Lineare Algebra II

Sommersemester 2015

Wolfgang Ebeling

Korrigierte Fassung

Institut für Algebraische Geometrie
Leibniz Universität Hannover
Postfach 6009
30060 Hannover
E-mail: ebeling@math.uni-hannover.de

1 Summen von Vektorräumen

Im Folgenden sei K zunächst ein beliebiger Körper. Wir betrachten verschiedene Summen von Vektorräumen.

Definition Es sei V ein K-Vektorraum und U_1, U_2 Unterräume von V. Dann heißt

$$U_1 + U_2 := \{u_1 + u_2 \mid u_1 \in U_1, u_2 \in U_2\}$$

die *Summe* von U_1 und U_2.

Lemma 1.1 *Für die oben definierte Summe U_1+U_2 der Unterräume U_1 und U_2 gilt:*

(i) $U_1 + U_2 = \mathrm{Span}(U_1 \cup U_2)$.

(ii) $U_1 + U_2 \subset V$ *ist ein Unterraum.*

(iii) $\dim(U_1 + U_2) \leq \dim U_1 + \dim U_2$.

Beweis. (i) $U_1 + U_2 \subset \mathrm{Span}(U_1 \cup U_2)$ ist klar. Zum Beweis der umgekehrten Inklusion sei $v \in \mathrm{Span}(U_1 \cup U_2)$. Dann gibt es $u_1, \ldots, u_k \in U_1$, $w_1, \ldots, w_m \in U_2$ und $\lambda_1, \ldots, \lambda_k, \mu_1, \ldots, \mu_m \in K$ mit

$$v = \lambda_1 u_1 + \cdots + \lambda_k u_k + \mu_1 w_1 + \cdots + \mu_m w_m.$$

Setze $v_1 := \lambda_1 u_1 + \cdots + \lambda_k u_k$ und $v_2 := \mu_1 w_1 + \cdots + \mu_m w_m$. Dann ist $v_1 \in U_1$ und $v_2 \in U_2$, also $v = v_1 + v_2 \in U_1 + U_2$.

(ii) folgt aus (i).

(iii) Ist $u_1, \ldots, u_k$ eine Basis von U_1 und $w_1, \ldots, w_m$ eine Basis von U_2, so ist $u_1, \ldots, u_k, w_1, \ldots, w_m$ ein Erzeugendensystem von $\mathrm{Span}(U_1 \cup U_2)$. Damit folgt die Behauptung aus (i). $\square$

Satz 1.1 (Dimensionsformel für Summen) *Für endlich dimensionale Unterräume $U_1, U_2 \subset V$ gilt*

$$\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2).$$

Beweis. Es sei $\{v_1, \ldots, v_m\}$ eine Basis von $U_1 \cap U_2$. Wir ergänzen diese Basis zu Basen

$$\{v_1, \ldots, v_m, u_1, \ldots, u_k\} \text{ von } U_1 \text{ und } \{v_1, \ldots, v_m, u'_1, \ldots, u'_\ell\} \text{ von } U_2.$$

Wir müssen zeigen:

$$B := \{v_1, \ldots, v_m, u_1, \ldots, u_k, u'_1, \ldots, u'_\ell\} \text{ ist eine Basis von } U_1 + U_2.$$

Es ist klar, dass B ein Erzeugendensystem von $\mathrm{Span}(U_1 \cup U_2)$ und nach Lemma 1.1(i) damit von $U_1 + U_2$ ist. Wir müssen also noch zeigen, dass B linear unabhängig ist. Dazu sei

$$\lambda_1 v_1 + \cdots + \lambda_m v_m + \mu_1 u_1 + \cdots + \mu_k u_k + \mu'_1 u'_1 + \cdots + \mu'_\ell u'_\ell = 0.$$

Wir setzen

$$v := \lambda_1 v_1 + \cdots + \lambda_m v_m + \mu_1 u_1 + \cdots + \mu_k u_k.$$

Dann ist $v \in U_1$ und $-v = \mu'_1 u'_1 + \cdots + \mu'_\ell u'_\ell \in U_2$. Daraus folgt $v \in U_1 \cap U_2$. Also ist

$$v = \lambda'_1 v_1 + \cdots + \lambda'_m v_m$$

für gewisse Skalare $\lambda'_1, \ldots, \lambda'_m$. Da $\{v_1, \ldots, v_m, u_1, \ldots, u_k\}$ eine Basis von U_1 bildet, folgt aus der Eindeutigkeit der Darstellung von v als Linearkombination der Vektoren dieser Basis insbesondere $\mu_1 = \cdots = \mu_k = 0$. Setzen wir dies in die obige Gleichung ein, so folgt

$$\lambda_1 = \cdots = \lambda_m = \mu'_1 = \cdots = \mu'_\ell = 0.$$

□

Lemma 1.2 *Ist $V = U_1 + U_2$, so sind die folgenden Aussagen äquivalent:*

(i) $U_1 \cap U_2 = \{0\}$.

