14 Reelle quadratische Formen

14.1 Der Trägheitssatz von Sylvester [James Joseph Sylvester *1814 †1897]

Satz und Definition 14.1.1 (Trägheitssatz). *Sei V ein endlichdimensionaler $\mathbb{R}$-Vektorraum und $q \in Q(V)$. Dann gibt es genau ein Paar $(r,s) \in \mathbb{N}_0^2$, genannt* Sylvester-Signatur *von q, derart, dass es eine geordnete Basis $\underline{v}$ von V gibt mit*

$$M(q,\underline{v}) = \begin{pmatrix} \overbrace{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}}^{r} & & \\ & \overbrace{\begin{matrix} -1 & & \\ & \ddots & \\ & & -1 \end{matrix}}^{s} & \\ & & \underbrace{\begin{matrix} 0 & & \\ & \ddots & \\ & & 0 \end{matrix}}_{} \end{pmatrix}$$

Beweis. Existenz Nach Korollar 13.5.12 gibt es eine Basis $\underline{w} = (w_1, \dots, w_n)$ von V derart, dass $M(q,\underline{w})$ Diagonalgestalt hat, Œ gibt es $r,s \in \mathbb{N}_0$ mit

$$q(w_1) > 0, \dots, q(w_r) > 0,\ q(w_{r+1}) < 0, \dots, q(w_{r+s}) < 0,\ q(w_{r+s+1}) = \dots = q(w_n) = 0.$$

Setze nun

$$\underline{v} := \left(\frac{w_1}{\sqrt{q(w_1)}}, \dots, \frac{w_r}{\sqrt{q(w_r)}}, \frac{w_{r+1}}{\sqrt{-q(w_{r+1})}}, \dots, \frac{w_{r+s}}{\sqrt{-q(w_{r+s})}}, w_{r+s+1}, \dots, w_n \right).$$

Dann ist $M(q,\underline{v})$ von der gewünschten Gestalt.

Eindeutigkeit Seien $(r,s),(t,u) \in \mathbb{N}_0^2$ und $\underline{v} = (v_1, \dots, v_n)$, $\underline{w} = (w_1, \dots, w_n)$ Basen von V mit

$$M(q,\underline{v}) = \begin{pmatrix} \overbrace{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}}^{r} & & \\ & \overbrace{\begin{matrix} -1 & & \\ & \ddots & \\ & & -1 \end{matrix}}^{s} & \\ & & \underbrace{\begin{matrix} 0 & & \\ & \ddots & \\ & & 0 \end{matrix}}_{} \end{pmatrix} \quad \text{und} \quad M(q,\underline{w}) = \begin{pmatrix} \overbrace{\begin{matrix} 1 & & \\ & \ddots & \\ & & 1 \end{matrix}}^{t} & & \\ & \overbrace{\begin{matrix} -1 & & \\ & \ddots & \\ & & -1 \end{matrix}}^{u} & \\ & & \underbrace{\begin{matrix} 0 & & \\ & \ddots & \\ & & 0 \end{matrix}}_{} \end{pmatrix}.$$

Zu zeigen: $(r,s) = (t,u)$. Setze $b := b_q$ [→ 13.4.6]. Dann gilt für $\lambda_1, \ldots, \lambda_n \in K$:

$$\begin{aligned}
\sum_{i=1}^{n} \lambda_i v_i \in \ker \vec{b} &\iff \vec{b}\left(\sum_{i=1}^{n} \lambda_i v_i\right) = 0 \\
&\iff b\left(\sum_{i=1}^{n} \lambda_i v_i, \ .\right) = 0 \\
&\iff \forall j \in \{1,\ldots,n\} : \underbrace{b\left(\sum_{i=1}^{n} \lambda_i v_i, v_j\right)}_{=\lambda_j q(v_j)} = 0 \\
&\iff \forall j \in \{1,\ldots,r+s\} : \lambda_j = 0.
\end{aligned}$$

Daher $\operatorname{span}(v_{r+s+1}, \ldots, v_n) = \ker \vec{b}$ und ebenso $\operatorname{span}(w_{t+u+1}, \ldots, w_n) = \ker \vec{b}$. Es folgt insbesondere $r + s = t + u$. Betrachte nun die Untervektorräume

$$\begin{aligned}
U &:= \operatorname{span}(v_1, \ldots, v_r, v_{r+s+1}, \ldots, v_n) \qquad \text{und} \\
W &:= \operatorname{span}(w_{t+1}, \ldots, w_{t+u}).
\end{aligned}$$

Es gilt $q(v) \geq 0$ für $v \in U$ und $q(v) < 0$ für $v \in W \setminus \{0\}$. Daher gilt $U \cap W = \{0\}$ und mit der Dimensionsformel 8.1.12 für Untervektorräume folgt

$$n \geq \dim(U + W) = (\dim U) + (\dim W) = (n - s) + u,$$

also $u \leq s$. Genauso zeigt man $s \leq u$. Es folgt $s = u$ und daher auch $r = t$. Somit $(r,s) = (t,u)$. □

Satz 14.1.2. [1] *Sei V ein endlichdimensionaler $\mathbb{R}$-Vektorraum und $q \in Q(V)$ mit Sylvester-Signatur (r,s).*

(a) *Ist $\underline{v} = (v_1, \ldots, v_n)$ eine Basis von V und sind $\lambda_1, \ldots, \lambda_n$ die Eigenwerte von $M(q, \underline{v})$, wobei jedes λ_i seiner algebraischen Vielfachheit entsprechend oft aufgeführt ist* [→ 10.1.14, 11.3.10]*, das heißt $\chi_{M(q,\underline{v})} = \det(M(q,\underline{v}) - XI_n) = (-1)^n \prod_{i=1}^{n}(X - \lambda_i)$, so gilt*

$$\begin{aligned}
r &= \#\{i \in \{1,\ldots,n\} \mid \lambda_i > 0\} \qquad \textit{und} \\
s &= \#\{i \in \{1,\ldots,n\} \mid \lambda_i < 0\}.
\end{aligned}$$

(b) *Ist $\underline{v} = (v_1, \ldots, v_n)$ eine Basis von V, $D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix} \in \mathbb{R}^{n\times n}$ eine Diagonalmatrix und $P \in \mathbb{R}^{n\times n}$ invertierbar mit $M(q,\underline{v}) = P^T D P$ oder $M(q,\underline{v}) = P^{-1} D P$, so gilt*

$$\begin{aligned}
r &= \#\{i \in \{1,\ldots,n\} \mid d_i > 0\} \qquad \textit{und} \\
s &= \#\{i \in \{1,\ldots,n\} \mid d_i < 0\}.
\end{aligned}$$

[1] Im Beweis dieses Satzes benutzen wir den Fundamentalsatz der Algebra 4.2.12 [→ 4.2.13(g)].

(c) *Sind $\ell_1, \ldots, \ell_m \in V^*$ linear unabhängig und $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ mit*

$$\forall v \in V : q(v) = \sum_{i=1}^{m} \lambda_i (\ell_i(v))^2,$$

so gilt

$$r = \#\{i \in \{1, \ldots, m\} \mid \lambda_i > 0\} \qquad \textit{und}$$
$$s = \#\{i \in \{1, \ldots, m\} \mid \lambda_i < 0\}.$$

Beweis. Wir zeigen zunächst den Fall $M(q, \underline{v}) = P^T D P$ von (b), dann (a), dann den Fall $M(q, \underline{v}) = P^{-1} D P$ von (b) und schließlich (c).

(b) Fall $\mathbf{M(q, \underline{v}) = P^T D P}$

Schritt 1: OE $P = I_n$

Begründung: Setze $w_i := \mathrm{vec}_{\underline{v}}(P^{-1} e_i)$ für $i \in \{1, \ldots, n\}$. Dann ist $\underline{w} := (w_1, \ldots, w_n)$ eine Basis von V [→ 6.3.7, 6.3.8] und es gilt $P^{-1} = M(\underline{w}, \underline{v})$ und somit

$$M(q, \underline{w}) \overset{13.3.10}{=} M(\underline{w}, \underline{v})^T M(q, \underline{v}) M(\underline{w}, \underline{v}) = (P^{-1})^T P^T D P P^{-1} = (P P^{-1})^T D (P P^{-1}) = D.$$

Schritt 2: OE D von der Gestalt $\begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & -1 & \\ & & & & \ddots \\ & & & & & -1 \\ & & & & & & 0 \\ & & & & & & & \ddots \\ & & & & & & & & 0 \end{pmatrix}$.

Begründung: Vertauschen der Basiselemente und Skalieren derselben mit positiven Skalaren ändert die Anzahl der positiven bzw. negativen Diagonaleinträge von $M(q, \underline{v})$ nicht.

(a) Da $M(q, \underline{v})$ eine reelle symmetrische Matrix ist, gibt es nach 11.3.10 eine orthogonale [→ 11.2.25, 11.2.27] Matrix $P \in \mathbb{R}^{n \times n}$ und eine Diagonalmatrix

$$D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix} \in \mathbb{R}^{n \times n}$$

mit $M(q, \underline{v}) = P^T D P = P^{-1} D P$. Da $M(q, \underline{v})$ und D ähnlich sind, haben sie dasselbe charakteristische Polynom [→ 10.1.4], das heißt $\prod_{i=1}^{n} (X - \lambda_i) = \prod_{i=1}^{n} (X - d_i)$. Dann gibt es nach 10.1.13 eine Permutation $\sigma \in S_n$ mit $(\lambda_1, \ldots, \lambda_n) = (d_{\sigma(1)}, \ldots, d_{\sigma(n)})$, weshalb man in der Behauptung die λ_i durch die d_i ersetzen kann. Wegen $M(q, \underline{v}) = P^T D P$ folgt aber dann die Behauptung nach dem bereits bewiesenen Teil von (b).

(b) Fall $\mathbf{M(q, \underline{v}) = P^{-1} D P}$

Setze $\lambda_i := d_i$ für $i \in \{1, \ldots, n\}$. Dann gilt $\chi_{M(q,\underline{v})} \overset{10.1.4}{=} \chi_D = (-1)^n \prod_{i=1}^n (X - \lambda_i)$ und die Behauptung folgt nun aus (a).

(c) Ergänze $\ell_1, \ldots, \ell_m$ zu einer Basis $\ell_1, \ldots, \ell_m, \ell_{m+1}, \ldots, \ell_n$ von V^* und setze $\lambda_{m+1} = \ldots = \lambda_n = 0$. Wähle eine Basis $\underline{v} = (v_1, \ldots, v_n)$ von V. Nach 13.5.2 ist dann $P := (\ell_i(v_j))_{1 \le i,j \le n} \in \mathbb{R}^{n \times n}$ invertierbar und mit $D := \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \in \mathbb{R}^{n \times n}$ gilt $M(q, \underline{v}) = P^T D P$. Nun folgt die Behauptung aus (b). □

Definition 14.1.3. Sei $n \in \mathbb{N}_0$ und $A \in S\mathbb{R}^{n \times n}$. Dann definiert man die Sylvester-Signatur von A als die Sylvester-Signatur der zu A gehörigen quadratischen Form $q_A \colon \mathbb{R}^n \to \mathbb{R},\ x \mapsto x^T A x$.

Bemerkung 14.1.4. Sei V ein $\mathbb{R}$-Vektorraum mit Basis $\underline{v} = (v_1, \ldots, v_n)$ und $q \in Q(V)$. Dann stimmen die Sylvester-Signaturen von q und von $M(q, \underline{v})$ natürlich überein, denn setzt man $A := M(q, \underline{v})$, so gilt $M(q_A, \underline{e}) = A = M(q, \underline{v})$ und es liefert zum Beispiel Teil (a) des obigen Satzes das Gewünschte.

Korollar 14.1.5. *Sei $n \in \mathbb{N}_0$ und $A \in S\mathbb{R}^{n \times n}$ mit Sylvester-Signatur (r, s).*

(a) *Gilt*

$$\chi_A = (-1)^n \prod_{i=1}^n (X - \lambda_i) \qquad \textit{mit} \qquad \lambda_1, \ldots, \lambda_n \in \mathbb{R},$$

so gilt $r = \#\{i \in \{1, \ldots, n\} \mid \lambda_i > 0\}$ und $s = \#\{i \in \{1, \ldots, n\} \mid \lambda_i < 0\}$.

(b) *Ist $D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix} \in \mathbb{R}^{n \times n}$ und $P \in \mathbb{R}^{n \times n}$ invertierbar mit*

$$M(q, \underline{v}) = P^T D P \qquad \text{oder} \qquad M(q, \underline{v}) = P^{-1} D P,$$

so gilt $r = \#\{i \in \{1, \ldots, n\} \mid d_i > 0\}$ und $s = \#\{i \in \{1, \ldots, n\} \mid d_i < 0\}$.

(c) *Sind $\ell_1, \ldots, \ell_m \in (\mathbb{R}^n)^*$ linear unabhängig und $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ mit*

$$\forall x \in \mathbb{R}^n : x^T A x = \sum_{i=1}^m \lambda_i (\ell_i(x))^2,$$

so gilt $r = \#\{i \in \{1, \ldots, n\} \mid \lambda_i > 0\}$ und $s = \#\{i \in \{1, \ldots, n\} \mid \lambda_i < 0\}$.

14.2 Positiv semidefinite Matrizen

Definition 14.2.1. Sei V ein $\mathbb{R}$-Vektorraum. Man nennt $q \in Q(V)$ $\left\{ \begin{matrix} \textit{positiv} \\ \textit{negativ} \end{matrix} \right\}$ *semidefinit* $\left\{ \begin{matrix} \textit{(psd)} \\ \textit{(nsd)} \end{matrix} \right\}$, wenn $\forall v \in V : q(v) \left\{ \begin{matrix} \ge \\ \le \end{matrix} \right\} 0$. Gilt zusätzlich $\forall v \in V : (q(v) = 0 \implies v = 0)$,

so nennt man q $\begin{Bmatrix} positiv \\ negativ \end{Bmatrix}$ *definit* $\begin{Bmatrix} (pd) \\ (nd) \end{Bmatrix}$. Man nennt $b \in \mathrm{Bil}(V)$ psd/nsd/pd/nd, wenn b symmetrisch und q_b psd/nsd/pd/nd ist.

Beispiel 14.2.2. Ein Skalaprodukt auf einem *reellen* Vektorraum ist per Definition nichts anderes als eine positive definite Bilinearform [→ 11.1.1].

Definition 14.2.3. Sei $n \in \mathbb{N}_0$ und $A \in \mathbb{R}^{n\times n}$. Es heißt A psd/nsd/pd/nd, wenn die zu A gehörige Bilinearform $b_A \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $(x, y) \mapsto x^T Ay$ [→ 13.3.2(a)] psd/nsd/pd/nd ist.

Bemerkung 14.2.4. Sei $n \in \mathbb{N}_0$ und $A \in \mathbb{R}^{n\times n}$. Dann ist A psd/nsd/pd/nd genau dann, wenn A symmetrisch ist [→ 13.4.2, 9.1.21] und $q_A \colon \mathbb{R}^n \to \mathbb{R}$, $x \mapsto x^T Ax$ psd/nsd/pd/nd ist.

Bemerkung 14.2.5. Sei V ein $\mathbb{R}$-Vektorraum mit Basis $\underline{v} = (v_1, \ldots, v_n)$ und $q \in Q(V)$. Dann q psd/nsd/pd/nd $\iff$ $M(q, \underline{v})$ psd/nsd/pd/nd.

Bemerkung 14.2.6. Für reelle quadratische Formen q gilt natürlich

$$q \text{ nsd} \iff -q \text{ psd} \qquad \text{und} \qquad q \text{ nd} \iff -q \text{ pd}.$$

Analoges gilt für Bilinearformen und für Matrizen. Daher betrachten wir im Folgenden nur noch die Begriffe „psd“ und „pd“.

Satz 14.2.7. *Sei V ein endlichdimensionaler $\mathbb{R}$-Vektorraum und $q \in Q(V)$. Dann sind äquivalent:*

(a) *q ist psd.*

(b) *Die Sylvester-Signatur von q ist $(r, 0)$ für ein $r \in \mathbb{N}_0$.*

(c) $\exists \ell_1, \ldots, \ell_n \in V^* : \forall v \in V : q(v) = \sum_{i=1}^n (\ell_i(v))^2$

(d) $\exists m \in \mathbb{N}_0 : \exists \ell_1, \ldots, \ell_m \in V^* : \forall v \in V : q(v) = \sum_{i=1}^m (\ell_i(v))^2$

Beweis. (a) $\Longrightarrow$ (b) folgt direkt aus der Definition der Sylvester-Signatur 14.1.1.

(b) $\Longrightarrow$ (c) folgt ebenfalls aus dieser Definition zusammen mit Lemma 13.5.1.

(c) $\Longrightarrow$ (d) $\Longrightarrow$ (a) sind trivial. □

Definition 14.2.8. [→ 13.5.7] Sei $A \in S\mathbb{R}^{n\times n}$. Unter einer *Cholesky-Zerlegung* [André Louis Cholesky *1875 †1918] von A verstehen wir ein Paar (P, D) von Matrizen $P, D \in \mathbb{R}^{n\times n}$ mit $A = P^T DP$, wobei P von oberer Dreiecksgestalt mit lauter Einsen auf der Diagonale und D von Diagonalgestalt ohne negative Einträge ist.

