longrightarrow$“: Sei f zyklisch. Nach Korollar 5.2.14 ist f genau dann zyklisch, wenn V_f maximal einen Elementarteiler besitzt. (Aus Theorem 5.3.7 folgt aber insbesondere, dass V_f mindestens einen Elementarteiler besitzt: Betrachte das charakteristische Polynom.) Theorem 5.3.7 impliziert nun $\chi_f = \mu_f$.

„$\Longleftarrow$“: Ist $\chi_f = \mu_f$, so besitzt V_f genau einen Elementarteiler, ist also zyklisch.

Die weiteren Äquivalenzen sind einfach zu sehen. $\square$

Bemerkung 5.3.11. Mit Hilfe des Elementarteilersatzes können wir einen Beweis für die Existenz der Jordanschen Normalform geben der statt expliziter Konstruktionen Modultheorie verwendet.

Seien f, V wie in Theorem 5.3.7. Zerfalle χ_f in Linearfaktoren. Nach Theorem 5.2.9 genügt es, einen primären Summanden von V_f und aufgrund des Beweises sogar, einen zyklischen Summanden zu betrachten und zu zeigen, dass dieser ein

Jordankästchen liefert. Der Vektorraum V lässt sich genauso als Summe schreiben. Sei also ohne Einschränkung V_f ein zyklischer primärer $K[t]$-Modul. Da χ_f in Linearfaktoren zerfällt gilt $V_f \cong K[t]/((t-\lambda)^n)$ für ein $n \geq 1$ und ein $\lambda \in K$. Da V_f zyklisch ist, gibt es ein $v \in V$ mit $V_f = K[t] \cdot v$. Definiere $v_j := (t-\lambda)^j v = (f - \lambda \mathbb{1})^j(v)$ für $j = 0, \ldots, n-1$. Nun bilden $v_0, \ldots, v_{n-1}$ eine Vektorraumbasis von V, da die Äquivalenzklassen von $1, t-\lambda, (t-\lambda)^2, \ldots, (t-\lambda)^{n-1}$ eine Vektorraumbasis von $K[t]/((t-\lambda)^n)$ bilden. Wir wollen nun die Basis $v_{n-1}, \ldots, v_1, v_0 = v$ benutzen. Es gilt

$$(f-\lambda)(v_j) = v_{j+1} \quad \text{für } 1 \leq j \leq n-2$$

sowie

$$(f-\lambda)(v_{n-1}) = 0.$$

Somit ist $f - \lambda$ bezüglich dieser Basis durch die linke Matrix und f durch die rechte Matrix in

$$\begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}, \qquad \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}$$

dargestellt, wobei sämtliche fehlende Einträge Nullen sind. Dies ist genau ein Jordankästchen. □

Literatur

1. Siegfried Bosch, *Algebra*, 6. Auflage, Springer-Verlag, Berlin, 2005.
2. Siegfried Bosch, *Lineare Algebra*, 4. Auflage, Springer-Verlag, Berlin, 2008.
3. Theodor Bröcker, *Lineare Algebra und analytische Geometrie*, Grundstudium Mathematik, Birkhäuser Verlag, Basel, 2003, Ein Lehrbuch für Physiker und Mathematiker.
4. Gerd Fischer, *Lineare Algebra*, fifth ed., Grundkurs Mathematik, vol. 17, Friedr. Vieweg & Sohn, Braunschweig, 1979.
5. Roger A. Horn and Charles R. Johnson, *Topics in matrix analysis*, Cambridge University Press, Cambridge, 1994, Corrected reprint of the 1991 original.
6. Serge Lang, *Algebra*, third ed., Graduate Texts in Mathematics, vol. 211, Springer-Verlag, New York, 2002.
7. O. Timothy O'Meara, *Introduction to quadratic forms*, Classics in Mathematics, Springer-Verlag, Berlin, 2000, Reprint of the 1973 edition.
8. Claus Scheiderer, *Lineare Algebra über Ringen*, 2009, Notizen zur Vorlesung.
9. Oliver C. Schnürer, *Lineare Algebra I*, 2010, Skript zur Vorlesung.
10. Urs Stammbach, *Lineare Algebra*, Teubner Studienskripten, vol. 82, B. G. Teubner, Stuttgart, 1980.
11. Wikipedia, `http://www.wikipedia.org`.

Oliver C. Schnürer, Fachbereich Mathematik und Statistik, Universität Konstanz, 78457 Konstanz, Germany

E-mail address: `Oliver.Schnuerer@uni-konstanz.de`