(ii) *Jedes $v \in V$ lässt sich eindeutig darstellen als $v = u_1 + u_2$ mit $u_1 \in U_1$ und $u_2 \in U_2$.*

(iii) *Je zwei vom Nullvektor verschiedene Vektoren $u_1 \in U_1$ und $u_2 \in U_2$ sind linear unabhängig.*

Beweis. (i) ⇒ (ii): Es sei

$$v = u_1 + u_2 = u'_1 + u'_2 \quad (u_1, u'_1 \in U_1, u_2, u'_2 \in U_2).$$

Dann folgt $u_1 - u'_1 = u'_2 - u_2 \in U_1 \cap U_2$, nach (i) also $u_1 - u'_1 = u'_2 - u_2 = 0$.

(ii) ⇒ (iii): Nach (ii) besitzt der Nullvektor eine eindeutige Darstellung $0 = 0u_1 + 0u_2$.

(iii) ⇒ (i): Es sei $0 \neq v \in U_1 \cap U_2$. Dann sind nach (iii) v und $-v$ linear unabhängig im Widerspruch zu $1v + (-1)v = 0$. □

Ist eine der drei äquivalenten Bedingungen von Lemma 1.2 erfüllt, so heißt V die *direkte Summe* von U_1 und U_2. Also gilt z. B.:

Definition Ein Vektorraum V heißt *direkte Summe* von zwei Unterräumen U_1 und U_2, in Zeichen $V = U_1 \oplus U_2$, wenn

$$V = U_1 + U_2 \text{ und } U_1 \cap U_2 = \{0\}.$$

Beispiel 1.1 Es sei $V = \mathbb{R}^3$. Ist $U_1 = \operatorname{Span}\{e_1, e_2\}$ und $U_2 = \operatorname{Span}\{e_3\}$, so ist $V = U_1 \oplus U_2$. Ist dagegen $U_3 = \operatorname{Span}\{e_2, e_3\}$, so ist zwar $V = U_1 + U_3$, die Summe ist aber nicht direkt, da $U_1 \cap U_3 = \operatorname{Span}\{e_2\}$.

Satz 1.2 *Es sei V ein endlich dimensionaler K-Vektorraum und U_1, U_2 Untervektorräume von V. Dann sind die folgenden Bedingungen äquivalent:*

(i) $V = U_1 \oplus U_2$.

(ii) *Es gibt Basen $\{u_1, \ldots, u_k\}$ von U_1 und $\{u'_1, \ldots, u'_\ell\}$ von U_2, so dass $\{u_1, \ldots, u_k, u'_1, \ldots, u'_\ell\}$ eine Basis von V ist.*

(iii) *Es gilt $V = U_1 + U_2$ und* $\dim V = \dim U_1 + \dim U_2$.

Beweis. (i) $\Rightarrow$ (ii) folgt aus dem Beweis der Dimensionsformel (Satz 1.1) und der Tatsache, dass $U_1 \cap U_2 = \{0\}$.

(ii) $\Rightarrow$ (iii) ist klar.

(iii) $\Rightarrow$ (i): Aus der Dimensionsformel folgt $\dim(U_1 \cap U_2) = 0$. Daraus folgt $U_1 \cap U_2 = \{0\}$. □

Definition Es sei $U \subset V$ ein Unterraum. Ein Unterraum $W \subset V$ heißt *Komplement* von U in V, falls

$$U \oplus W = V.$$

Bemerkung 1.1 Zu einem Unterraum U ist ein Komplement W im Allgemeinen nicht eindeutig bestimmt: Ist zum Beispiel $V = \mathbb{R}^3$ und $U = \operatorname{Span}\{e_1, e_2\}$, so sind $W_1 = \operatorname{Span}\{e_3\}$ und $W_2 = \operatorname{Span}\{e_1 + e_3\}$ Komplemente von U.