Definition 14.2.9. Sei K ein kommutativer Ring, $n \in \mathbb{N}_0$ und $A \in K^{n\times n}$. Für jedes $I \subseteq \{1, \ldots, n\}$ bezeichne $A_I \in K^{(\#I)\times(\#I)}$ die Matrix, die aus A durch Streichen aller Zeilen i und Spalten i mit $i \notin I$ entsteht. Wir bezeichnen die Determinanten der n Matrizen

$$A_{\{1\}}, A_{\{1,2\}}, \ldots, A_{\{1,\ldots,n\}}$$

als die *Leithauptminoren (oder führende Hauptminoren)* von A und die Determinanten der $2^n - 1$ Matrizen

$$A_I \qquad (\emptyset \neq I \subseteq \{1, \ldots, n\})$$

als die *Hauptminoren* von A. [*Vorsicht:* Manche deutschsprachigen Autoren bezeichnen nur die Leithauptminoren als Hauptminoren und haben keine Bezeichnung für unsere Hauptminoren.]

Beispiel 14.2.10. Die Leithauptminoren von $A := \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \in \mathbb{R}^{3\times 3}$ sind 1, 0, 0 und ihre Hauptminoren sind die Diagonaleinträge 1, 1, -1, die Determinanten $\det \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = 0$, $\det \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = -1$, $\det \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = -1$ und $\det(A) = 0$.

Satz 14.2.11. [2] *Sei $A \in S\mathbb{R}^{n\times n}$. Dann sind äquivalent:*

(a) *A ist psd.*

(b) $\forall x \in \mathbb{R}^n : x^T A x \geq 0$

(c) *Die Sylvester-Signatur von A ist $(r, 0)$ für ein $r \in \mathbb{N}_0$.*

(d) *Alle Eigenwerte von A sind ≥ 0.*

(e) *Alle Koeffizienten von* $\det(A + XI_n) = \chi_A(-X) \in \mathbb{R}[X]$ *sind* ≥ 0.

(f) *Alle Hauptminoren von A sind ≥ 0.*

(g) *A besitzt eine Cholesky-Zerlegung.*

(h) *Es gibt eine obere Dreiecksmatrix* [$\rightarrow$ 10.3.1] $B \in \mathbb{R}^{n\times n}$ *mit* $A = B^T B$.

(i) $\exists B \in \mathbb{R}^{n\times n} : A = B^T B$

(j) $\exists m \in \mathbb{N}_0 : \exists B \in \mathbb{R}^{m\times n} : A = B^T B$

(k) $\exists v_1, \ldots, v_n \in \mathbb{R}^n : A = \begin{pmatrix} \langle v_1, v_1\rangle & \ldots & \langle v_1, v_n\rangle \\ \vdots & & \vdots \\ \langle v_n, v_1\rangle & \ldots & \langle v_n, v_n\rangle \end{pmatrix}$

(l) $\exists m \in \mathbb{N}_0 : \exists v_1, \ldots, v_n \in \mathbb{R}^m : A = \begin{pmatrix} \langle v_1, v_1\rangle & \ldots & \langle v_1, v_n\rangle \\ \vdots & & \vdots \\ \langle v_n, v_1\rangle & \ldots & \langle v_n, v_n\rangle \end{pmatrix}$

(m) $\exists v_1, \ldots, v_n \in \mathbb{R}^n : A = \sum_{i=1}^n v_i v_i^T$

(n) $\exists m \in \mathbb{N}_0 : \exists v_1, \ldots, v_m \in \mathbb{R}^n : A = \sum_{i=1}^m v_i v_i^T$

[2] Im Beweis dieses Satzes benutzen wir den Fundamentalsatz der Algebra 4.2.12 [$\rightarrow$ 4.2.13(g)].

Beweis. (a) $\Longleftrightarrow$ (b) ist trivial, da A symmetrisch ist.

(b) $\Longleftrightarrow$ (c) ist klar nach Definition 14.1.1 der Sylvester-Signatur.

(c) $\Longleftrightarrow$ (d) folgt aus 14.1.2(a).

(d) $\Longleftrightarrow$ (e) „$\Longrightarrow$“ Sind $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ die Eigenwerte von A mit algebraischen Vielfachheiten [$\rightarrow$ 10.1.14, 11.3.10], so gilt $\chi_A = \prod_{i=1}^{n}(\lambda_i - X)$ und daher $\chi_A(-X) = \prod_{i=1}^{n}(\lambda_i + X)$. Die Koeffizienten von $\chi_A(-X)$ sind daher Summen von Produkten der λ_i.

„$\Longleftarrow$“ Sei $\lambda \in \mathbb{R}$ ein Eigenwert von A. Dann $\det(A - \lambda I_n) = \chi_A(\lambda) = 0$. Setzt man also $-\lambda$ anstelle von X in das Polynom $0 \neq \det(A + XI_n)$ ein, so erhält man 0. Hat dieses Polynom nur nichtnegative Koeffizienten, so folgt $-\lambda \leq 0$.

(e) $\Longleftrightarrow$ (f) „$\Longleftarrow$“ Schreibt man $\det(A + XI_n) = X^n + a_{n-1}X^{n-1} + \ldots + a_0$ mit $a_0, \ldots, a_{n-1} \in K$, so sieht man mit scharfem Auge direkt an der Definition einer Determinante 9.1.9, dass

$$a_i = \sum_{\substack{I \subseteq \{1,\ldots,n\} \\ \#(\{1,\ldots,n\} \setminus I) = i}} \det(A_I)$$

für $i \in \{0, \ldots, n-1\}$.

„$\Longrightarrow$“ Gelte (e). Sei $\emptyset \neq I \subseteq \{1, \ldots, n\}$. Zu zeigen ist $\det(A_I) \geq 0$. Wegen (b) $\Longleftrightarrow$ (e) gilt dann auch (b). Insbesondere $\forall x \in \mathbb{R}^{\#I} : x^T A_I x \geq 0$. Wieder wegen (b) $\Longleftrightarrow$ (e) hat $\det(A_I + XI_{\#I}) \in \mathbb{R}[X]$ keine negativen Koeffizienten. Insbesondere ist der konstante Koeffizient $\det(A_I)$ dieses Polynoms ≥ 0.

Es ist nun die Äquivalenz der Aussagen (a)–(f) gezeigt.

(b) $\Longrightarrow$ (g) folgt wie in Bemerkung 13.5.10(c) angekündigt durch Inspektion des Beweises von Satz 13.5.9.

(g) $\Longrightarrow$ (h) Ist $A = P^T D P$ mit $P = \begin{pmatrix} 1 & & * \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \in \mathbb{R}^{n \times n}$ und $D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix} \in \mathbb{R}^{n \times n}$ mit $d_i \geq 0$, so ist $B := \begin{pmatrix} \sqrt{d_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{d_n} \end{pmatrix} P \in \mathbb{R}^{n \times n}$ und $A = B^T B$.

(h) $\Longrightarrow$ (i) $\Longrightarrow$ (j) ist trivial.

(j) $\Longrightarrow$ (b) Ist $A = B^T B$ mit $B \in \mathbb{R}^{m \times n}$, so gilt $x^T A x = x^T B^T B x = (Bx)^T Bx = \langle Bx, Bx \rangle \geq 0$ für alle $x \in \mathbb{R}^n$.

Es ist nun die Äquivalenz der Aussagen (a)–(j) gezeigt.

Die Äquivalenzen (i) $\Longleftrightarrow$ (k) und (j) $\Longleftrightarrow$ (l) ergeben sich sofort, indem man die v_i als die Spalten von B auffasst, denn für $v_1, \ldots, v_n \in \mathbb{R}^m$ gilt

$$\begin{pmatrix} v_1^T \\ \vdots \\ v_n^T \end{pmatrix} \begin{pmatrix} v_1 & \ldots & v_n \end{pmatrix} = \begin{pmatrix} v_1^T v_1 & \ldots & v_1^T v_n \\ \vdots & & \vdots \\ v_n^T v_1 & \ldots & v_n^T v_n \end{pmatrix}.$$

Die Äquivalenzen (i) $\iff$ (m) und (j) $\iff$ (m) ergeben sich, indem man die v_i^T als die Zeilen von B auffasst, denn für $v_1, \dots, v_m \in \mathbb{R}^n$ gilt

$$\begin{aligned}
\begin{pmatrix} v_1 & \dots & v_m \end{pmatrix} \begin{pmatrix} v_1^T \\ \vdots \\ v_m^T \end{pmatrix} &= \sum_{i=1}^{m} \begin{pmatrix} 0 & \dots & 0 & \overset{i}{\underset{\downarrow}{v_i}} & 0 \dots & 0 \end{pmatrix} \begin{pmatrix} v_1^T \\ \vdots \\ v_m^T \end{pmatrix} \\
&= \sum_{i=1}^{m} \begin{pmatrix} 0 & \dots & 0 & \overset{i}{\underset{\downarrow}{v_i}} & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ v_i^T \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i \\
&= \sum_{i=1}^{m} v_i v_i^T .
\end{aligned}$$

□

Satz 14.2.12. [→ 14.2.7] *Sei V ein endlichdimensionaler $\mathbb{R}$-Vektorraum, $n := \dim V$ und $q \in Q(V)$. Dann sind äquivalent:*

(a) *q ist pd.*

(b) *Die Sylvester-Signatur von q ist $(n, 0)$.*

(c) *Es gibt eine Basis $\ell_1, \dots, \ell_n$ von V^* mit $\forall v \in V : q(v) = \sum_{i=1}^{n} (\ell_i(v))^2$*

Beweis. (a) $\implies$ (b) folgt direkt aus der Definition der Sylvester-Signatur 14.1.1.

(b) $\implies$ (c) folgt ebenfalls aus dieser Definition zusammen mit Lemma 13.5.1.

(c) $\implies$ (a) Gelte (c) und sei $0 \neq v \in V$. Zu zeigen ist $q(v) > 0$. Da $V \to V^{**}$ injektiv ist [→ 13.1.14], gibt es $\ell \in V^*$ mit $\ell(v) \neq 0$. Wegen $\ell \in \operatorname{span}(\ell_1, \dots, \ell_n)$ gibt es $i \in \{1, \dots, n\}$ mit $\ell_i(v) \neq 0$. Daraus folgt $q(v) \geq (\ell_i(x))^2 > 0$. □

Satz 14.2.13. [→ 14.2.11] [3] *Sei $A \in S\mathbb{R}^{n \times n}$. Dann sind äquivalent:*

(a) *A ist pd.*

(b) $\forall x \in \mathbb{R}^n \setminus \{0\} : x^T A x > 0$

(c) *Die Sylvester-Signatur von A ist $(n, 0)$.*

(d) *Alle Eigenwerte von A sind positiv.*

[3] Im Beweis dieses Satzes benutzen wir den Fundamentalsatz der Algebra 4.2.12 [→ 4.2.13(g)].

(e) *Die Koeffizienten zu den Monomen* $1, X, \ldots, X^{n-1}$ *in* $\det(A + XI_n) = \chi_A(-X) \in \mathbb{R}[X]$ *sind positiv.*

(f) *Alle Leithauptminoren von* A *sind positiv.*

(g) *Alle Hauptminoren von* A *sind positiv.*

(h) A *besitzt eine Cholesky-Zerlegung* (P, D) *derart, dass alle Diagonaleinträge von* D *positiv sind.*

Beweis. Die Äquivalenz aller Aussagen mit Ausnahme von (f) zeigt man analog zum Beweis von Satz 14.2.11. Es ist $\underline{\text{(g)} \Longrightarrow \text{(f)}}$ trivial. Wir zeigen schließlich $\underline{\text{(f)} \Longrightarrow \text{(b)}}$ durch Induktion nach $n \in \mathbb{N}_0$:

$\underline{n = 0}$ Dann ist (b) die leere Aussage, da $\mathbb{R}^n = \mathbb{R}^0 = \{0\}$.

$\underline{n \to n+1 \quad (n \in \mathbb{N}_0)}$ Seien alle Leithauptminoren der Matrix $A \in S\mathbb{R}^{(n+1)\times(n+1)}$ positiv. Schreibt man

$$A = \left(\begin{array}{ccc|c} & & & a_1 \\ & B & & \vdots \\ & & & a_n \\ \hline a_1 & \ldots & a_n & c \end{array}\right)$$

mit $B \in S\mathbb{R}^{n\times n}$ und $a_1, \ldots, a_n, c \in \mathbb{R}$, so gilt nach Induktionsvoraussetzung $x^T B x > 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$ (denn insbesondere sind alle Leithauptminoren von B positiv). Wähle nun

$$0 \neq v \in \ker \underbrace{\left(\begin{array}{ccc} & & a_1 \\ & B & \vdots \\ & & a_n \end{array}\right)}_{\in \mathbb{R}^{n\times(n+1)}}.$$

Wegen $\ker B = \{0\}$ kann nicht $v \in \operatorname{span}(e_1, \ldots, e_n)$ gelten, wobei $\underline{e} = (e_1, \ldots, e_{n+1})$ die Standardbasis des $\mathbb{R}^{n+1}$ bezeichne. Dann ist $\underline{v} := (e_1, \ldots, e_n, v)$ eine Basis des $\mathbb{R}^{n+1}$ und es gilt $e_i^T A v = 0$ für $i \in \{1, \ldots, n\}$. Es folgt, dass $M(q_A, \underline{v})$ von der Form

$$M(q_A, \underline{v}) = \left(\begin{array}{ccc|c} & & & 0 \\ & B & & \vdots \\ & & & 0 \\ \hline 0 & \ldots & 0 & d \end{array}\right)$$

mit $d \in \mathbb{R}$ ist. Wegen $A = M(q_A, \underline{e}) \overset{13.3.10}{=} M(\underline{e}, \underline{v})^T M(q_A, \underline{v}) M(\underline{e}, \underline{v})$ gilt

$$0 < \det A \overset{9.1.15}{\underset{9.1.23}{=}} (\det M(\underline{e}, \underline{v}))^2 M(q_A, \underline{v}) \overset{9.1.11}{=} \underbrace{(\det M(\underline{e}, \underline{v}))^2}_{>0} \underbrace{(\det B)}_{>0} d$$

und daher $d > 0$. Nun gilt für alle $x \in \mathbb{R}^n$ und $y \in \mathbb{R}$

$$\begin{pmatrix} x \\ y \end{pmatrix}^T M(q_A, \underline{v}) \begin{pmatrix} x \\ y \end{pmatrix} = x^T B x + d y^2 > 0 \qquad \text{falls} \qquad \begin{pmatrix} x \\ y \end{pmatrix} \neq 0.$$

Somit ist $M(q_A, \underline{v})$ und damit q_A pd, das heißt A ist pd. □

Bemerkung 14.2.14. Wie man eine Cholesky-Zerlegung einer positiv semidefiniten Matrix berechnet, ist aus dem Beweis von 13.5.9 wegen Bemerkung 13.5.10(c) klar. Ist die Matrix sogar positiv definit, so kann auch der dortige Fall 1 nicht auftreten. Da im dortigen Fall 2 die Wahl der Linearform ℓ_1 zwingend ist, sieht man mit Hilfe von 13.5.2 leicht, dass *die* Cholesky-Zerlegung einer *positiv definiten* Matrix *eindeutig* ist.

Die in 13.5.12 bewiesene Diagonalisierung quadratischer Formen über beliebigen Körpern mit $0 \neq 2$ kann über dem Körper der reellen Zahlen zu folgender in §11.3 in einer anderen Sprache formulierten Aussage verschärft werden („simultane Diagonalisierung"):

Satz 14.2.15. [4] *Seien V ein endlichdimensionaler $\mathbb{R}$-Vektorraum und $q_1, q_2 \in Q(V)$. Ist q_1 pd oder nd, so gibt es eine geordnete Basis $\underline{v}$ von V derart, dass $M(q_1, \underline{v})$ und $M(q_2, \underline{v})$ beide Diagonalgestalt haben.*

Beweis. Œ q_1 pd. Es ist b_{q_1} [→ 13.4.6] ein Skalarprodukt auf V vermöge dessen V zu einem Vektorraum mit Skalarprodukt wird [→ 11.1.1]. Wähle eine ONB $\underline{w}$ von V [→ 11.2.5, 11.2.15]. Wähle $f \in \mathrm{End}(V)$ mit $M(f, \underline{w}) = M(q_2, \underline{w})$ (nämlich $f = \mathrm{vec}_{\underline{w}} \circ f_{M(q_2,\underline{w})} \circ \mathrm{coord}_{\underline{w}}$). Da $M(f, \underline{w})$ symmetrisch (also selbstadjungiert [→ 11.3.3]) und $\underline{w}$ eine ONB ist, ist f nach 11.3.4 selbstadjungiert. Nach Satz 11.3.9 gibt es eine ONB $\underline{v}$ von V, die aus Eigenvektoren von f besteht. Dann ist $M(q_1, \underline{v}) = M(b_{q_1}, \underline{v})$ die Einheitsmatrix (da $\underline{v}$ ONB) und

$$\begin{array}{rcl} M(q_2, \underline{v}) & = & M(\underline{v}, \underline{w})^T M(q_2, \underline{w}) M(\underline{v}, \underline{w}) \\ & = & M(\underline{v}, \underline{w})^T M(f, \underline{w}) M(\underline{v}, \underline{w}) \\ & \underset{11.2.26}{\overset{\underline{v},\underline{w}\ \mathrm{ONB}}{=}} & M(\underline{v}, \underline{w})^{-1} M(f, \underline{w}) M(\underline{v}, \underline{w}) \\ & \underset{7.2.11}{=} & M(\underline{w}, \underline{v}) M(f, \underline{w}) M(\underline{v}, \underline{w}) \underset{7.2.5}{=} M(f, \underline{v}) \end{array}$$

von Diagonalgestalt, da $\underline{v}$ aus Eigenvektoren von f besteht. □

Jedoch ist obige simultane Diagonalisierung viel schwieriger zu berechnen (siehe Beweis von Satz 11.3.9) als einfach nur eine Diagonalisierung (siehe §13.5).