Satz 1.3 *Ist V endlich dimensional und $U \subset V$ ein Unterraum, so besitzt U ein Komplement in V.*

Beweis. Man nehme eine Basis $\{v_1, \ldots, v_r\}$ von U und ergänze sie nach I, Satz 11.4, zu einer Basis $\{v_1, \ldots, v_r, v_{r+1}, \ldots, v_n\}$ von V. Man setze

$$W := \operatorname{Span}\{v_{r+1}, \ldots, v_n\}.$$

□

Nun sei $K = \mathbb{R}$ oder $K = \mathbb{C}$ und $(V, \langle\ ,\ \rangle)$ ein euklidischer ($K = \mathbb{R}$) oder unitärer ($K = \mathbb{C}$) Vektorraum .

Definition Es sei V ein euklidischer (unitärer) Vektorraum.

(a) Zwei Vektoren $v, w \in V$ heißen *orthogonal*, in Zeichen $v \perp w$, falls $\langle v.w \rangle = 0$.

(b) Zwei Unterräume $U, W \subset V$ heißen *orthogonal*, in Zeichen $U \perp W$, falls $u \perp w$ für alle $u \in U$, $w \in W$.

(c) Ist U ein Unterraum, so heißt

$$U^\perp := \{v \in V \mid \langle u, v \rangle = 0 \text{ für alle } u \in U\}$$

das *orthogonale Komplement* von U.

Bemerkung 1.2 Das orthogonale Komplement $U^\perp$ eines Unterraums U ist ein Unterraum. Aus dem unten stehenden Satz 1.4 folgt, dass es ein Komplement von U ist.

Definition Es sei V ein euklidischer (unitärer) Vektorraum und U_1, U_2 Unterräume von V. Man sagt, V ist die *orthogonale direkte Summe* der Unterräume U_1 und U_2, in Zeichen $V = U_1 \perp U_2$, falls

(i) $V = U_1 \oplus U_2$ und

(ii) $U_1 \perp U_2$.

Lemma 1.3 *Es sei V ein euklidischer (unitärer) Vektorraum und U_1, U_2 Unterräume von V. Gilt*

(i) $V = U_1 + U_2$ *und*

(ii) $U_1 \perp U_2$,

so ist V die orthogonale direkte Summe der Unterräume U_1 und U_2.

Beweis. Wir haben zu zeigen: $U_1 \cap U_2 = \{0\}$. Es sei $v \in U_1 \cap U_2$, $v \neq 0$. Wegen (i) ist $v = u_1 + u_2$ mit $u_1 \in U_1$ und $u_2 \in U_2$. Dann gilt

$$0 \neq \langle v, v \rangle = \langle v, u_1 \rangle + \langle v, u_2 \rangle = 0 \quad \text{wegen (ii)},$$

ein Widerspruch. □

Satz 1.4 *Es sei V ein endlich dimensionaler euklidischer (unitärer) Vektorraum und $W \subset V$ ein Unterraum. Dann gilt*

$$V = W \perp W^\perp.$$

Insbesondere ist

$$\dim V = \dim W + \dim W^\perp.$$

Beweis. Es sei $\{w_1, \dots, w_m\}$ eine ON-Basis von W. Diese ergänze man nach I, Satz 20.2, zu einer ON-Basis $\{w_1, \dots, w_m, w_{m+1}, \dots, w_n\}$ von V. Dann ist $w_{m+1}, \dots, w_n \in W^\perp$. Es sei nun $v \in V$. Dann können wir v schreiben als

$$v = \lambda_1 w_1 + \cdots + \lambda_m w_m + \lambda_{m+1} w_{m+1} + \cdots + \lambda_n w_n$$

mit

$$\lambda_1 w_1 + \cdots + \lambda_m w_m \in W, \quad \lambda_{m+1} w_{m+1} + \cdots + \lambda_n w_n \in W^\perp.$$

Daraus folgt die Behauptung. □

Wir wollen nun auch Summen von mehr als zwei Unterräumen betrachten.

Definition Es sei V ein K-Vektorraum und $U_1, \dots, U_s$ Unterräume von V. Dann heißt

$$U_1 + \cdots + U_s := \{u_1 + \cdots + u_s \,|\, u_i \in U_i, i = 1, \dots, s\}$$

die *Summe* von $U_1, \dots, U_s$.

Wie oben beweist man:

Lemma 1.4 *Für die Summe $U_1 + \cdots + U_s$ der Unterräume $U_1, \dots, U_s$ gilt:*

(i) $U_1 + \cdots + U_s = \mathrm{Span}(U_1 \cup \cdots \cup U_s)$.

(ii) $U_1 + \cdots + U_s \subset V$ *ist ein Unterraum.*

(iii) $\dim(U_1 + \cdots + U_s) \leq \dim U_1 + \cdots + \dim U_s$.

Satz 1.5 *Ist $V = U_1 + \cdots + U_s$, so sind die folgenden Aussagen äquivalent:*

(i) *Für jedes $i = 1, \dots, s$ gilt: Ist $W_i := U_1 + \cdots + \widehat{U_i} + \cdots + U_s$, so ist $U_i \cap W_i = \{0\}$. (Hierbei bedeutet $\widehat{U_i}$: In der Summe wird U_i weggelassen, "nimmt seinen Hut und geht".)*

(ii) *Jedes Element $v \in V$ lässt sich eindeutig darstellen als $v = u_1 + \cdots + u_s$ mit $u_i \in U_i$.*

(iii) *Für jede Teilmenge $I \subset \{1, \dots, s\}$ gilt: Ist für $i \in I$ $u_i \in U_i$, $u_i \neq 0$, so ist die Teilmenge $S_I := \{u_i \,|\, i \in I\}$ linear unabhängig.*

Definition Ist eine der äquivalenten Bedingungen von Satz 1.5 erfüllt, so heißt V die *direkte Summe* von $U_1, \dots, U_s$, in Zeichen $V = U_1 \oplus \cdots \oplus U_s$.

Beweis. (i) $\Rightarrow$ (ii): Es sei

$$v = u_1 + \cdots + u_s = u'_1 + \cdots + u'_s \quad (u_i, u'_i \in U_i).$$

Dann folgt

$$u_i - u'_i = (u'_1 - u_1) + \cdots + (u'_{i-1} - u_{i-1}) + (u'_{i+1} - u_{i+1}) + \cdots + (u'_s - u_s) \in W_i \cap U_i.$$

Nach (i) folgt $u_i - u'_i = 0$.

(ii) $\Rightarrow$ (iii): Es sei $I = \{i_1, \ldots, i_r\} \subset \{1, \ldots, s\}$ und

$$\lambda_1 u_{i_1} + \cdots + \lambda_r u_{i_r} = 0.$$

Da nach (ii) auch der Nullvektor $0 \in V$ eine eindeutige Darstellung

$$0 = 0u_{i_1} + \cdots + 0u_{i_r}$$

besitzt, folgt $\lambda_1 = \ldots = \lambda_r = 0$.

(iii) $\Rightarrow$ (i): Es sei $W_i \cap U_i \neq \{0\}$. Dann gibt es ein $u_i \in U_i$ mit $u_i \neq 0$ und

$$u_i = u_1 + \cdots + u_{i-1} + u_{i+1} + \cdots + u_s \text{ mit } u_j \in U_j.$$

Es sei I die Menge aller Indizes $j \in \{1, \ldots, s\}$ mit $u_j \neq 0$. Dann ist die Teilmenge S_I linear abhängig im Widerspruch zu (iii). □

Satz 1.6 *Es sei V ein endlich dimensionaler K-Vektorraum und $U_1, \ldots, U_s$ Unterräume von V. Dann sind die folgenden Bedingungen äquivalent:*

(i) $V = U_1 \oplus \cdots \oplus U_s$.

(ii) *Ist für jedes $i \in \{1, \ldots, s\}$ eine Basis $\{u_1^{(i)}, \ldots, u_{k_i}^{(i)}\}$ von U_i gegeben, so ist*

$$\{u_1^{(1)}, \ldots, u_{k_1}^{(1)}, \ldots, u_1^{(s)}, \ldots, u_{k_s}^{(s)}\}$$

eine Basis von V.

(iii) *Es gilt $V = U_1 + \cdots + U_s$ und* $\dim V = \dim U_1 + \cdots + \dim U_s$.

Beweis. (i) $\Rightarrow$ (ii): Es sei

$$B := \{u_1^{(1)}, \ldots, u_{k_1}^{(1)}, \ldots, u_1^{(s)}, \ldots, u_{k_s}^{(s)}\}.$$

Offensichtlich ist B ein Erzeugendensystem von V. Es reicht daher zu zeigen, dass B linear unabhängig ist. Dazu sei

$$\lambda_1^{(1)} u_1^{(1)} + \cdots + \lambda_{k_1}^{(1)} u_{k_1}^{(1)} + \cdots + \lambda_1^{(s)} u_1^{(s)} + \cdots + \lambda_{k_s}^{(s)} u_{k_s}^{(s)} = 0.$$

Setzen wir $w_i := \lambda_1^{(i)} u_1^{(i)} + \cdots + \lambda_{k_i}^{(i)} u_{k_i}^{(i)}$, so folgt

$$w_1 + \cdots + w_s = 0.$$

Aus Satz 1.5 (iii) folgt $w_1 = \ldots = w_s = 0$. Also ist

$$\lambda_1^{(i)} u_1^{(i)} + \cdots + \lambda_{k_i}^{(i)} u_{k_i}^{(i)} = 0 \text{ für } i = 1, \ldots, s.$$

Daraus folgt $\lambda_1^{(i)} = \cdots = \lambda_{k_i}^{(i)} = 0$.