[4] Im Beweis dieses Satzes benutzen wir den Fundamentalsatz der Algebra 4.2.12 [→ 4.2.13(g)].

15 Skalarprodukte

Dieses Kapitel ist eine Fortsetzung von §11. Wie dort, so sei auch hier stets $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

15.1 Die adjungierte Abbildung

Proposition und Definition 15.1.1. *Seien V und W $\mathbb{K}$-Vektorräume mit Skalarprodukt und $f\colon V \to W$ linear. Dann gibt es zu jedem $w \in W$ höchstens ein $v' \in V$ mit $\forall v \in V : \langle f(v), w\rangle = \langle v, v'\rangle$. Setzt man $W' := \{w \in W \mid \exists v' \in V : \forall v \in V : \langle f(v), w\rangle = \langle v, v'\rangle\}$, so gibt es also genau eine Abbildung $f^*\colon W' \to V$ mit $\forall v \in V\colon \langle f(v), w\rangle = \langle v, f^*(w)\rangle$. Man nennt f^* die zu f* adjungierte Abbildung. *Es ist W' ein Untervektorraum von W und f^* linear.*

Beweis. Zu zeigen:

(a) Eindeutigkeit von v'.

(b) $\forall w_1, w_2 \in W' : (w_1 + w_2 \in W' \;\&\; f^*(w_1 + w_2) = f^*(w_1) + f^*(w_2))$

(c) $\forall w \in W' : \forall \lambda \in \mathbb{K} : (\lambda w \in W' \;\&\; f^*(\lambda w) = \lambda f^*(w))$

Zu (a). Sei $w \in W$. Sind $v_1', v_2' \in V$ mit $\forall v \in V : \langle v, v_1'\rangle = \langle f(v), w\rangle = \langle v, v_2'\rangle$, so folgt $\forall v \in V : \langle v, v_1' - v_2'\rangle = 0$, insbesondere $\langle v_1' - v_2', v_1' - v_2'\rangle = 0$, also $v_1' = v_2'$.

Zu (b). Seien $w_1, w_2 \in W'$. Dann gilt für alle $v \in V$ $\langle f(v), w_1 + w_2\rangle = \langle f(v), w_1\rangle + \langle f(v), w_2\rangle = \langle v, f^*(w_1)\rangle + \langle v, f^*(w_2)\rangle = \langle v, f^*(w_1) + f^*(w_2)\rangle$.

Zu (c). Sei $w \in W'$ und $\lambda \in \mathbb{K}$. Dann gilt für alle $v \in V$

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

□

Beispiel 15.1.2. Sei $V := C^\infty([0,1], \mathbb{R})$ der Vektorraum der unendlich oft differenzierbaren reellen Funktionen auf $[0,1]$ mit dem durch

$$\langle f, g\rangle := \int_0^1 fg \qquad (f, g \in V)$$

definierten Skalarprodukt. Die Ableitung $D\colon V \to V,\ f \mapsto f'$ ist eine lineare Abbildung. Ist g im Definitionsbereich von D^*, so gilt

$$\forall f \in V : \int_0^1 f'g = \int_0^1 fD^*(g).$$

Andererseits gilt gemäß partieller Integration auch

$$\forall f \in V : \int_0^1 f'g = [fg]_0^1 - \int_0^1 fg'$$

und daher $\forall f \in V : \int_0^1 f(g' + D^*(g)) = [fg]_0^1$. Durch Einsetzen von „Nadelfunktionen“ (siehe etwa Wikipedia-Eintrag zu “bump function” [male Bild]) für f sieht man nun $D^*(g) = -g'$. Daher ist der Definitionsbereich von D^* gleich

$$\begin{aligned}\left\{g \in V \mid \forall f \in V : \int_0^1 f'g = \int_0^1 f(-g')\right\} &= \{g \in V \mid \forall f \in V : [fg]_0^1 = 0\} \\ &= \{g \in V \mid g(0) = g(1) = 0\}.\end{aligned}$$

Es gilt also $D^* \colon \{g \in V \mid g(0) = g(1) = 0\} \to V,\ g \mapsto -g'$.

Erinnerung 15.1.3. [→ 11.3.1] Sei V ein Vektorraum mit Skalarprodukt und $f \colon V \to V$ linear. Dann heißt f selbstadjungiert, wenn $\forall v, w \in V \colon \langle f(v), w\rangle = \langle v, f(w)\rangle$.

Proposition 15.1.4. *Sei V ein Vektorraum mit Skalarprodukt und $f \colon V \to V$ linear. Dann ist f selbstadjungiert genau dann, wenn $f = f^*$ (was natürlich beinhaltet, dass f^* auf V definiert ist).*

Beispiel 15.1.5. Sei $V := \{f \in C^\infty([0,1], \mathbb{C}) \mid f(0) = f(1)\}$ mit dem durch

$$\langle f, g\rangle := \int_0^1 f^* g \qquad (f, g \in V)$$

gegebenen Skalarprodukt, wobei für $f, g \in V$ jeweils f^* die zu f punktweise komplex-konjugierte Funktion und fg das punktweise Produkt von f und g sei. Betrachte $T \colon V \to V,\ f \mapsto \overset{\circ}{\imath} f'$. Es gilt für alle $f, g \in V$

$$\begin{aligned}\langle T(f), g\rangle = \int_0^1 (\overset{\circ}{\imath} f'(x))^* g(x)\,dx = -\overset{\circ}{\imath} \int_0^1 (f^*)' g &\overset{\text{partielle Integration}}{=} -\overset{\circ}{\imath} \left(\underbrace{[f^* g]_0^1}_{\substack{=0,\\ \text{da } f,g \in V}} - \int_0^1 f^* g' \right) \\ &= \int_0^1 f^* T(g) = \langle f, T(g)\rangle.\end{aligned}$$

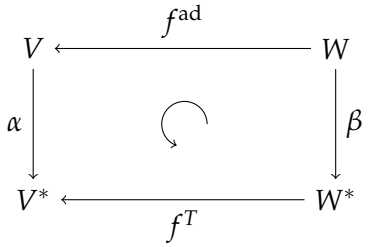
Daher ist $T = T^*$.

Satz 15.1.6. *Seien V und W $\mathbb{K}$-Vektorräume mit Skalarprodukt. Sei $f \colon V \to W$ linear und* $\dim V < \infty$. *Dann ist f^* auf ganz W definiert.*

Beweis. Wähle mit 11.2.5 eine ONB $\underline{v} = (v_1, \ldots, v_n)$ von V. Dann gilt für jede lineare Abbildung $g \colon W \to V$:

$$\begin{aligned} g = f^* &\iff \forall v \in V : \forall w \in W : \langle f(v), w\rangle = \langle v, g(w)\rangle \\ &\iff \forall i \in \{1, \ldots, n\} : \forall w \in W : \langle f(v_i), w\rangle = \langle v_i, g(w)\rangle \\ &\iff \forall w \in W : \sum_{i=1}^n \langle f(v_i), w\rangle v_i = \sum_{i=1}^n \langle v_i, g(w)\rangle v_i \overset{11.2.13}{=} g(w).\end{aligned}$$

Bemerkung 15.1.7. Die Notation f^* hatten wir in 13.1.8 anders verwendet, nämlich für die zu einer linearen Abbildung gehörige duale Abbildung $f^*\colon W^* \to V^*$. Man verwendet fast nie die adjungierte und die duale Abbildung gleichzeitig und aus dem Kontext ist fast immer klar, welche gemeint ist. Wenn V und W endlichdimensionale $\mathbb{R}$-Vektorräume mit Skalarprodukt sind und $f\colon V \to W$ linear ist, dann sind ausserdem die zu f dualen Abbildungen $f^T := f^*\colon W^* \to V^*$ und die zu f adjungierte Abbildung $f^{\mathrm{ad}} := f^*\colon W \to V$ „im Prinzip dieselben", denn das Diagramm



mit den kanonischen Isomorphien [→ 13.1.7]

$$\alpha\colon V \to V^*,\ v_1 \mapsto (v_2 \mapsto \langle v_1, v_2\rangle) \qquad \text{und}$$
$$\beta\colon W \to W^*,\ w_1 \mapsto (w_2 \mapsto \langle w_1, w_2\rangle)$$

kommutiert. In der Tat: Es gilt für alle $w \in W$ und $v \in V$

$$\begin{aligned}(\alpha(f^{\mathrm{ad}}(w)))(v) &= \langle v, f^{\mathrm{ad}}(w)\rangle = \langle f(v), w\rangle = \langle w, f(v)\rangle \\ &= (\beta(w))(f(v)) = ((\beta(w)) \circ f)(v) = (f^T(\beta(w)))(v),\end{aligned}$$

das heißt $\alpha \circ f^{\mathrm{ad}} = f^T \circ \beta$.

□

Lemma 15.1.8. *Sei $A \in \mathbb{K}^{m\times n}$, $x \in \mathbb{K}^n$ und $y \in \mathbb{K}^m$. Dann $\langle Ax, y\rangle = \langle x, A^*y\rangle$.*

Beweis. $\langle Ax, y\rangle = (Ax)^*y = x^*A^*y = \langle x, A^*y\rangle$ □

Proposition 15.1.9. *Seien V und W endlichdimensionale $\mathbb{K}$-Vektorräume mit Skalarprodukt und $f\colon V \to W$ linear. Sei $\underline{v}$ eine ONB von V und $\underline{w}$ eine ONB von W. Dann gilt*

$$M(f^*, \underline{w}, \underline{v}) = M(f, \underline{v}, \underline{w})^*.$$

Beweis. Es ist

$$(*) \qquad f^* = \mathrm{vec}_{\underline{v}} \circ f_{M(f,\underline{v},\underline{w})^*} \circ \mathrm{coord}_{\underline{w}}$$

zu zeigen. Es gilt

$$\begin{aligned}(*) &\iff \forall v \in V : \forall w \in W : \langle f(v), w\rangle = \langle v, \mathrm{vec}_{\underline{v}}(M(f,\underline{v},\underline{w})^* \,\mathrm{coord}_{\underline{w}}(w))\rangle \\ &\overset{11.2.24}{\iff} \forall v \in V : \forall w \in W : \langle \overbrace{\mathrm{coord}_{\underline{w}}(f(v))}^{M(f,\underline{v},\underline{w})\,\mathrm{coord}_{\underline{v}}(v)}, \mathrm{coord}_{\underline{w}}(w)\rangle = \\ &\qquad\qquad \langle \mathrm{coord}_{\underline{v}}(v), M(f,\underline{v},\underline{w})^* \,\mathrm{coord}_{\underline{w}}(w)\rangle.\end{aligned}$$

□

Proposition 15.1.10. *Seien U, V und W endlichdimensionale $\mathbb{K}$-Vektorräume mit Skalarprodukt, $\lambda \in \mathbb{K}$ und $f, f_1, f_2 \colon U \to V$ sowie $g \colon V \to W$ linear. Dann gilt $(f_1 + f_2)^* = f_1^* + f_2^*$, $(\lambda f)^* = \lambda^* f^*$, $(g \circ f)^* = f^* \circ g^*$, $\mathrm{id}_V^* = \mathrm{id}_V$ und $f^{**} = f$.*

Beweis. Zu zeigen:

(a) $\forall u \in U : \forall v \in V : \langle (f_1 + f_2)(u), v \rangle = \langle u, (f_1^* + f_2^*)(v) \rangle$

(b) $\forall u \in U : \forall v \in V : \langle (\lambda f)(u), v \rangle = \langle u, (\lambda^* f^*)(v) \rangle$

(c) $\forall u \in U : \forall w \in W : \langle (g \circ f)(u), w \rangle = \langle u, (f^* \circ g^*)(w) \rangle$

(d) $\forall v_1, v_2 \in V : \langle \mathrm{id}_V(v_1), v_2 \rangle = \langle v_1, \mathrm{id}_V(v_2) \rangle$

(e) $\forall v \in V : \forall u \in U : \langle f^*(v), u \rangle = \langle v, f(u) \rangle$

Zu (a). Seien $u \in U$ und $v \in V$. Dann $\langle (f_1 + f_2)(u), v \rangle = \langle f_1(u), v \rangle + \langle f_2(u), v \rangle = \langle u, f_1^*(v) \rangle + \langle u, f_2^*(v) \rangle = \langle u, f_1^*(v) + f_2^*(v) \rangle = \langle u, (f_1^* + f_2^*)(v) \rangle$.

Zu (b). Seien $u \in U$ und $v \in V$. Dann $\langle (\lambda f)(u), v \rangle = \langle \lambda(f(u)), v \rangle = \lambda^* \langle f(u), v \rangle = \langle f(u), \lambda^* v \rangle = \langle u, f^*(\lambda^* v) \rangle = \langle u, \lambda^*(f^*(v)) \rangle = \langle u, (\lambda^* f^*)(v) \rangle$.

Zu (c). Seien $u \in U$ und $w \in W$. Dann $\langle (g \circ f)(u), w \rangle = \langle g(f(u)), w \rangle = \langle f(u), g^*(w) \rangle = \langle u, f^*(g^*(w)) \rangle = \langle u, (f^* \circ g^*)(w) \rangle$.

(d) ist trivial.

Zu (e). Seien $u \in U$ und $v \in V$. Dann $\langle f^*(v), u \rangle = \langle u, f^*(v) \rangle^* = \langle f(u), v \rangle^* = \langle v, f(u) \rangle$.

Alternativ kann man den Beweis mit Hilfe von Proposition 15.1.9 durch Rückführung auf die entsprechenden Tatsachen für Matrizen führen. □

Proposition 15.1.11. *Seien V und W $\mathbb{K}$-Vektorräume mit Skalarprodukt und $f \colon V \to W$ linear. Dann gilt* $\ker(f^*) = (\operatorname{im} f)^\perp$.

Beweis. Für $w \in W$ gilt gemäß Definition 15.1.1 der adjungierten Abbildung

$$w \in \ker(f^*) \iff \forall v \in V : \langle f(v), w \rangle = \langle v, 0 \rangle.$$

□

Satz 15.1.12. *Seien V und W $\mathbb{K}$-Vektorräume mit Skalarprodukt und $f \colon V \to W$ linear. Dann sind äquivalent:*

(a) *f ist ein Isomorphismus von Vektorräumen mit Skalarprodukt.*

(b) *f^* ist auf ganz W definiert und es gilt sowohl $f^* \circ f = \mathrm{id}_V$ als auch $f \circ f^* = \mathrm{id}_W$.*

(c) *f^* ist eine Bijektion $W \to V$ mit $(f^*)^{-1} = f$.*

Beweis. (c) ist nur eine Umformulierung von (b).

(a) $\Longrightarrow$ (b) Gelte (a). Sei $v' \in V$. Setze $w := f(v')$. Wegen $\langle f(v), w\rangle \overset{(a)}{=} \langle v, v'\rangle$ für alle $v \in V$ ist f^* in w definiert und es gilt $f^*(w) = v'$, das heißt $f^*(f(v')) = v'$. Da $v' \in V$ beliebig war, ist f^* auf $\operatorname{im} f \overset{(a)}{=} W$ definiert und es gilt $f^* \circ f = \mathrm{id}_V$. Weiter gilt $f^* = f^* \circ (f \circ f^{-1}) = (f^* \circ f) \circ f^{-1} = f^{-1}$ und daher auch $f \circ f^* = f \circ f^{-1} = \mathrm{id}_W$.

(b) $\Longrightarrow$ (a) Gelte (b). Dann hat f eine Umkehrabbildung und ist bijektiv. Da f auch linear ist, ist f ein Isomorphismus von Vektorräumen. Es bleibt zu zeigen, dass für alle $v, v' \in V$ gilt $\langle f(v), f(v')\rangle = \langle v, v'\rangle$. Seien also $v, v' \in V$. Dann $\langle f(v), f(v')\rangle = \langle v, f^*(f(v'))\rangle = \langle v, (f^* \circ f)(v')\rangle = \langle v, v'\rangle$. □

15.2 Normale Abbildungen

Definition 15.2.1. Sei V ein Vektorraum mit Skalarprodukt und $f\colon V \to V$ linear. Dann heißt f *normal*, wenn f^* auf ganz V definiert ist [$\to$ 15.1.1] und $f \circ f^* = f^* \circ f$.

Beispiel 15.2.2. Sei V ein Vektorraum mit Skalarprodukt und $f\colon V \to V$ linear.

(a) Ist f selbstadjungiert (das heißt $f = f^*$), so ist f normal, denn $f \circ f^* = f^2 = f^* \circ f$.

(b) Ist f ein Automorphismus von Vektorräumen mit Skalarprodukt [$\to$ 11.2.19], so ist f ebenfalls normal, denn $f \circ f^* = \mathrm{id}_V = f^* \circ f$.

Lemma 15.2.3. *Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum mit Skalarprodukt, $\lambda \in \mathbb{K}$ und $f\colon V \to V$ normal. Dann gilt* $\ker(f - \lambda \,\mathrm{id}_V) = \ker(f^* - \lambda^* \,\mathrm{id}_V)$.

Beweis. Wegen

$$\begin{array}{rcl}(f - \lambda \,\mathrm{id}_V) \circ (f - \lambda \,\mathrm{id}_V)^* & \overset{15.1.10}{=} & (f - \lambda \,\mathrm{id}_V) \circ (f^* - \lambda^* \,\mathrm{id}_V) \\ & = & f \circ f^* - \lambda f^* - \lambda^* f + \lambda\lambda^* \,\mathrm{id}_V \\ & = & f^* \circ f - \lambda f^* - \lambda^* f + \lambda^* \lambda \,\mathrm{id}_V \\ & = & (f^* - \lambda^* \,\mathrm{id}_V) \circ (f - \lambda \,\mathrm{id}_V)\end{array}$$

ist auch $f - \lambda \,\mathrm{id}_V$ normal. Daher reicht es, $\ker(f) = \ker(f^*)$ zu zeigen. Dies folgt aus $\|f(v)\|^2 = \langle f(v), f(v)\rangle = \langle v, f^*(f(v))\rangle = \langle v, (f^* \circ f)(v)\rangle = \langle v, (f \circ f^*)(v)\rangle = \langle v, f(f^*(v))\rangle = \langle f^*(v), f^*(v)\rangle = \|f^*(v)\|^2$ für alle $v \in V$. □

Satz 15.2.4. [5][$\to$ 11.3.9] *Sei V ein endlichdimensionaler $\mathbb{C}$-Vektorraum mit Skalarprodukt und $f\colon V \to V$ linear. Dann sind folgende Aussagen äquivalent:*

(a) *f ist normal.*

(b) *V hat eine ONB, die aus Eigenvektoren von f besteht.*

(c) *Es gibt eine ONB $\underline{v}$ von V derart, dass $M(f, \underline{v})$ Diagonalgestalt hat.*

Beweis. (b) $\Longrightarrow$ (c) ist klar.

(c) $\Longrightarrow$ (a) Sei $\underline{v}$ eine ONB von V mit $M(f, \underline{v})$ in Diagonalgestalt. Nach 15.1.6 und 7.1.8 reicht es $M(f \circ f^*, \underline{v}) = M(f^* \circ f, \underline{v})$ zu zeigen. Es gilt aber

$$M(f \circ f^*, \underline{v}) = M(f, \underline{v}) M(f^*, \underline{v}) \overset{\underline{v}\ \mathrm{ONB}}{\underset{15.1.9}{=}} M(f, \underline{v}) M(f, \underline{v})^* \overset{\text{Diagonalgestalt}}{=} M(f, \underline{v})^* M(f, \underline{v})$$
$$\overset{\underline{v}\ \mathrm{ONB}}{\underset{15.1.9}{=}} M(f^*, \underline{v}) M(f, \underline{v}) = M(f^* \circ f, \underline{v}).$$

(a) $\Longrightarrow$ (b) Induktion nach $n := \dim V \in \mathbb{N}_0$:
$n = 0$ nichts zu zeigen

[5] Im Beweis dieses Satzes benutzen wir den Fundamentalsatz der Algebra 4.2.12 [$\to$ 4.2.13(g)].

$n-1 \to n \quad (n \in \mathbb{N})$ Sei $f\colon V \to V$ normal. Wegen $\deg \chi_f = n \geq 1$ gibt es nach dem Fundamentalsatz der Algebra einen Eigenwert λ von f. Wähle dazu einen Eigenvektor $u \in V$, das heißt $u \in V \setminus \{0\}$ mit $f(u) = \lambda u$. Setze $U := \operatorname{span}(u)$. Es gilt $f(U^\perp) \subseteq U^\perp$ und $f^*(U^\perp) \subseteq U^\perp$, denn ist $v \in U^\perp$ und $u \in U$, so gilt

$$\begin{aligned} \langle f(v), u\rangle = \langle v, f^*(u)\rangle &\overset{15.2.3}{=} \langle v, \lambda^* u\rangle = \lambda^* \langle v, u\rangle = \lambda^* \cdot 0 = 0 \qquad \text{und} \\ \langle f^*(v), u\rangle = \langle v, f(u)\rangle &= \langle v, \lambda u\rangle = \lambda \langle v, u\rangle = \lambda \cdot 0 = 0. \end{aligned}$$

Betrachte nun $f|_{U^\perp}\colon U^\perp \to U^\perp$ und $f^*|_{U^\perp}\colon U^\perp \to U^\perp$. Anhand von 15.1.1 sieht man leicht, dass $f|_{U^\perp}^* = f^*|_{U^\perp}$ gilt. Daher hat man

$$f|_{U^\perp}^* \circ f|_{U^\perp} = f^*|_{U^\perp} \circ f|_{U^\perp} = (f^* \circ f)|_{U^\perp} = (f \circ f^*)|_{U^\perp} = f|_{U^\perp} \circ f^*|_{U^\perp} = f|_{U^\perp} \circ f|_{U^\perp}^*,$$

weswegen $f|_{U^\perp}$ ebenfalls normal ist. Wegen $\dim(U^\perp) \overset{11.2.17}{=} n-1$ gibt es nach IV eine ONB $(v_2, \ldots, v_n)$ von $U^\perp$, die aus Eigenvektoren von f besteht. Setzt man $v_1 := \frac{u}{\|u\|}$, so erhält man eine ONB $(v_1, \ldots, v_n)$ von V, die aus Eigenvektoren von f besteht. □

Korollar 15.2.5. [6] *Sei V ein endlichdimensionaler $\mathbb{C}$-Vektorraum mit Skalarprodukt und $f\colon V \to V$ normal. Dann ist f diagonalisierbar* [→ 10.3.2(a)].

Korollar 15.2.6. [7] *Sei V ein endlichdimensionaler $\mathbb{C}$-Vektorraum mit Skalarprodukt und $f\colon V \to V$ linear. Dann sind die folgenden Aussagen äquivalent:*

(a) f ist ein orthogonal (auch unitär genannt [→ 11.2.19]*), das heißt*

$$\forall v, w \in V : \langle f(v), f(w)\rangle = \langle v, w\rangle.$$

(b) f ist ein Isomorphismus von Vektorräumen mit Skalarprodukt.

(c) V besitzt eine ONB, die aus Eigenvektoren von f zu Eigenwerten vom Betrag 1 besteht.

(d) Es gibt eine ONB $\underline{v}$ von V derart, dass $M(f, \underline{v})$ Diagonalgestalt mit Diagonaleinträgen vom Betrag 1 hat.

Beweis. (a) $\Longrightarrow$ (b) ist klar, denn wenn f injektiv ist, so auch surjektiv wegen $\dim V < \infty$ [→ 7.2.12].

(b) $\Longrightarrow$ (c) folgt sofort aus Satz 15.2.4, denn wenn (b) gilt, so ist f ist normal und alle Eigenwerte von f haben Absolutbetrag 1: Sei $\lambda \in \mathbb{C}$ und $v \in V \setminus \{0\}$ mit $f(v) = \lambda v$. Dann $|\lambda|\|v\| = \|\lambda v\| = \|f(v)\| = \|v\|$ und daher $|\lambda| = 1$.

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

[6] Im Beweis dieses Korollars benutzen wir den Fundamentalsatz der Algebra 4.2.12 [→ 4.2.13(g)].
[7] Im Beweis dieses Korollars benutzen wir den Fundamentalsatz der Algebra 4.2.12 [→ 4.2.13(g)].

(d) $\Longrightarrow$ (a) Sei $\underline{v}$ eine ONB von V mit $M(f,\underline{v}) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ und $|\lambda_i| = 1$ für $i \in \{1,\dots,n\}$. Dann

$$M(f\circ f^*,\underline{v}) = M(f,\underline{v})M(f^*,\underline{v}) \overset{\underline{v}\text{ ONB}}{\underset{15.1.9}{=}} M(f,\underline{v})M(f,\underline{v})^*$$
$$= \begin{pmatrix} \lambda_1\lambda_1^* & & 0 \\ & \ddots & \\ 0 & & \lambda_n\lambda_n^* \end{pmatrix} = \begin{pmatrix} |\lambda_1|^2 & & 0 \\ & \ddots & \\ 0 & & |\lambda_n|^2 \end{pmatrix} = I_n$$

und daher $f\circ f^* = \mathrm{id}_V$. Analog folgt $f^*\circ f = \mathrm{id}_V$ [alternativ: aus $f\circ f^* = \mathrm{id}_V$ folgt f^* injektiv und damit f^* bijektiv, wodurch $f^*\circ f = f^*\circ f\circ f^*\circ (f^*)^{-1} = f^*\circ \mathrm{id}_V\circ(f^*)^{-1} = \mathrm{id}_V$]. □

Um dem Leser zu helfen, die Resultate einzuordnen, formulieren wir Satz 11.3.9 noch einmal leicht anders. Er wurde damals durch eine kleine Variante des Beweises von 15.2.4 gezeigt, aber zumindest für $\mathbb{K} = \mathbb{C}$ erhält man ihn auch leicht als Korollar aus dem obigen Satz 15.2.4.

Satz 15.2.7. [8][→ 11.3.9] *Sei V ein endlichdimensionaler $\mathbb{K}$-Vektorraum mit Skalarprodukt und $f\colon V\to V$ linear. Dann sind folgende Aussagen äquivalent:*

(a) *f ist selbstadjungiert (im Fall $\mathbb{K} = \mathbb{C}$ auch hermitesch genannt* [→ 11.3.3]*).*

(b) *V hat eine ONB, die aus Eigenvektoren von f zu reellen Eigenwerten besteht.*

(c) *Es gibt eine ONB $\underline{v}$ von V derart, dass $M(f,\underline{v})$ eine reelle Diagonalmatrix ist.*

Notation 15.2.8. Für den Rest dieses Abschnitts notieren wir

$$\overline{x} := \begin{pmatrix} x_1^* \\ \vdots \\ x_n^* \end{pmatrix} = (x^*)^T \qquad \text{für} \qquad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{C}^n.$$

Lemma 15.2.9. *Seien $x,y\in\mathbb{R}^n$ und $w\in\mathbb{C}^n$ mit $\sqrt{2}w = x + \overset{\circ}{i}y$. Ist dann $(w,\overline{w})$ ein ONS in $\mathbb{C}^n$, so ist (x,y) eine ONB von* $\mathrm{span}(w,\overline{w})$ *in $\mathbb{C}^n$.*

[8] Im Beweis dieses Satzes benutzen wir den Fundamentalsatz der Algebra 4.2.12 [→ 4.2.13(g)].

Beweis. Sei $(w, \overline{w})$ ein ONS in $\mathbb{C}^n$, das heißt $\langle w, w\rangle = \langle \overline{w}, \overline{w}\rangle = 1$ und $\langle w, \overline{w}\rangle = 0$. Dann

$$\begin{aligned} 1 = \langle w, w\rangle &= \frac{1}{2}\langle x + \mathring{\imath} y, x + \mathring{\imath} y\rangle = \frac{1}{2}\Big(\langle x, x\rangle + \mathring{\imath}\langle x, y\rangle - \mathring{\imath}\langle x, y\rangle - \mathring{\imath}^2\langle y, y\rangle\Big) \\ &= \frac{1}{2}(\langle x, x\rangle + \langle y, y\rangle + \mathring{\imath}(\underbrace{\underbrace{\langle x, x\rangle}_{\in\mathbb{R}} - \overline{\langle x, y\rangle}}_{=0})) = \frac{1}{2}(\langle x, x\rangle + \langle y, y\rangle) \qquad \text{und} \\ 0 = \langle w, \overline{w}\rangle &= \frac{1}{2}\langle x + \mathring{\imath} y, x - \mathring{\imath} y\rangle = \frac{1}{2}\Big(\langle x, x\rangle - \mathring{\imath}\langle x, y\rangle + \mathring{\imath}\langle x, y\rangle + \mathring{\imath}^2\langle y, y\rangle\Big) \\ &= \frac{1}{2}(\langle x, x\rangle - \langle y, y\rangle - 2\mathring{\imath}\langle x, y\rangle), \end{aligned}$$

woraus folgt $0 = \frac{1}{2}(\langle x, x\rangle - \langle y, y\rangle)$ und $\langle x, y\rangle = 0$. Insgesamt folgt $\langle x, x\rangle = \langle y, y\rangle = 1$ und $\langle x, y\rangle = 0$, weshalb (x, y) ein ONS in $\mathbb{C}^n$ ist mit $w, \overline{w} \in \operatorname{span}_{\mathbb{C}}(x, y)$. Wegen $\operatorname{span}_{\mathbb{C}}(w, \overline{w}) \subseteq \operatorname{span}_{\mathbb{C}}(x, y)$ und $\dim(\operatorname{span}_{\mathbb{C}}(w, \overline{w})) = 2$ folgt

$$\operatorname{span}_{\mathbb{C}}(w, \overline{w}) = \operatorname{span}_{\mathbb{C}}(x, y).$$

□

Satz 15.2.10. [9][→ 11.3.9] *Sei V ein endlichdimensionaler $\mathbb{R}$-Vektorraum mit Skalarprodukt und $f\colon V \to V$ linear. Dann sind folgende Aussagen äquivalent:*

(a) *f ist normal* [→ 15.2.1].

(b) *Es gibt eine ONB $\underline{v}$ von V derart, dass $M(f, \underline{v})$ von der Gestalt*

$$\begin{pmatrix} \lambda_1 & & & & \\ & \ddots & & & \\ & & \lambda_k & & \\ & & & \boxed{\begin{matrix} a_1 & b_1 \\ -b_1 & a_1 \end{matrix}} & \\ & & & & \ddots \\ & & & & & \boxed{\begin{matrix} a_\ell & b_\ell \\ -b_\ell & a_\ell \end{matrix}} \end{pmatrix}$$

ist.

Beweis. (b) $\Longrightarrow$ (a) Für $a, b \in \mathbb{R}$ gilt $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}^* \begin{pmatrix} a & b \\ -b & a \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}\begin{pmatrix} a & b \\ -b & a \end{pmatrix} = \begin{pmatrix} a^2 + b^2 & 0 \\ 0 & a^2 + b^2 \end{pmatrix} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}\begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}\begin{pmatrix} a & b \\ -b & a \end{pmatrix}^*$ und der Beweis geht daher wie der Beweis von (c) $\Longrightarrow$ (a) in Satz 15.2.4.

[9] Im Beweis dieses Satzes benutzen wir den Fundamentalsatz der Algebra 4.2.12 [→ 4.2.13(g)].

(a) $\Longrightarrow$ (b) Nach 11.2.23 ist V als Vektorraum mit Skalarprodukt isomorph zu $\mathbb{R}^n$ ($n := \dim V$) mit dem Standardskalarprodukt. Daher Œ $V = \mathbb{R}^n$ und $f = f_A$ mit $A := M(f, \underline{e}) \in \mathbb{R}^{n\times n}$. Betrachte $g\colon \mathbb{C}^n \to \mathbb{C}^n,\ x \mapsto Ax$. Offensichtlich gilt $A = M(g, \underline{e})$ und wegen $A^*A = AA^*$ ist g normal. Daher gibt es eine ONB $\underline{w} = (w_1, \ldots, w_n)$ von $\mathbb{C}^n$, die aus Eigenvektoren von g besteht. Bezeichne $\lambda_i \in \mathbb{C}$ den zu w_i gehörigen Eigenwert von g ($i \in \{1, \ldots, n\}$). Dann gilt $\chi_g = \det(A - XI_n) = (-1)^n \prod_{i=1}^n (X - \lambda_i)$ und wegen $\chi_g \in \mathbb{R}[X]$ (da $A \in \mathbb{R}^{n\times n}$) auch $\chi_g = (-1)^n \prod_{i=1}^n (X - \lambda_i^*)$. Es folgt [→ 10.1.13] dass die Tupel $(\lambda_1, \ldots, \lambda_n)$ und $(\lambda_1^*, \ldots, \lambda_n^*)$ bis auf Permutation der Einträge dieselben sind. Nach allfälligem Umnummerieren der λ_i und w_i können wir daher davon ausgehen, dass

- $\lambda_1, \ldots, \lambda_k$ reell sind,
- $\lambda_{k+1}, \lambda_{k+3}, \lambda_{k+5}, \ldots$ positiven Imaginärteil haben und
- $\lambda_{k+1}^* = \lambda_{k+2}, \lambda_{k+3}^* = \lambda_{k+4}, \lambda_{k+5}^* = \lambda_{k+6}, \ldots$

Wir ersetzen nun Œ in der ONB $\underline{w}$ mehrmals jeweils endlich viele $w_{i_1}, \ldots, w_{i_r}$ ($1 \leq i_1 < \ldots < i_r \leq n$) durch eine ONB $(u_1, \ldots, u_r)$ von $\operatorname{span}(w_{i_1}, \ldots, w_{i_r})$ mit $f(u_j) = \lambda_{i_j} u_j$ für $j \in \{1, \ldots, r\}$.