(ii) $\Leftrightarrow$ (iii) ist klar.

(ii) $\Rightarrow$ (i) folgt aus Satz 1.5 (ii). □

Es sei nun wieder $K = \mathbb{R}, \mathbb{C}$ und V ein euklidischer (unitärer) Vektorraum.

Definition Es sei V ein euklidischer (unitärer) Vektorraum und $U_1, \ldots, U_s$ Unterräume von V. Man sagt, V ist die *orthogonale direkte Summe* der Unterräume $U_1 \ldots, U_s$, in Zeichen $V = U_1 \perp \ldots \perp U_s$, falls

(i) $V = U_1 \oplus \cdots \oplus U_s$ und

(ii) $U_i \perp U_j$ für $i \neq j$.

Lemma 1.5 *Es sei V ein euklidischer (unitärer) Vektorraum und $U_1, \ldots, U_s$ Unterräume von V. Gilt*

(i) $V = U_1 + \cdots + U_s$ und

(ii) $U_i \perp U_j$ für $i \neq j$,

so ist V die orthogonale direkte Summe der Unterräume $U_1, \ldots, U_s$.

Beweis. Wir haben zu zeigen: $U_i \cap W_i = \{0\}$. Es sei $v \in U_i \cap W_i$, $v \neq 0$. Dann gilt

$$v = u_1 + \cdots + u_{i-1} + u_{i+1} + \cdots + u_s \text{ mit } u_j \in U_j.$$

Da $U_i \perp U_j$ für $i \neq j$ gilt, folgt dann

$$0 \neq \langle v, v \rangle = \langle v, u_1 \rangle + \cdots + \langle v, u_{i-1} \rangle + \langle v, u_{i+1} \rangle + \cdots + \langle v, u_s \rangle = 0,$$

ein Widerspruch. □

2 Normierte Vektorräume

Es sei $K = \mathbb{R}, \mathbb{C}$ und V ein K-Vektorraum. Ist $\langle\ ,\ \rangle : V \times V \to K$ eine symmetrische Bilinearform (hermitesche Sesquilinearform), so erhält man daraus eine Abbildung

$$\begin{array}{rccl} q: & V & \longrightarrow & K \\ & v & \longmapsto & q(v) := \langle v, v \rangle \end{array} .$$

Sie heißt die zu $\langle\ ,\ \rangle$ gehörige *quadratische Form.*

Man kann $\langle\ ,\ \rangle$ aus q zurückgewinnen. Dies nennt man *Polarisierung*:

$$\begin{array}{ll} K = \mathbb{R}: & \langle v, w \rangle = \dfrac{1}{4}\left(q(v+w) - q(v-w)\right), \\ K = \mathbb{C}: & \langle v, w \rangle = \dfrac{1}{4}\left(q(v+w) - q(v-w) + iq(v+iw) - iq(v-iw)\right). \end{array}$$

(Beweis durch Nachrechnen.)

Definition Es sei V ein K-Vektorraum. Unter einer *Norm* auf V versteht man eine Funktion

$$\begin{array}{rccl} ||\ ||: & V & \longrightarrow & \mathbb{R} \\ & v & \longmapsto & ||v|| \end{array}$$

mit folgenden Eigenschaften:

(i) $||v|| \geq 0, ||v|| = 0 \Leftrightarrow v = 0$,

(ii) $||\lambda v|| = |\lambda| \cdot ||v||$ für alle $\lambda \in K, v \in V$.

(iii) $||v + w|| \leq ||v|| + ||w||$ für alle $v, w \in V$ (*Dreiecksungleichung*).

Ein *normierter Vektorraum* ist ein Paar $(V, ||\ ||)$, das aus einem Vektorraum V und einer Norm $||\ ||$ auf V besteht.

Definition Es sei X eine Menge. Unter einer *Metrik* auf X versteht man eine Abbildung

$$\begin{array}{rccl} d: & X \times X & \longrightarrow & \mathbb{R} \\ & (x,y) & \longmapsto & d(x,y) \end{array}$$

mit folgenden Eigenschaften:

(i) $d(x,y) = 0 \Leftrightarrow x = y$

(ii) $d(x,y) = d(y,x)$ für alle $x, y \in X$ (*Symmetrie*)

(iii) $d(x,z) \leq d(x,y) + d(y,z)$ für alle $x, y, z \in X$ (*Dreiecksungleichung*).