Wir behaupten, dass wir so

$$w_1, \ldots, w_k \in \mathbb{R}^n \text{ und } \overline{w_{k+1}} = w_{k+2}, \overline{w_{k+3}} = w_{k+4}, \overline{w_{k+5}} = w_{k+6}, \ldots$$

erreichen können [→ 15.2.8].

Schritt 1: Œ $w_1, \ldots, w_k \in \mathbb{R}^n$

Begründung: Für jeden reellen Eigenwert λ von g seien $r \in \mathbb{N}$ und $1 \leq i_1 < \ldots < i_r \leq n$ derart, dass $\{i_1, \ldots, i_r\} = \{i \mid \lambda_i = \lambda\} \subseteq \{1, \ldots, k\}$. Dann gilt $\operatorname{span}(w_{i_1}, \ldots, w_{i_r}) = \ker(g - \lambda \operatorname{id}_{\mathbb{C}^n})$, denn „$\subseteq$“ ist trivial und $r \geq \dim \ker(g - \lambda \operatorname{id}_{\mathbb{C}^n})$ nach 10.1.15 („geometrische Vielfachheit $\leq$ algebraische Vielfachheit“). Nun gilt

$$\ker(g - \lambda \operatorname{id}_{\mathbb{C}^n}) = \ker_{\mathbb{C}^n}(A - \lambda I_n) \supseteq \ker_{\mathbb{R}^n}(A - \lambda I_n).$$

Da $A - \lambda I_n$ als reelle Matrix denselben Rang hat wie als komplexe Matrix (der Rang kann durch Überführung in Stufenform über $\mathbb{R}$ ermittelt werden [→ 9.1.14], dieselben Zeilenoperationen sind erst recht über $\mathbb{C}$ durchführbar), gilt

$$r = \dim \ker_{\mathbb{C}^n}(A - \lambda I_n) = \dim \ker_{\mathbb{R}^n}(A - \lambda I_n).$$

Wähle nun eine ONB $(u_1, \ldots, u_r)$ des $\mathbb{R}$-Vektorraums $\ker_{\mathbb{R}^n}(A - \lambda I_n)$ (mit Standardskalarprodukt). Dann ist $(u_1, \ldots, u_r)$ auch eine ONB des $\mathbb{C}$-Vektorraums

$$\ker_{\mathbb{C}^n}(A - \lambda I_n) = \operatorname{span}(w_{i_1}, \ldots, w_{i_r}).$$

Schritt 2: Œ $\overline{w_{k+1}} = w_{k+2}, \overline{w_{k+3}} = w_{k+4}, \ldots, \overline{w_{n-1}} = w_n$

Begründung: Für jede Eigenwert λ von g mit negativem Imaginärteil seien $r \in \mathbb{N}$ und $1 \leq i_1 < \ldots < i_r \leq n$ derart, dass $\{i_1, \ldots, i_r\} = \{i \mid \lambda_i = \lambda\} \subseteq \{k+2, k+4, \ldots, n\}$. Dann gilt wieder

$$\operatorname{span}(w_{i_1}, \ldots, w_{i_r}) = \ker(g - \lambda \operatorname{id}_{\mathbb{C}^n}).$$

Nun gilt

$$\begin{aligned}\ker(g - \lambda \operatorname{id}_{\mathbb{C}^n}) = \ker(A - \lambda I_n) = \{x \in \mathbb{C}^n \mid Ax = \lambda x\} = \{\overline{x} \mid x \in \mathbb{C}^n, A\overline{x} = \lambda \overline{x}\} \\ \overset{A \in \mathbb{R}^{n\times n}}{=} \{\overline{x} \mid x \in \mathbb{C}^n, Ax = \lambda^* x\} = \{\overline{x} \mid x \in \ker(g - \lambda^* \operatorname{id}_{\mathbb{C}^n})\}.\end{aligned}$$

Wegen $\{i \mid \lambda_i = \lambda^*\} = \{i_1 - 1, \ldots, i_r - 1\} \subseteq \{k+1, k+3, \ldots, n-1\}$ ist

$$\operatorname{span}(w_{i_1-1}, \ldots, w_{i_r-1}) = \ker(g - \lambda^* \operatorname{id}_{\mathbb{C}^n})$$

und mit $(u_1, \ldots, u_r) := (\overline{w_{i_1-1}}, \ldots, \overline{w_{i_r-1}})$ daher $\operatorname{span}(u_1, \ldots, u_r) = \operatorname{span}(g - \lambda \operatorname{id}_{\mathbb{C}^n}) = \operatorname{span}(w_{i_1}, \ldots, w_{i_r})$. Es ist $(u_1, \ldots, u_r)$ auch ein ONS und damit eine ONB von

$$\operatorname{span}(w_{i_1}, \ldots, w_{i_r}).$$

Schritt 3: Für $v_1, \ldots, v_n \in \mathbb{R}^n$ definiert durch

$$\begin{aligned}&w_1 = v_1, \ldots, w_k = v_k, \\ &\qquad \sqrt{2}w_{k+1} = v_{k+1} + \mathring{\imath} v_{k+2}, \sqrt{2}w_{k+3} = v_{k+3} + \mathring{\imath} v_{k+4}, \ldots, \sqrt{2}w_{n-1} = v_{n-1} + \mathring{\imath} v_n.\end{aligned}$$

ist $\underline{v} := (v_1, \ldots, v_n)$ eine ONB des $\mathbb{R}^n$ mit $M(f, \underline{v})$ von der gewünschten Gestalt.

Begründung: Mit Lemma 15.2.9 sieht man leicht, dass $\underline{v}$ eine ONB des $\mathbb{C}^n$ und damit auch des $\mathbb{R}^n$ ist. Dass $M(f, \underline{v})$ die gewünschte Gestalt hat, folgt nun leicht wie folgt: Sei $j \in \{k+1, k+3, \ldots, n-1\}$ und schreibe $\lambda_j = a + \mathring{\imath} b$ mit $a, b \in \mathbb{R}$. Dann gilt

$$\begin{aligned}f(v_j) + \mathring{\imath} f(v_{j+1}) &= g(v_j) + \mathring{\imath} g(v_{j+1}) = g(v_j + \mathring{\imath} v_{j+1}) = g(\sqrt{2}w_j) = \sqrt{2}g(w_j) = \sqrt{2}\lambda_j w_j \\ &= \lambda_j \sqrt{2} w_j = \lambda_j (v_j + \mathring{\imath} v_{j+1}) = (a + \mathring{\imath} b)(v_j + \mathring{\imath} v_{j+1}) \\ &= a v_j - b v_{j+1} + \mathring{\imath}(a v_{j+1} + b v_j).\end{aligned}$$

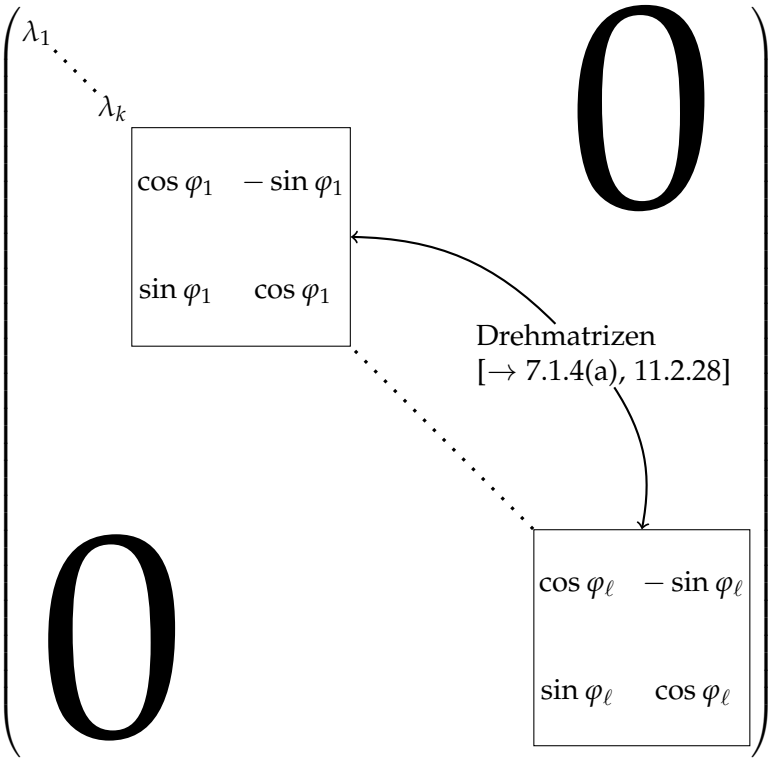
und daher $f(v_j) = a v_j - b v_{j+1}$ und $f(v_{j+1}) = b v_j + a v_{j+1}$. □

Satz 15.2.11. [10][→ 11.3.9] *Sei V ein endlichdimensionaler $\mathbb{R}$-Vektorraum mit Skalarprodukt und $f\colon V \to V$ linear. Dann sind folgende Aussagen äquivalent:*

(a) *f ist orthogonal* [→ 11.2.19].

[10]Im Beweis dieses Satzes benutzen wir den Fundamentalsatz der Algebra 4.2.12 [→ 4.2.13(g)].

(b) *Es gibt eine ONB* $\underline{v}$ *von* V *derart, dass* $M(f, \underline{v})$ *von der Gestalt*



mit $\lambda_1, \ldots, \lambda_k \in \{-1, 1\}$ *und* $\varphi_1, \ldots, \varphi_\ell \in \mathbb{R}$ *ist.*

Beweis. (b) $\Longrightarrow$ (a) ist einfach.

(a) $\Longrightarrow$ (b) Es reicht zu zeigen, dass es für $a, b \in \mathbb{R}$ mit

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}^* \begin{pmatrix} a & b \\ -b & a \end{pmatrix} = I_2$$

ein $\varphi \in \mathbb{R}$ gibt mit

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

Seien also $a, b \in \mathbb{R}$ mit

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} a & b \\ -b & a \end{pmatrix} = I_2,$$

das heißt $a^2 + (-b)^2 = a^2 + b^2 = 1$. Dann gibt es $\varphi \in \mathbb{R}$ mit

$$\begin{pmatrix} a \\ -b \end{pmatrix} = \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix}$$

und somit

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

□

Definition 15.2.12. [→ 11.2.25, 11.3.3] Eine Matrix $A \in \mathbb{K}^{n\times n}$ heißt *normal*, wenn f_A normal ist.

Proposition 15.2.13. [→ 11.2.27, 11.3.7] *Sei $A \in \mathbb{K}^{n\times n}$. Dann gilt*

$$A \text{ normal} \iff A^*A = AA^*.$$

Beweis.

$$\begin{aligned}
A \text{ normal} &\iff f_A \text{ normal} \\
&\iff f_A^* \circ f_A = f_A \circ f_A^* \\
&\iff M(f_A^* \circ f_A, \underline{e}) = M(f_A \circ f_A^*, \underline{e}) \\
&\iff M(f_A^*, \underline{e})M(f_A, \underline{e}) = M(f_A, \underline{e})M(f_A^*, \underline{e}) \\
&\overset{15.1.9}{\iff} M(f_A, \underline{e})^*M(f_A, \underline{e}) = M(f_A, \underline{e})M(f_A, \underline{e})^* \\
&\iff A^*A = AA^*
\end{aligned}$$

□

Proposition 15.2.14. [→ 11.2.26, 11.3.4] *Seien V ein Vektorraum mit Skalarprodukt und $\underline{v} = (v_1, \ldots, v_n)$ eine ONB. Sei $f\colon V \to V$ linear. Dann gilt*

$$f \text{ ist normal} \iff M(f, \underline{v}) \text{ ist normal.}$$

Beweis.

$$\begin{aligned}
f \text{ normal} &\iff f^* \circ f = f \circ f^* \\
&\iff M(f^* \circ f, \underline{v}) = M(f \circ f^*, \underline{v}) \\
&\iff M(f^*, \underline{v})M(f, \underline{v}) = M(f, \underline{v})M(f^*, \underline{v}) \\
&\overset{15.1.9}{\iff} M(f, \underline{v})^*M(f, \underline{v}) = M(f, \underline{v})M(f, \underline{v})^* \overset{15.2.13}{\iff} M(f, \underline{v}) \text{ normal}
\end{aligned}$$

□

Ähnlich wie man zum Beispiel 11.3.9 in 11.3.10 übersetzen konnte, kann man auch die Sätze in diesem Abschnitt matrizentheoretisch fomulieren. Wir überlassen dies dem Leser.

16 Teilbarkeit in kommutativen Ringen

In diesem Kapitel sei stets A ein kommutativer Ring [→ 3.1.1].

16.1 Teilerbeziehung und Ideale

Definition 16.1.1. Wir führen auf A die Relationen „$|$“ und „$\mathrel{\widehat{=}}$“ ein durch

$$a \mid b :\Longleftrightarrow \exists c \in A : ac = b \qquad (a, b \in A)$$

„a teilt b (in A)“

„a ist Teiler von b (in A)“

„b ist Vielfaches von a (in A)“

$$a \mathrel{\widehat{=}} b :\Longleftrightarrow (a \mid b \;\&\; b \mid a) \qquad (a, b \in A)$$

„a ist assoziiert zu b (in A)“

„a und b sind assoziiert (zueinander) (in A)“

Erinnerung 16.1.2. [→ 3.3.4] Eine Untergruppe I der additiven Gruppe von A heißt Ideal von A, wenn $\forall a \in A : \forall b \in I : ab \in I$ [→ 3.3.4]. Für $a \in A$ ist $(a) := \{ca \mid c \in A\}$ das kleinste Ideal I von A mit $a \in I$, genannt das von a erzeugte (Haupt-)Ideal. Man nennt ein Ideal I von A ein Hauptideal, wenn es $a \in A$ gibt mit $I = (a)$ [→ 3.3.11].

Bemerkung 16.1.3. Seien $a, b \in A$. Dann gilt

$$\begin{aligned} a \mid b &\iff b \in (a) \iff (b) \subseteq (a) \qquad \text{und} \\ a \mathrel{\widehat{=}} b &\iff (a) = (b). \end{aligned}$$

Insbesondere ist $\mathrel{\widehat{=}}$ eine Äquivalenzrelation [→ 1.3.1(b)] auf A und auf der Quotientenmenge [→ 1.3.5(a)] $A/\mathrel{\widehat{=}}$ können wir eine Halbordnung [→ 12.1.1] $\preceq$ definieren durch

$$\hat{a} \preceq \hat{b} :\Longleftrightarrow a \mid b \qquad (a, b \in A),$$

wobei für jedes $a \in A$ mit $\widehat{a} := \widehat{\widehat{=}}\, a$ die Äquivalenzklasse [→ 1.3.1(b), 2.3.1] von a bezeichnet wird. In der Tat: Sind $a, a', b, b' \in A$ mit $\widehat{a} = \widehat{a'}$ und $\widehat{b} = \widehat{b'}$, so gilt $(a) = (a')$ und $(b) = (b')$ und daher

$$a \mid b \iff (b) \subseteq (a) \iff (b') \subseteq (a') \iff a' \mid b'.$$

Definition 16.1.4. Sei $B \subseteq A$ und $a \in A$. Es heißt a ein $\left\{\begin{array}{c} \textit{gemeinsamer Teiler (gT)} \\ \textit{gemeinsames Vielfaches (gV)} \end{array}\right\}$ der Elemente von B (in A), wenn $\widehat{a}$ eine $\left\{\begin{array}{c} \text{untere} \\ \text{obere} \end{array}\right\}$ Schranke [→ 12.1.6] von $\{\widehat{b} \mid b \in B\}$ in $(A/\mathrel{\widehat{=}}, \preceq)$ ist. Es heißt a ein $\left\{\begin{array}{c} \text{größter gT (ggT)} \\ \text{kleinstes gV (kgV)} \end{array}\right\}$ der Elemente von B (in A), wenn $\widehat{a}$ die $\left\{\begin{array}{c} \text{größte untere} \\ \text{kleinste obere} \end{array}\right\}$ Schranke [→ 12.1.6] von $\{\widehat{b} \mid b \in B\}$ in $(A/\mathrel{\widehat{=}}, \preceq)$ ist.

Definition 16.1.5. Seien $(A_1, \preceq_1)$ und $(A_2, \preceq_2)$ halbgeordnete Mengen. Eine Abbildung $f\colon A_1 \to A_2$ heißt *Isomorphismus halbgeordneter Mengen*, wenn f bijektiv ist und

$$\forall a, b \in A_1 : (a \preceq_1 b \iff f(a) \preceq_2 f(b)).$$

Bemerkung 16.1.6. Sei $B \subseteq A$ und $a \in A$.

(a) Es ist klar, wie man die Definition von gT, gV, ggT, kgV ohne die Verwendung von $\widehat{=}$ lesen kann:

$$a \text{ gT der Elemente von } B \iff \forall b \in B : a \mid b$$

$$a \text{ ggT der Elemente von } B \iff \begin{pmatrix} a \text{ gT der Elemente von } B & \& \\ \forall b \in A : (b \text{ gT der Elemente von } B \implies b \mid a) \end{pmatrix}$$

(b) Wegen der Eindeutigkeit von Infima und Suprema in halbgeordneten Mengen [→ 12.1.6] sind ggT und kgV in kommutativen Ringen genau bis auf Assoziiertheit eindeutig, sofern sie existieren.

(c) Betrachte die durch die Obermengenbeziehung (umgekehrte Inklusion) halbgeordnete Menge aller Hauptideale $H := \{(a) \mid a \in A\}$ von A. Nach Definition von $\widehat{=}$ ist die Abbildung $A/\mathrel{\widehat{=}} \to H$, $\widehat{a} \mapsto (a)$ wohldefiniert und injektiv. Offenbar ist sie auch surjektiv. Nach Definition von $\preceq$ ist sie ein Isomorphismus halbgeordneter Mengen. In der Definition von gT/gV/ggT/kgV [→ 16.1.4] könnte man daher genausogut (a) statt $\widehat{a}$, (b) statt $\widehat{b}$ und $(H, \supseteq)$ statt $(A/\mathrel{\widehat{=}}, \preceq)$ schreiben.

Erinnerung 16.1.7. (a) Die Schnittmenge einer Menge von Idealen in A ist wieder ein Ideal von A. Den leeren Schnitt definieren wir dabei als A. [→ 3.3.8]

(b) Für jede Teilmenge E von A gibt es eine kleinstes Ideal, welches E enthält [→ 3.3.9]. Man nennt es das von E erzeugte Ideal und notiert es mit $(E) = (E)_A$. Es besteht gerade aus allen Summen von Vielfachen von Elementen von E [→ 3.3.10].

Proposition 16.1.8. *Sei $B \subseteq A$ und $a \in A$.*

(a) *a gT der Elemente von B $\iff$ $B \subseteq (a)$ $\iff$ $(B) \subseteq (a)$*

(b) *a gV der Elemente von B $\iff$ $a \in \bigcap\{(b) \mid b \in B\}$ $\iff$ $(a) \subseteq \bigcap\{(b) \mid b \in B\}$*

(c) *a ggT der Elemente von B $\impliedby$ $(a) = (B)$*
und wenn (B) ein Hauptideal ist, gilt auch „$\implies$“

(d) *a kgV der Elemente von B $\iff$ $(a) = \bigcap\{(b) \mid b \in B\}$*

Beweis. (a) und (b) sind klar.

Zu (c). „$\impliedby$“ Gelte $(a) = (B)$. Nach (a) ist a ein gT der Elemente von B. Sei b ein gT der Elemente von B. Zu zeigen: $b \mid a$. Dann $a \in (a) = (B) \overset{\text{(a)}}{\subseteq} (b)$ und daher $b \mid a$.

„$\Longrightarrow$“ Sei a ein ggT der Elemente von B und $b \in A$ mit $(b) = (B)$. Nach (a) gilt $(B) \subseteq (a)$. Noch zu zeigen: $(a) \subseteq (B)$. Da b ein gT und a ein ggT der Elemente von B ist, gilt $b \mid a$, also $(a) \subseteq (b) = (B)$.

Zu (d). Wegen (b) genügt es zu zeigen:

$$(\forall c \in A : (c \text{ gV der Elemente von } B \implies a \mid c)) \iff (a) \supseteq \bigcap\{(b) \mid b \in B\}.$$

Dies ist klar. □

Bemerkung 16.1.9. In 3.3.13 haben wir gezeigt, dass in $\mathbb{Z}$ jedes Ideal ein Hauptideal ist. In 10.2.2 haben wir dasselbe für den Polynomring $K[X]$ über einem Körper K gezeigt. Nach 16.1.8(c) läuft also in diesen Ringen das bestimmen eines ggT auf die Berechnung eines Erzeugers eines Ideals hinaus. Im Polynomring $\mathbb{Z}[X]$ zum Beispiel ist dies jedoch nicht so, wie Beispiel 16.1.11 unten zeigt.

Sprechweise 16.1.10. Seien $b_1, \ldots, b_n \in A$. Wenn wir von einem gT/gV/ggT/kgV von $b_1, \ldots, b_n$ schreiben, so meinen wir einen gT/gV/ggT/kgV von $\{b_1, \ldots, b_n\}$. Weiter nennen wir $(b_1, \ldots, b_n) := (\{b_1, \ldots, b_n\}) = \{\sum_{i=1}^{n} a_i b_i \mid a_1, \ldots, a_n \in A\}$ [→ 3.3.11] das von $b_1, \ldots, b_n$ erzeugte Ideal.

Beispiel 16.1.11. Die Teiler von 2 in $\mathbb{Z}[X]$ sind $-2, -1, 1, 2$, Die Teiler von X in $\mathbb{Z}[X]$ sind $-X, -1, 1, X$. Die gT von 2 und X sind daher -1 und 1. Die ggT von 2 und X sind daher ebenfalls -1 und 1. Allerdings gilt $(2, X) = \{2p + Xq \mid p, q \in \mathbb{Z}[X]\} = \{2a + Xp \mid a \in \mathbb{Z}, p \in \mathbb{Z}[X]\} \neq (1) = \mathbb{Z}[X]$.

Beispiel 16.1.12. Sei A ein kommutativer Ring.

(a) $\forall a \in A : (1 \mid a \;\&\; a \mid 0)$

(b) $\forall a \in A : (a \mid 1 \iff a \in A^\times)$, wobei $A^\times = \{a \in A \mid \exists b \in A : ab = 1\}$ die Einheitengruppe von A ist [→ 4.1.1].

(c) $\forall a \in A : (0 \mid a \iff a = 0)$

(d) $\widehat{1} = A^\times, \widehat{0} = \{0\}$

(e) Wegen $(\emptyset) = (0)$ ist 0 ein ggT der Elemente von $\emptyset$. Wegen $\widehat{0} = \{0\}$ ist es der einzige.

(f) Wegen $(A) = A = (1)$ ist 1 ein ggT der Elemente von A. Wegen $\widehat{1} = A^\times$ sind diese ggT genau die Einheiten.

16.2 Integritäts- und Hauptidealringe

Definition 16.2.1. Ein kommutativer Ring A heißt *Integritätsring* (auch *Integritätsring*), wenn $1 \neq 0$ in A und $\forall a, b \in A : (ab = 0 \implies (a = 0 \text{ oder } b = 0))$. Ein Integritätsring A heißt *Hauptidealring* (auch *Hauptidealbereich*), wenn jedes Ideal von A ein Hauptideal ist.

Beispiel 16.2.2. (a) Jeder Unterring [→ 3.2.1, 3.2.2] eines Körpers [→ 4.1.4] ist ein Integritätsring.

(b) $\mathbb{Z}$ ist ein Hauptidealring [→ 3.3.13].

(c) Für jeden Körper K ist $K[X]$ ein Hauptidealring [→ 10.2.2].

Proposition 16.2.3. (a) *Sei A ein Integritätsring. Dann gilt die* Kürzungsregel*:*

$$\forall a, b, c \in A : ((ac = bc \;\&\; c \neq 0) \implies a = b)$$

(b) *Sei A ein Hauptidealring und $B \subseteq A$. Dann gibt es einen ggT und ein kgV der Elemente von B.*

Beweis. Zu (a). Seien $a, b, c \in A$ mit $ac = bc$ und $c \neq 0$. Dann gilt $(a-b)c = ac - bc = 0$ und daher $a - b = 0$ oder $c = 0$. Wegen $c \neq 0$ folgt $a = b$.

(b) folgt sofort aus 16.1.8(c)(d), da die Ideale (B) und $\bigcap\{(b) \mid b \in B\}$ Hauptideale sind. □

Proposition 16.2.4. *Sei A ein Integritätsring und $a, b \in A$. Dann gilt*

$$a \mathrel{\widehat{=}} b \iff \exists c \in A^\times : a = bc.$$

Beweis. „$\Longrightarrow$“ Gelte $a \mathrel{\widehat{=}} b$, also $(a) = (b)$. Dann gibt es $c, d \in A$ mit $a = bc$ und $b = ad$. Es folgt $b = bcd$. Ist $b = 0$, so $(a) = (0)$ und daher $a = 0 = 0 \cdot 1 = b \cdot 1$ (und $1 \in A^\times$). Sei also $b \neq 0$. Dann $1 = cd$ wegen 16.2.3(a) und daher $c \in A^\times$.

„$\Longleftarrow$“ Sei $c \in A^\times$ mit $a = bc$. Dann $b \mid a$. Wegen $c^{-1}a = b$ aber auch $a \mid b$. Daher $a \mathrel{\widehat{=}} b$. □

Beispiel 16.2.5. (a) $\mathbb{N}_0 \to \mathbb{Z}/\mathrel{\widehat{=}},\ n \mapsto \widehat{n}$ ist eine Bijektion, denn $\mathbb{Z}^\times = \{-1, 1\}$.

(b) Sei K ein Körper. Dann ist $\{p \in K[X] \mid p \text{ normiert oder null}\} \to K[X]/\mathrel{\widehat{=}}$ eine Bijektion, denn $K[X]^\times = K^\times$ [→ 10.2.3].

Bemerkung 16.2.6. Aus 16.2.2(b)(c) folgt nun:

(a) Zu jeder Menge $B \subseteq \mathbb{Z}$ gibt es genau $\left\{\begin{matrix}\text{einen ggT}\\ \text{ein kgV}\end{matrix}\right\}$ $a \in \mathbb{N}_0$ der Elemente von B in $\mathbb{Z}$, oft notiert mit $\left\{\begin{matrix}\gcd(B)\\ \operatorname{lcm}(B)\end{matrix}\right\}$.

(b) Sei K ein Körper. Zu jeder Menge $B \subseteq K[X]$ gibt es genau $\left\{\begin{matrix}\text{einen ggT}\\ \text{ein kgV}\end{matrix}\right\}$ $p \in K[X]$ der Elemente von B in $K[X]$ mit $p = 0$ oder p normiert, oft notiert mit $\left\{\begin{matrix}\gcd(B)\\ \operatorname{lcm}(B)\end{matrix}\right\}$.

greatest common divisor

least common multiple

Beispiel 16.2.7. (a) $\operatorname{lcm}(\emptyset) = 1$ und $\gcd(\emptyset) = 0$ sowohl in $\mathbb{Z}$ als auch in $K[X]$.

(b) $\operatorname{lcm}(\{-5\}) = 5$ in $\mathbb{Z}$ und $\operatorname{lcm}(\{-5\}) = 1$ in $\mathbb{R}[X]$.

(c) $\operatorname{lcm}(\{5,3,2,6\} = 30$ in $\mathbb{Z}$, denn $\underbrace{\underbrace{(5)\cap(3)}_{(15)}\cap(2)\cap(6)}_{(30)} = (30)$.

(d) $\operatorname{lcm}(\{2X^2-2, X^2-2X+1\}) = (X-1)^2(X+1)$, denn $(2X^2-2)\cap(X^2-2X+1) = (X^2-1)\cap((X-1)^2) = ((X-1)(X+1))\cap((X-1)^2) \overset{4.2.10}{=} ((X+1)(X-1)^2)$.

(e) $\gcd(\{153, 204, -357, 0\}) = 51$ in $\mathbb{Z}$, denn $(153, 204, -357, 0) = (153, 204, 357) = (153, 204, 357-153) = (153, 204, 204) = (153, 204) = (153, 204-153) = (153, 51) = (102, 51) = (51)$.

(f) $\gcd(\{2X^2-2, X^2-2X+1\}) = X-1$, denn $(2X^2-2, X^2-2X+1) = (X^2-1, -2X+2) = (X^2-1, X-1) = (X-1)$ in $\mathbb{Q}[X]$.

(g) $\operatorname{lcm}(\{2^m \mid m \in \mathbb{N}\}) = 0$ in $\mathbb{Z}$.

16.3 Zur Berechnung größter gemeinsamer Teiler

Gegeben seien $b_1, \ldots, b_n \in A$. Um die gT von $b_1, \ldots, b_n$ zu bestimmen oder sogar einen ggT von $b_1, \ldots, b_n$ zu berechnen, falls er existiert, ist es generell eine gute Strategie „einfachere“ $c_1, \ldots, c_m \in A$ zu suchen mit $(b_1, \ldots, b_n) = (c_1, \ldots, c_m)$. Gilt nämlich diese Gleichung, so haben $b_1, \ldots, b_n$ und $c_1, \ldots, c_m$ jeweils dieselben gT nach 16.1.8(a). Existiert sogar ein $c \in A$ mit $(b_1, \ldots, b_n) = (c)$, was in Hauptidealringen immer der Fall ist, so ist c nach 16.1.8(c) ein ggT von $b_1, \ldots, b_n$. Wir werden sehen, wie man in $\mathbb{Z}$ und in $K[X]$ (modulo dem Problem, die Rechenoperationen im Körper K durchzuführen) ein solches c und damit $\gcd\{b_1, \ldots, b_n\}$ immer berechnen kann. Weiter werden wir sehen, wie man in $\mathbb{Z}$ und in $K[X]$, allerdings mit erheblich mehr Aufwand, sogar $a_1, \ldots, a_n$ berechnen kann mit $a_1b_1 + \ldots + a_nb_n = \gcd\{b_1, \ldots, b_n\}$.

Satz 16.3.1. *Seien $b_1, \ldots, b_n \in A$, $i \in \{1, \ldots, n\}$ und sei $d \in A$ derart, dass $\overline{b_i} \mathrel{\widehat{=}} \overline{d}$ in $A/(b_1, \ldots, b_{i-1}, b_{i+1}, \ldots, b_n)$. Dann gilt*

$$(b_1, \ldots, b_n) = (b_1, \ldots, b_{i-1}, d, b_{i+1}, \ldots, b_n).$$

Beweis. Da man in der Behauptung b_i und d vertauschen kann, reicht es „$\subseteq$“ zu zeigen. Dafür reicht es, $b_i \in (b_1, \ldots, b_{i-1}, d, b_{i+1}, \ldots, b_n)$ zu zeigen. Wegen $(\overline{b_i}) = (\overline{d})$ gibt es $e \in A$ mit $\overline{b_i} = \overline{e}\overline{d}$ in $A/(b_1, \ldots, b_{i-1}, b_{i+1}, \ldots, b_n)$. Daher $b_i - ed \in (b_1, \ldots, b_{i-1}, b_{i+1}, \ldots, b_n)$ und somit $b_i \in (b_1, \ldots, b_{i-1}, d, b_{i+1}, \ldots, b_n)$. □

Beispiel 16.3.2. (a) In $\mathbb{Z}$ gilt

$$\begin{aligned}(38\,321\,783\,292, 27, 45) &= (38\,321\,783\,292, 27, 18) = (38\,321\,783\,292, 9, 18)\\ &= (38\,321\,783\,292, 9)\\ &= (3+8+3+2+1+7+8+3+2+9+2, 9)\\ &= (3, 9) = (3)\end{aligned}$$

und daher $\gcd\{38\,321\,783\,292, 27, 45\} = 3$.

(b) In $\mathbb{Z}$ gilt $(43733, 17473, 27977) = (43733, 17473, 10504) = (1717, 17473, 10504) = (1717, 303, 10504) = (1717, 303, 1414) = (303, 303, 1414) = (303, 1414) = (303, 202) = (101)$ und daher $\gcd\{43733, 17473, 27977\} = 101$.

(c) In $\mathbb{Z}$ gilt $(\underbrace{348}_{10\,000}\ \underbrace{390}_{15}\ \underbrace{458}_{04}\ \underbrace{502}_{02}, 24, 54) = (\underbrace{150}_{03}\ \underbrace{404\,022}_{040404}, 24, 6) = (\underbrace{30}_{00}\ \underbrace{040}_{04}\ \underbrace{404}_{04}, 24, 6) = (\underbrace{4044}_{0408}, 24, 6) = (\underbrace{408}_{04}, 24, 6) = (24, 6) = (6)$ und daher

$$\gcd\{348\,390\,458\,502, 24, 54\} = 6.$$

(d) In $\mathbb{Z}$ gilt $(3^{30} - 1, 3^{45} - 1) = (3^{30} - 1, 3^{15} - 1) = (1 - 1, 3^{15} - 1)$ und daher

$$\gcd\{3^{30} - 1, 3^{45} - 1\} = 3^{15} - 1.$$

(e) In $\mathbb{Q}[X]$ gilt

$$\begin{aligned}(-X^7 + X^5 - X^3 + X, X^8 - 1) &\overset{\overline{X} \in (\mathbb{Q}[X]/(X^8-1))^\times}{=} (-X^6 + X^4 - X^2 + 1, X^8 - 1)\\ &= (-X^6 + X^4 - X^2 + 1, X^6 - X^4 + X^2 - 1) = (-X^6 + X^4 - X^2 + 1)\end{aligned}$$

und daher $\gcd\{-X^7 + X^5 - X^3 + X, X^8 - 1\} = X^6 - X^4 + X^2 - 1$.