Ein *metrischer Raum* ist ein Paar (X, d), X Menge, d Metrik. Man nennt $d(x, y)$ den *Abstand* oder die *Distanz* der Punkte x und y bzgl. d.

Bemerkung 2.1 Aus den Axiomen folgt, dass $d(x, y) \geq 0$ für alle $x, y \in X$.

Beweis. Wende Dreiecksungleichung auf x, y, x an:

$$0 \stackrel{\text{(i)}}{=} d(x, x) \leq d(x, y) + d(y, x) \stackrel{\text{(ii)}}{=} 2d(x, y).$$

□

Satz 2.1 *Es sei* $(V, || \ ||)$ *ein normierter Vektorraum. Dann wird durch*

$$d(x, y) := ||x - y|| \text{ für } x, y \in V$$

eine Metrik d *auf* V *definiert.*

Beweis.

(i) $d(x, y) = 0 \Leftrightarrow ||x - y|| = 0 \Leftrightarrow x - y = 0 \Leftrightarrow x = y$.

(ii) $d(x, y) = ||x - y|| = |-1| \, ||x - y|| = ||y - x|| = d(y, x)$.

(iii) $d(x, z) = ||x-z|| = ||x-y+y-z|| \leq ||x-y||+||y-z|| = d(x, y)+d(y, z)$.

□

Satz 2.2 *Ist* $(V, \langle \ , \ \rangle)$ *ein euklidischer (unitärer) Vektorraum, so wird durch*

$$\|v\| := \sqrt{\langle v, v \rangle}$$

eine Norm auf V *definiert.*

Für den Beweis dieses Satzes brauchen wir das folgende Resultat:

Satz 2.3 (Cauchy-Schwarzsche Ungleichung) *Für* $v, w \in V$ *gilt*

$$|\langle v, w \rangle| \leq \|v\| \|w\|$$

und Gleichheit gilt genau dann, wenn v *und* w *linear abhängig sind.*

Beweis. Für $w = 0$ sind beide Seiten der Ungleichung gleich 0, die Ungleichung ist daher erfüllt. Es genügt daher, den Fall $w \neq 0$ zu behandeln.

Für $\lambda, \mu \in K$ gilt

$$\begin{array}{rcl} 0 & \leq & \langle \lambda v + \mu w, \lambda v + \mu w \rangle \\ & = & \lambda \overline{\lambda} \langle v, v \rangle + \lambda \overline{\mu} \langle v, w \rangle + \overline{\lambda} \mu \langle w, v \rangle + \mu \overline{\mu} \langle w, w \rangle. \end{array}$$

Setzen wir nun $\lambda := \langle w, w\rangle$ und $\mu := -\langle v, w\rangle$, so folgt

$$0 \le \lambda(\langle v, v\rangle\langle w, w\rangle - \langle v, w\rangle\overline{\langle v, w\rangle}) = \lambda(\|v\|^2\|w\|^2 - |\langle v, w\rangle|^2).$$

Wegen $\lambda \ge 0$ folgt daraus

$$0 \le \|v\|^2\|w\|^2 - |\langle v, w\rangle|^2.$$

Nun ziehen wir auf beiden Seiten die Quadratwurzel. Dann bleibt das Ungleichheitszeichen erhalten und wir erhalten die behauptete Ungleichung.

Für den Beweis des Zusatzes bemerken wir (für $\lambda := \langle w, w\rangle$ und $\mu := -\langle v, w\rangle$) :

$$\begin{aligned} & |\langle v, w\rangle| = \|v\|\|w\| \\ \Leftrightarrow\quad & \langle \lambda v + \mu w, \lambda v + \mu w\rangle = 0 \\ \Leftrightarrow\quad & \lambda v + \mu w = 0 \\ \Leftrightarrow\quad & v = -\frac{\mu}{\lambda} w. \end{aligned}$$

Man beachte, dass dies der gleiche Beweis wie für I, Satz 3.4 ist, nur dass wir statt einer symmetrischen Bilinearform auch eine hermitesche Sesquilinearform zugelassen haben. □

Beweis von Satz 2.2. (i) und (ii) sind einfach (siehe Vorlesung).

(iii): Um die Dreiecksungleichung

$$\sqrt{\langle v + w, v + w\rangle} \le \sqrt{\langle v, v\rangle} + \sqrt{\langle w, w\rangle}$$

zu beweisen, geht man durch Quadrieren zu der äquivalenten Ungleichung

$$\langle v + w, v + w\rangle \le \langle v, v\rangle + 2\sqrt{\langle v, v\rangle\langle w, w\rangle} + \langle w, w\rangle$$

über, die gleichbedeutend ist mit

$$\langle v, v\rangle + 2|\langle v, w\rangle| + \langle w, w\rangle \le \langle v, v\rangle + 2\sqrt{\langle v, v\rangle\langle w, w\rangle} + \langle w, w\rangle.$$

Diese Ungleichung ist äquivalent zu der Cauchy-Schwarzschen Ungleichung. □

3 Normalform orthogonaler und unitärer Endomorphismen

Nun sei im Folgenden wieder $K = \mathbb{R}, \mathbb{C}$ und $(V, \langle\ ,\ \rangle)$ ein euklidischer (unitärer) Vektorraum. Wir erinnern an die folgende Definition.

Definition Es sei V ein euklidischer (unitärer) Vektorraum. Ein Endomorphismus $f : V \to V$ heißt *orthogonal* (*unitär*), falls gilt:

$$\langle f(v), f(w) \rangle = \langle v, w \rangle \quad \text{für alle } v, w \in V.$$

Theorem 3.1 *Es sei V ein unitärer Vektorraum der Dimension n und $f : V \to V$ ein unitärer Endomorphismus. Dann besitzt V eine ON-Basis, die aus Eigenvektoren von f besteht. Insbesondere ist f diagonalisierbar.*

Beweis. Der zugrundeliegende Körper ist $\mathbb{C}$ und die Eigenwerte von f sind die Nullstellen des charakteristischen Polynoms $P_f(x)$, das ein komplexes Polynom ist. Nach dem Fundamentalsatz der Algebra hat $P_f(x)$ genau n Nullstellen $\lambda_1, \ldots \lambda_n \in \mathbb{C}$. Also gilt

$$P_f(x) = (x - \lambda_1) \cdots (x - \lambda_n).$$

Wir führen nun Induktion nach $n = \dim V$ durch.

Der Induktionsanfang $n = 1$ ist klar.

Wir nehmen nun an, dass die Behauptung bereits für $n - 1$ bewiesen ist. Es sei v_1 ein Eigenvektor von f zum Eigenwert λ_1. Ohne Einschränkung können wir annehmen, dass $\|v_1\| = 1$. Es sei

$$W := \operatorname{Span}\{v_1\}^\perp = \{w \in V \mid \langle v_1, w \rangle = 0\}.$$

Dann gilt nach Satz 1.4

$$V = \operatorname{Span}\{v_1\} \perp W.$$

Behauptung $f(W) = W$.

Beweis. Da f ein Isomorphismus ist, reicht es zu zeigen: $f(W) \subset W$. Nach I, Satz 21.1 (v), gilt $|\lambda_1| = 1$. Damit gilt für $w \in W$:

$$\lambda_1 \langle v_1, f(w) \rangle = \langle \lambda_1 v_1, f(w) \rangle = \langle f(v_1), f(w) \rangle = \langle v_1, w \rangle = 0.$$

Da $\lambda_1 \neq 0$ folgt $\langle v_1, f(w) \rangle = 0$, also $f(w) \in W$. $\square$

Nun betrachten wir den Endomorphismus $f|_W : W \to W$. Da $f|_W$ die Einschränkung eines unitären Endomorphismus ist, ist $f|_W$ auch wieder unitär. Da $\dim W = n - 1$ können wir auf $f|_W : W \to W$ die Induktionsvoraussetzung anwenden. Danach besitzt W eine ON-Basis $\{v_2, \ldots, v_n\}$ aus Eigenvektoren. Dann ist

$$B := \{v_1, v_2, \ldots, v_n\}$$

eine ON-Basis von V aus Eigenvektoren. $\square$

Korollar 3.1 *Eine unitäre Matrix A ist diagonalisierbar. Genauer gilt: Es gibt eine unitäre Matrix S mit*

$$\overline{S}^T AS = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Hierbei gilt $|\lambda_i| = 1$ für $i = 1, \ldots, n$.