(f) In $\mathbb{Q}[X]$ gilt

$$\begin{aligned}&(X^4 - 2X^3 + 2X^2 - 2X + 1, \underbrace{4X^3 - 6X^2 + 4X - 2}_{=:p}, \underbrace{X^3 + 3X^2 - 9X + 5}_{=:q}) =\\ &(-5X^3 + 11X^2 - 7X + 1, p, q) = (8X^2 - 12X + 4, -18X^2 + 40X - 22, q) =\\ &(2X^2 - 3X + 1, 13X - 13, q) = (2 - 3 + 1, X - 1, 1 + 3 - 9 + 5) =\\ &(0, X - 1, 0) = (X - 1)\end{aligned}$$

und daher $\gcd\{X^4 - 2X^3 + 2X^2 - 2X + 1, p, q\} = X - 1$.

Es sollte nun klar sein, dass man in $\mathbb{Z}$ und in $K[X]$ (sofern K ein Körper ist, in dem man zu rechnen weiß) durch mehrfache Anwendung von Satz 16.3.1 zu gegebenen Elementen $b_1, \ldots, b_n$ stets $\gcd\{b_1, \ldots, b_n\}$ berechnen kann. Solange man nämlich das Ideal $(b_1, \ldots, b_n)$ noch nicht als Hauptideal geschrieben hat (also mindestens zwei Erzeuger $\neq 0$ hat), kann man die Summe der Absolutbeträge (für $\mathbb{Z}$) bzw. die Summe der Grade (für $K[X]$) der Erzeuger $\neq 0$ in jedem Schritt verringern.

Will man zu gegebenen $b_1, \ldots, b_n \in A$ nicht nur, falls es existiert, ein $b \in A$ finden mit $(b_1, \ldots, b_n) = (b)$, sondern auch noch $a_1, \ldots, a_n \in A$ mit $b = a_1 b_1 + \ldots + a_n b_n$, so kann man das wie folgt versuchen: Man bildet die Matrix

$$\begin{pmatrix} b_1 & & \\ \vdots & & I_n \\ b_n & & \end{pmatrix} = \begin{pmatrix} b_1 & 1 & & 0 \\ \vdots & & \ddots & \\ b_n & 0 & & 1 \end{pmatrix} \in A^{n \times (n+1)}.$$

Die i-te Zeile dieser Matrix kann man als die (trivialerweise) gültige Gleichung

$$b_i = 0 \cdot b_1 + \ldots + 0 \cdot b_{i-1} + 1 \cdot b_i + 0 \cdot b_{i+1} + \ldots + 0 \cdot b_n$$

interpretieren. Nun überträgt man die vom Gauß-Verfahren aus §5.2 bekannten elementaren Zeilenoperationen von Matrizen über Körpern auf Matrizen über kommutativen Ringen:

- $Z_i \leftarrow Z_i + \lambda Z_j \qquad (i \neq j, \lambda \in A)$
  (Addieren des λ-fachen einer Zeile zu einer anderen)
- $Z_i \leftarrow \lambda Z_i \qquad (\lambda \in A^\times)$
  (Multiplizieren einer Zeile mit einer Einheit).

Nach Satz 16.3.1 ändert sich das von den Einträgen der ersten Spalte aufgespannte Ideal hierbei nicht. Auch kann man die dabei entstehenden Zeilen jeweils als eine gültige Gleichung interpretieren.

Sowohl für $A = \mathbb{Z}$ als auch für $A = K[X]$ (K Körper) kann man durch endlich viele dieser Zeilenoperationen die Matrix $\begin{pmatrix} b_1 & & \\ \vdots & & I_n \\ b_n & & \end{pmatrix}$ überführen in eine Matrix $\begin{pmatrix} \gcd\{b_1, \ldots, b_n\} & a_1 & \ldots & a_n \\ 0 & & & \\ \vdots & & * & \\ 0 & & & \end{pmatrix}$, deren erste Zeile als Gleichung gelesen besagt, dass

$$\gcd\{b_1, \ldots, b_n\} = a_1 b_1 + \ldots + a_n b_n.$$

Beispiel 16.3.3. Durch Anwendung elementarer Zeilenoperationen über $\mathbb{Z}$ erhält man

folgende Sequenz von Matrizen:

$$\begin{pmatrix} 43733 & 1 & 0 & 0 \\ 17473 & 0 & 1 & 0 \\ 27977 & 0 & 0 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 43733 & 1 & 0 & 0 \\ 17473 & 0 & 1 & 0 \\ 10504 & 0 & -1 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1717 & 1 & 4 & -4 \\ 17473 & 0 & 1 & 0 \\ 10504 & 0 & -1 & 1 \end{pmatrix} \rightsquigarrow$$
$$\begin{pmatrix} 1717 & 1 & 4 & -4 \\ 303 & -10 & -39 & 40 \\ 10504 & 0 & -1 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1717 & 1 & 4 & -4 \\ 303 & -10 & -39 & 40 \\ 1414 & 300 & 1169 & -1199 \end{pmatrix} \rightsquigarrow$$
$$\begin{pmatrix} 303 & * & * & * \\ 303 & -10 & -39 & 40 \\ 1414 & 300 & 1169 & -1199 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 303 & * & * & * \\ 303 & -10 & -39 & 40 \\ 202 & 340 & 1325 & -1359 \end{pmatrix} \rightsquigarrow$$
$$\begin{pmatrix} 303 & * & * & * \\ 101 & -350 & -1364 & 1399 \\ 202 & * & * & * \end{pmatrix} \rightsquigarrow \begin{pmatrix} 0 & * & * & * \\ 101 & -350 & -1364 & 1399 \\ 0 & * & * & * \end{pmatrix}.$$

Also $\gcd\{43733, 17473, 27977\} = 101 = (-350) \cdot 43733 + (-1364) \cdot 17473 + 1399 \cdot 27977$.

16.4 Faktorielle Ringe

Definition 16.4.1. Sei $p \in A$. Es heißt p *irreduzibel* (in A), wenn

$$p \notin A^\times \ \ \& \ \ \forall a, b \in A : (p = ab \implies (a \in A^\times \text{ oder } b \in A^\times)),$$

und *prim* (in A) (oder *Primelement* von A), wenn

$$p \notin A^\times \ \ \& \ \ \forall a, b \in A : (p \mid ab \implies (p \mid a \text{ oder } p \mid b)).$$

Bemerkung 16.4.2. (a) 0 ist niemals irreduzibel, denn sonst erhielte man aus $0 = 0 \cdot 0$, dass $0 \in A^\times$ im Widerspruch zur angenommenen Irreduzibilität.

(b) 0 ist prim in $A \iff A$ Integritätsring [$\rightarrow$ 16.2.1]

Proposition 16.4.3. *Sei A ein Integritätsring. Dann ist jedes Primelement $\neq 0$ von A irreduzibel in A.*

Beweis. Sei $p \neq 0$ prim in A. Zu zeigen:

(a) $p \notin A^\times$

(b) $\forall a, b \in A : (p = ab \implies (a \in A^\times \text{ oder } b \in A^\times))$

(a) ist Teil der Definition eines Primelements.

Zu (b). Seien $a, b \in A$ mit $p = ab$. Wegen $p \mid ab$ gilt dann $p \mid a$ oder $p \mid b$, Œ $p \mid a$, etwa $a = a'p$ mit $a' \in A$. Es folgt $p = ab = a'pb$ und daher $1 = a'b$ wegen $p \neq 0$. Somit $b \in A^\times$. □

Die Voraussetzung, dass A Integritätsring ist, ist nicht überflüssig:

Beispiel 16.4.4. $\overline{2} = \overline{2} \cdot \overline{4}$ in $\mathbb{Z}/(6)$ und $\overline{2}, \overline{4} \notin (\mathbb{Z}/(6))^\times$. Daher ist $\overline{2}$ nicht irreduzibel in $\mathbb{Z}/(6)$. Es ist aber $\overline{2}$ ein Primelement $\neq 0$ in $\mathbb{Z}/(6)$, denn $\overline{2} \notin (\mathbb{Z}/(6))^\times$ und sind $a, b \in \mathbb{Z}$ mit $\overline{2} \mid \overline{a}\overline{b}$ in $\mathbb{Z}/(6)$, so ist a oder b eine gerade ganze Zahl und daher $\overline{2} \mid \overline{a}$ oder $\overline{2} \mid \overline{b}$ in $\mathbb{Z}/(6)$.

Erinnerung 16.4.5. [$\to$ 4.1.6] Wir nennen $n \in \mathbb{N}$ mit $n \geq 2$ eine Primzahl, wenn es nicht $s, t \in \mathbb{N}$ mit $s, t \geq 2$ und $n = st$ gibt. Wir schreiben $\mathbb{P} := \{2, 3, 5, 7, 11, \ldots\}$ für die Menge der Primzahlen.

Proposition 16.4.6. *Für $p \in \mathbb{Z}$ sind äquivalent:*

(a) $p \in \mathbb{P}$

(b) *p ist irreduzibel in $\mathbb{Z}$ und $p \geq 0$*

(c) *p ist prim in $\mathbb{Z}$ und $p > 0$*

Beweis. (a) $\iff$ (b) ist sehr einfach.
(c) $\implies$ (b) folgt aus Proposition 16.4.3.
(a) $\implies$ (c) Gelte $p \in \mathbb{P}$. Dann $p \notin \{-1, 1\} = \mathbb{Z}^\times$. Noch zu zeigen:

$$\forall a, b \in \mathbb{Z} : (p \mid ab \implies (p \mid a \text{ oder } p \mid b)).$$

Mit Hilfe von $\mathbb{Z}/(p)$ kann man das anders schreiben als:

$$\forall a, b \in \mathbb{Z}/(p) : (ab = 0 \implies (a = 0 \text{ oder } b = 0)).$$

Wegen $p \in \mathbb{P}$ ist aber nach 4.1.7 der kommutative Ring $\mathbb{Z}/(p)$ ein Körper, woraus dies sofort folgt. □

Beispiel 16.4.7. Im Unterring $\mathbb{Z}[2\mathring{i}] = \{a + b2\mathring{i} \mid a, b \in \mathbb{Z}\}$ von $\mathbb{C}$ [$\to$ 3.2.4, 4.2.6] ist 2 irreduzibel, aber nicht prim. Wegen $(2\mathring{i})(2\mathring{i}) = -4 = (-2) \cdot 2$ gilt nämlich $2 \mid (2\mathring{i})(2\mathring{i})$ in $\mathbb{Z}[2\mathring{i}]$, aber offensichtlich gilt nicht $2 \mid 2\mathring{i}$ in $\mathbb{Z}[2\mathring{i}]$. Daher ist 2 nicht prim in $\mathbb{Z}[2\mathring{i}]$. Offensichtlich ist 2 keine Einheit in $\mathbb{Z}[2\mathring{i}]$. um zu zeigen, dass 2 irreduzibel in $\mathbb{Z}[2\mathring{i}]$ ist, reicht es schließlich $a, b, c, d \in \mathbb{Z}$ zu betrachten mit $2 = (a + 2b\mathring{i})(c + 2d\mathring{i})$. Zu zeigen ist $a + 2b\mathring{i} \in \mathbb{Z}[2\mathring{i}]^\times$ oder $c + 2d\mathring{i} \in \mathbb{Z}[2\mathring{i}]^\times$. Es gilt

$$4 = 2 \cdot 2^* = (a + 2b\mathring{i})(c + 2d\mathring{i})(a - 2b\mathring{i})(c - 2d\mathring{i}) = (a^2 + 4b^2)(c^2 + 4d^2).$$

Da $a^2 + 4b^2 \neq 2$ und $c^2 + 4d^2 \neq 2$ folgt $a^2 + 4b^2 = 1$ oder $c^2 + 4d^2 = 1$, also

$$(a \in \{-1, 1\} \;\&\; b = 0) \text{ oder } (c \in \{-1, 1\} \;\&\; d = 0).$$

Notation 16.4.8. Im Folgenden fixieren wir eine Menge $\mathbb{P}_A$ von Primelementen $\neq 0$ in A derart, dass jedes Primelement $\neq 0$ von A zu genau einem Element von $\mathbb{P}_A$ assoziiert ist. Es enthält $\mathbb{P}_A$ also aus jedem $\widehat{p}$ mit $p \in A \setminus \{0\}$ prim genau einen Vertreter.

Beispiel 16.4.9. (a) In der Regel wird man $\mathbb{P}_{\mathbb{Z}} = \mathbb{P}$ nehmen. Man könnte aber auch $\mathbb{P}_{\mathbb{Z}} = \{2, -3, 5, 7, -11, 13, \ldots\}$ nehmen.

(b) Ist K ein Körper, so nimmt man in der Regel

$$\mathbb{P}_{K[X]} = \{p \in K[X] \mid p \neq 0, p \text{ normiert}, p \text{ prim in } K[X]\}.$$

Notation 16.4.10. Es bezeichne $\mathbb{N}_0^{(\mathbb{P}_A)}$ die Menge der Funktionen $\alpha\colon \mathbb{P}_A \to \mathbb{N}_0$ mit *endlichem* Träger $\operatorname{supp}(\alpha) := \{p \in \mathbb{P}_A \mid \alpha(p) \neq 0\}$. Für jedes $\alpha \in \mathbb{N}_0^{(\mathbb{P}_A)}$ setzen wir

$$\mathbb{P}_A^{\alpha} := \prod_{p \in \operatorname{supp}(\alpha)} p^{\alpha(p)}.$$

Wir nennen $(c, \alpha) \in A^{\times} \times \mathbb{N}_0^{(\mathbb{P}_A)}$ eine *Primfaktorzerlegung* von $a \in A$, wenn $a = c\mathbb{P}_A^{\alpha}$.

Beispiel 16.4.11. Mit $\mathbb{P}_{\mathbb{Z}} = \mathbb{P}$ ist $(-1, \alpha)$ eine Primfaktorzerlegung von -20 in $\mathbb{Z}$, wenn man $\alpha \in \mathbb{N}_0^{(\mathbb{P})}$ definiert durch $\operatorname{supp}(\alpha) = \{2, 5\}$, $\alpha(2) = 2$ und $\alpha(5) = 1$. Informell würde man auch sagen, $-20 = (-1) \cdot 2^2 \cdot 5$ ist eine Primfaktorzerlegung von -20.

Lemma 16.4.12. *Sei A ein Integritätsring und $p, q \in \mathbb{P}_A$ mit $p \mid q$. Dann $p = q$.*

Beweis. Schreibe $q = pa$ mit $a \in A$. Wegen $q \mid pa$ gilt $q \mid p$ oder $q \mid a$. Falls $q \mid p$, so $p \mathrel{\widehat{=}} q$ und daher $p = q$. Es reicht also, die Annahme $q \mid a$ zum Widerspruch zu führen. Wäre aber $b \in A$ mit $a = qb$, so folgte $q = pa = pqb$ und daher $1 = pb$ wegen $q \neq 0$. Dann wäre aber $p \in A^{\times}$ und damit p nicht prim. □

Proposition 16.4.13. *In Integritätsringen sind Primfaktorzerlegungen eindeutig, das heißt ist A ein Integritätsring und sind $(c, \alpha), (d, \beta) \in A^{\times} \times \mathbb{N}_0^{(\mathbb{P}_A)}$ mit $c\mathbb{P}_A^{\alpha} = d\mathbb{P}_A^{\beta}$, so folgt $(c, \alpha) = (d, \beta)$.*

Beweis. Induktion nach der Anzahl der Primfaktoren in der ersten Primfaktorzerlegung, das heißt wir zeigen

$$\forall (c, \alpha), (d, \beta) \in A^{\times} \times \mathbb{N}_0^{(\mathbb{P}_A)} : \left(\left(\sum_{p \in \operatorname{supp}(\alpha)} \alpha(p) = n \;\&\; c\mathbb{P}_A^{\alpha} = d\mathbb{P}_A^{\beta} \right) \implies (c, \alpha) = (d, \beta) \right)$$

durch Induktion nach $n \in \mathbb{N}_0$.

$\underline{n = 0}$ Seien $(c, \alpha), (d, \beta) \in A^{\times} \times \mathbb{N}_0^{(P_A)}$ mit $\alpha = 0$ und $c\mathbb{P}_A^{\alpha} = d\mathbb{P}_A^{\beta}$. Dann $\mathbb{P}_A^{\beta} = cd^{-1} \in A^{\times}$. Da kein Primelement Einheit ist, folgt $\beta = 0$ und $c = d$.

$\underline{n - 1 \to n \quad (n \in \mathbb{N})}$ Seien $(c, \alpha), (d, \beta) \in A^{\times} \times \mathbb{N}_0^{(\mathbb{P}_A)}$ mit $\sum_{p \in \operatorname{supp}(\alpha)} \alpha(p) = n$ und $c\mathbb{P}_A^{\alpha} = d\mathbb{P}_A^{\beta}$. Wähle $p \in \operatorname{supp}(\alpha)$. Wegen $p \mid \mathbb{P}_A^{\beta}$ gibt es ein $q \in \operatorname{supp}(\beta)$ mit $p \mid q$, denn p ist prim. Gemäß Lemma 16.4.12 gilt $p = q$. Definiere $\alpha', \beta' \in \mathbb{N}_0^{(\mathbb{P}_A)}$ durch

$$\alpha'|_{\mathbb{P}_A \setminus \{p\}} := \alpha|_{\mathbb{P}_A \setminus \{p\}},\ \beta'|_{\mathbb{P}_A \setminus \{p\}} := \beta|_{\mathbb{P}_A \setminus \{p\}},\ \alpha'(p) := \alpha(p) - 1 \text{ und } \beta'(p) := \beta(p) - 1.$$

Es folgt $c\mathbb{P}_A^{\alpha'} = d\mathbb{P}_A^{\beta'}$ und nach IV daher $(c, \alpha') = (d, \beta')$. Somit auch $(c, \alpha) = (d, \beta)$. □

Definition 16.4.14. Ein Integritätsring A heißt *faktoriell*, wenn jedes $a \in A \setminus \{0\}$ eine Primfaktorzerlegung in A besitzt, das heißt es gibt $(c, \alpha) \in A^\times \times \mathbb{N}_0^{(\mathbb{P}_A)}$ mit $a = c\mathbb{P}_A^\alpha$.

Bemerkung 16.4.15. Sei A ein Integritätsring.

(a) Definition 16.4.14 ist wegen Proposition 16.2.4 unabhängig von der Wahl von $\mathbb{P}_A$.

(b) Proposition 16.4.13 besagt, dass

$$A^\times \times \mathbb{N}_0^{(\mathbb{P}_A)} \to A \setminus \{0\},\ (c, \alpha) \mapsto c\mathbb{P}_A^\alpha$$

injektiv ist. Gemäß Definition 16.4.14 ist A genau dann faktoriell wenn diese Abbildung auch surjektiv (und damit bijektiv) ist.

Notation 16.4.16. Wir führen auf $\mathbb{N}_0^{(\mathbb{P}_A)}$ die Halbordnung $\preceq$ definiert durch

$$\alpha \preceq \beta :\iff \forall p \in \mathbb{P}_A : \alpha(p) \leq \beta(p) \qquad (\alpha, \beta \in \mathbb{N}_0^{(\mathbb{P}_A)})$$

ein.

Proposition 16.4.17. *Sei A ein faktorieller Ring. Dann gilt für alle $(c, \alpha), (d, \beta) \in A^\times \times \mathbb{N}_0^{(\mathbb{P}_A)}$*

$$c\mathbb{P}_A^\alpha \mid d\mathbb{P}_A^\beta \iff \alpha \preceq \beta.$$

Beweis. „$\Longleftarrow$“ Ist $\alpha \preceq \beta$, so ist $\gamma := \beta - \alpha \in \mathbb{N}_0^{(\mathbb{P}_A)}$ und $d\mathbb{P}_A^\beta = \left(\frac{d}{c}\mathbb{P}_A^\gamma\right)(c\mathbb{P}_A^\alpha)$.

„$\Longrightarrow$“ Gelte $c\mathbb{P}_A^\alpha \cdot a = d\mathbb{P}_A^\beta$ für ein $a \in A$. Zu zeigen: $\alpha \preceq \beta$. Da A faktoriell und $a \neq 0$ ist, gibt es $(e, \gamma) \in A^\times \times \mathbb{N}_0^{(\mathbb{P}_A)}$ mit $a = e\mathbb{P}_A^\gamma$. Es folgt $ce\mathbb{P}_A^{\alpha+\gamma} = d\mathbb{P}_A^\beta$ und wegen 16.4.13 daher $\alpha + \gamma = \beta$ (und $ce = d$). □

Satz 16.4.18. *Sei A ein Integritätsring. Dann ist A faktoriell genau dann, wenn in A jedes irreduzible Element prim ist und die folgende „Teilerkettenbedingung“ gilt: Ist $(a_n)_{n\in\mathbb{N}}$ eine Folge von Elementen $a_n \in A$ mit $(a_1) \subseteq (a_2) \subseteq (a_3) \subseteq \ldots$, so gibt es ein $k \in \mathbb{N}$ mit $(a_k) = (a_{k+1}) = (a_{k+2}) = \ldots$.*

Beweis. Sei zunächst A faktoriell. Da ein irreduzibles Element per Definition keine Einheit ist, muss in seiner Primfaktorzerlegung mindestens ein Primfaktor auftauchen. Da Primelemente keine Einheiten sind, kann aber dort auch höchstens ein Primfaktor auftauchen. Sei nun $(a_n)_{n\in\mathbb{N}}$ eine Folge in A mit $(a_1) \subseteq (a_2) \subseteq (a_3) \subseteq \ldots$ Œ $a_1 \neq 0$ und damit $\forall n \in \mathbb{N} : a_n \neq 0$. Schreibe $a_n = c_n\mathbb{P}_A^{\alpha_n}$ mit $(c_n, \alpha_n) \in A^\times \times \mathbb{N}_0^{(\mathbb{P}_A)}$ für alle $n \in \mathbb{N}$. Mit Proposition 16.4.17 folgt $\alpha_1 \succeq \alpha_2 \succeq \alpha_3 \succeq \ldots$. Weil $\operatorname{supp}(\alpha_1)$ endlich ist, folgt hieraus, dass es $k \in \mathbb{N}$ gibt mit $\alpha_k = \alpha_{k+1} = \ldots$ Dann $(a_k) = (a_{k+1}) = \ldots$

Sei nun umgekehrt jedes irreduzible Element prim und gelte die Teilerkettenbedingung. Es reicht dann zu zeigen, dass jedes $a \in A \setminus \{0\}$ ein Produkt von irreduziblen Elementen und einer Einheit ist. Sei M die durch Inklusion halbgeordnete Menge von Hauptidealen (a) derart, dass $a \in A \setminus \{0\}$ kein solches Produkt ist.

Annahme: $M \neq \emptyset$.

Wegen der Teilerkettenbedingung besitzt M mindestens ein maximales Element [→ 12.2.2] (a) mit $a \in A \setminus \{0\}$, welches kein solches Produkt ist. Insbesondere ist a weder irreduzibel noch eine Einheit ist, weswegen es $b, c \in A \setminus A^\times$ gibt mit $a = bc$. Es folgt $(a) \subset (b)$ und $(a) \subset (c)$ (wäre etwa $(a) = (b)$, das heißt $a \mathrel{\widehat{=}} b$, so gäbe es nach Proposition 16.2.4 ein $c' \in A^\times$ mit $a = bc'$ und es folgte $c = c' \in A^\times$ wegen $b \neq 0$). Wegen der Maximalität von (a) sind sowohl b als auch c Produkte von irreduziblen Elementen und einer Einheit. Dann aber auch a. $\lightning$ □

Korollar 16.4.19. *Jeder Hauptidealring ist faktoriell.*

Beweis. Sei A ein Hauptidealring. Zu zeigen:

(a) In A ist jedes irreduzible Element prim.

(b) Teilerkettenbedingung

Zu (a). Sei $p \in A$ irreduzibel. Zu zeigen: p prim. Per Definition ist $p \notin A^\times$. Seien $a, b \in A$ mit $p \mid ab$. Zu zeigen: $p \mid a$ oder $p \mid b$. Wähle $c \in A$ mit $(c) = (p, a)$. Wegen $p \in (c)$ und der Irreduzibilität von p folgt $c \in A^\times$ oder $p \mathrel{\widehat{=}} c$. Falls $p \mathrel{\widehat{=}} c$, so $p \mid c$ und $c \mid a$, also $p \mid a$. Gelte also $c \in A^\times$. Dann $(1) = (p, a)$ und es gibt $s, t \in A$ mit $1 = sp + ta$. Dann $b = sbp + tab$ und mit $p \mid (sbp + tab)$ (wegen $p \mid ab$) folgt $p \mid b$.

Zu (b). Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in A mit $(a_1) \subseteq (a_2) \subseteq (a_3) \subseteq \ldots$ Dann ist $I := \bigcup\{(a_n) \mid n \in \mathbb{N}\}$ ein Ideal von A, wie man sich leicht überlegt. Wähle $a \in A$ mit $I = (a)$. Wähle $k \in \mathbb{N}$ mit $a \in (a_k)$. Dann $(a_k) \subseteq (a_{k+1}) \subseteq \ldots \subseteq I = (a_k)$, also $(a_k) = (a_{k+1}) = \ldots$ □

Beispiel 16.4.20. (a) $\mathbb{Z}$ ist faktoriell: Für alle $a \in \mathbb{Z} \setminus \{0\}$ gibt es genau ein $(c, \alpha) \in \{-1, 1\} \times \mathbb{N}_0^{(\mathbb{P})}$ mit $a = c\mathbb{P}^\alpha$.

(b) Sei K ein Körper. Dann ist $K[X]$ faktoriell. Insbesondere sind die Primelemente $p \in K[X] \setminus \{0\}$ genau die irreduziblen Polynome in $K[X]$. Man setzt in Übereinstimmung mit 16.4.9(b)

$$\mathbb{P}_{K[X]} = \{p \in K[X] \mid p \text{ normiert und irreduzibel}\},$$

sofern nichts anderes erwähnt ist. Es gilt[11]

$$\mathbb{P}_{\mathbb{C}[X]} = \{X - z \mid z \in \mathbb{C}\} \qquad \text{und}$$
$$\mathbb{P}_{\mathbb{R}[X]} = \{X - a \mid a \in \mathbb{R}\} \cup \{(X + a)^2 + c \mid a, c \in \mathbb{R}, c > 0\}.$$

Letzteres folgt, da in der Primfaktorzerlegung eines $p \in \mathbb{R}[X]$ in $\mathbb{C}[X]$ mit jedem $X + a + \mathring{\imath}b \in \mathbb{P}_{\mathbb{C}[X]}$ $(a, b \in \mathbb{R}, b \neq 0)$ auch $X + a - \mathring{\imath}b \in \mathbb{P}_{\mathbb{C}[X]}$ auftaucht und mit $c := b^2 > 0$ gilt $(X + a + \mathring{\imath}b)(X + a - \mathring{\imath}b) = (X + a)^2 + c$ (vergleiche auch 10.1.13 und Beweis von 15.2.10).

[11] Hier benutzen wir den Fundamentalsatz der Algebra 4.2.12 [→ 4.2.13(g)].

(c) Sei K ein Körper und $p \in K[X] \setminus \{0\}$ mit Primfaktorzerlegung $p = c\mathbb{P}_{K[X]}^{\alpha}$ ($c \in A^\times, \alpha \in \mathbb{N}_0^{(\mathbb{P}_A)}$). Ist $\lambda \in K$, so ist λ eine Nullstelle von p genau dann, wenn $\alpha(X - \lambda) \geq 1$. In diesem Fall ist $\alpha(X - \lambda)$ die in 10.1.13 definierte Vielfachheit der Nullstelle λ von p.

Notation 16.4.21. Wir fügen zu $\mathbb{N}_0^{(P_A)}$ ein neues Element ∞ hinzu und erweitern die Halbordnung $\preceq$ aus 16.4.16 auf $\mathbb{N}_0^{(P_A)} \cup \{\infty\}$, indem wir festlegen, dass ∞ das größte Element der Halbordnung wird:

$$\alpha \preceq \beta :\Longleftrightarrow (\beta = \infty \text{ oder } (\alpha \neq \infty \neq \beta \;\&\; \forall p \in \mathbb{P}_A : \alpha(p) \leq \beta(p))) \qquad (\alpha, \beta \in \mathbb{N}_0^{(\mathbb{P}_A)} \cup \{\infty\})$$

Proposition 16.4.22. *Sei A faktoriell. Dann ist die Abbildung*

$$\mathbb{N}_0^{(\mathbb{P}_A)} \cup \{\infty\} \to A/\widehat{=}, \quad \alpha \mapsto \begin{cases} \widehat{0} & \text{falls } \alpha = \infty \\ \widehat{P_A^\alpha} & \text{sonst} \end{cases}$$

ein Isomorphismus halbgeordneter Mengen [→ 16.1.5, 16.1.3, 16.4.21].

Beweis. surjektiv Sei $a \in A$. Ist $a = 0$, so ist $\widehat{a}$ Bild von ∞. Sei also $a \neq 0$ Schreibe $a = c\mathbb{P}_A^\alpha$ mit $c \in A^\times$ und $\alpha \in \mathbb{N}_0^{(P_A)}$. Dann $a \mathrel{\widehat{=}} \mathbb{P}_A^\alpha$, also ist $\widehat{a} = \widehat{\mathbb{P}_A^\alpha}$ Bild von α.

injektiv $\widehat{0}$ hat nur ∞ als Urbild. Sind $\alpha, \beta \in \mathbb{N}_0^{(\mathbb{P}_A)}$ mit $\widehat{\mathbb{P}_A^\alpha} = \widehat{\mathbb{P}_A^\beta}$, so $\mathbb{P}_A^\alpha \mathrel{\widehat{=}} \mathbb{P}_A^\beta$ und es gibt $c \in A^\times$ mit $c\mathbb{P}_A^\alpha = \mathbb{P}_A^\beta$. Aus 16.4.13 ergibt sich $\alpha = \beta$ (und $c = 1$).

Isomorphismus ∞ ist das größte Element von $\mathbb{N}_0^{(\mathbb{P}_A)} \cup \{\infty\}$ und sein Bild $\widehat{0}$ das größte Element von $A/\widehat{=}$. Für $\alpha, \beta \in \mathbb{N}_0^{(\mathbb{P}_A)}$ gilt

$$\alpha \preceq \beta \iff \mathbb{P}_A^\alpha \mid \mathbb{P}_A^\beta \iff \widehat{\mathbb{P}_A^\alpha} \preceq \widehat{\mathbb{P}_A^\beta}.$$

□

In Verallgemeinerung [→ 16.4.19] von Proposition 16.2.3(b) halten wir fest:

Korollar 16.4.23. *Ist A ein faktorieller Ring und $B \subseteq A$, so gibt es einen ggT und ein kgV der Elemente von B.*

Beweis. Lauf Definition 16.1.4 ist zu zeigen, dass in der halbgeordneten Menge

$$(A/\widehat{=}, \preceq)$$

jede Teilmenge ein Infimum und ein Supremum besitzt. Nach der letzten Proposition reicht es, dasselbe für die halbgeordnete Menge

$$(\mathbb{N}_0^{(\mathbb{P}_A)} \cup \{\infty\}, \preceq)$$

zu zeigen. Sei also $B \subseteq \mathbb{N}_0^{(\mathbb{P}_A)} \cup \{\infty\}$. Setze

$$\alpha_{\mathrm{ggT}} := \begin{cases} \infty & \text{falls } B \subseteq \{\infty\}, \\ \begin{pmatrix} \mathbb{P}_A \to \mathbb{N}_0 \\ p \mapsto \min\{\beta(p) \mid \beta \in B \setminus \{\infty\}\} \end{pmatrix} & \text{falls } B \not\subseteq \{\infty\}. \end{cases}$$

Dann ist α_{ggT} nach 12.1.7(c) das Infimum von B, denn es gilt

$$\forall \beta \in \mathbb{N}_0^{(\mathbb{P}_A)} \cup \{\infty\} : (\beta \text{ untere Schranke von } B \iff \beta \preceq \alpha_{\mathrm{ggT}}).$$

Setze weiter

$$\alpha_{\mathrm{kgV}} := \begin{cases} 0 & \text{falls } B = \emptyset, \\ \infty & \text{falls } B \text{ keine obere Schranke } \beta \in \mathbb{N}_0^{(\mathbb{P}_A)} \text{ besitzt}, \\ \begin{pmatrix} \mathbb{P}_A \to \mathbb{N}_0 \\ p \mapsto \max\{\beta(p) \mid \beta \in B \setminus \{\infty\}\} \end{pmatrix} & \text{sonst.} \end{cases}$$

Dann

$$\forall \beta \in \mathbb{N}_0^{(\mathbb{P}_A)} \cup \{\infty\} : (\beta \text{ obere Schranke von } B \iff \alpha_{\mathrm{kgV}} \preceq \beta),$$

weswegen α_{kgV} das Supremum von B ist. □

Obiger Beweis zeigt, wie man in einem faktoriellen Ring ggT und kgV einer Menge von Elementen, deren Primfaktorzerlegungen man kennt, sofort berechnen kann.

Beispiel 16.4.24. $\gcd\{3^{17}5^{13}, 3^{14}5^{9}7\} = 3^{14}5^{9}$ und $\mathrm{lcm}\{3^{17}5^{13}, 3^{14}5^{9}7\} = 3^{17}5^{13}7$ in $\mathbb{Z}$