Beweis. Siehe Vorlesung. □

In LA I hatten wir bereits orthogonale Abbildungen $f : \mathbb{R}^n \to \mathbb{R}^n$ betrachtet und für $n = 1, 2, 3$ klassifiziert. Allgemeiner wollen wir nun beweisen:

Theorem 3.2 *Es sei V ein euklidischer Vektorraum der Dimension n und $f : V \to V$ ein orthogonaler Endomorphismus. Dann besitzt V eine ON-Basis B, bezüglich der f die Darstellungsmatrix*

$$M_B^B(f) = \begin{pmatrix} +1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & +1 & & & & 0 & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & 0 & & & & \boxed{A_1} & & \\ & & & & & & & \ddots & \\ & & & & & & & & \boxed{A_k} \end{pmatrix},$$

besitzt, wobei für $j = 1, \ldots, k$

$$A_j = \begin{pmatrix} \cos\alpha_j & -\sin\alpha_j \\ \sin\alpha_j & \cos\alpha_j \end{pmatrix} \text{ mit } \alpha_j \in [0, 2\pi), \text{ aber } \alpha_j \neq 0, \pi.$$

Für den Beweis dieses Theorems brauchen wir ein Lemma.

Lemma 3.1 *Jedes Polynom $P(x)$ mit reellen Koeffizienten besitzt eine Zerlegung*

$$P(x) = (x - \lambda_1) \cdots (x - \lambda_r) Q_1(x) \cdots Q_k(x),$$

wobei $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ und $Q_1(x), \ldots, Q_k(x)$ Polynome vom Grad 2 sind, die keine reelle Nullstelle haben.

Beweis. Das Polynom $P(x)$ hat n komplexe Nullstellen. Ist $\lambda \in \mathbb{C}$ eine Nullstelle von $P(x)$, so auch $\overline{\lambda}$:

$$P(\overline{\lambda}) = a_0 + a_1\overline{\lambda} + \cdots + a_n\overline{\lambda}^n = \overline{a_0} + \overline{a_1}\overline{\lambda} + \cdots + \overline{a_n}\overline{\lambda}^n = \overline{P(\lambda)} = \overline{0} = 0.$$

Also hat man eine Zerlegung

$$P(x) = (x - \lambda_1) \cdots (x - \lambda_r)(x - \mu_1)(x - \overline{\mu_1}) \cdots (x - \mu_k)(x - \overline{\mu_k}),$$

wobei $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ und $\mu_1, \ldots, \mu_k \notin \mathbb{R}$. Es sei $j = 1, \ldots, k$ und $\mu_j = \xi_j + i\eta_j$ mit $\xi_j, \eta_j \in \mathbb{R}$. Setze

$$Q_j(x) = (x - \mu_j)(x - \overline{\mu_j}) = x^2 - 2\xi_j x + (\xi_j^2 + \eta_j^2).$$

□

Beweis von Theorem 3.2. Wir führen Theorem 3.2 auf Theorem 3.1 zurück. Dazu *komplexifizieren* wir f. Es sei B' irgendeine ON-Basis von V und $A := M_{B'}^{B'}(f)$ die Darstellungsmatrix von f bezüglich B'. Dann ist A orthogonal und als reelle Matrix auch unitär. Also ist der Endomorphismus

$$A : \mathbb{C}^n \to \mathbb{C}^n, \quad z \mapsto Az,$$

unitär. Es sei

$$P_A(x) = P(x) = (x - \lambda_1) \cdots (x - \lambda_r)(x - \mu_1)(x - \overline{\mu_1}) \cdots (x - \mu_k)(x - \overline{\mu_k})$$

die Zerlegung des charakteristischen Polynoms von A, die nach Lemma 3.1 existiert. Nach I, Satz 21.1 (v) gilt $\lambda_i = \pm 1$, $i = 1, \ldots, r$, $\mu_j = \cos\alpha_j + i \sin\alpha_j$, $\alpha_j \in [0, 2\pi)$, $\alpha_j \neq 0, \pi$, $j = 1, \ldots, k$. Nach Theorem 3.1 erhalten wir für A eine ON-Basis $\widetilde{B}$ von $\mathbb{C}^n$ von Eigenvektoren von A. Es sei nun z ein Eigenvektor zu einem nicht reellen Eigenwert μ. Dann ist $\overline{z}$ ein Eigenvektor zum Eigenwert $\overline{\mu}$, denn

$$A\overline{z} = \overline{Az} = \overline{\mu z} = \overline{\mu}\,\overline{z}.$$

Deswegen können wir die Basis $\widetilde{B}$ so anordnen:

$v_1, \ldots, v_p$ die Eigenvektoren zum Eigenwert $+1$,
$w_1, \ldots, w_q$ die Eigenvektoren zum Eigenwert -1,
$z_1, \ldots, z_k$ die Eigenvektoren zu den Eigenwerten $\mu_1, \ldots, \mu_k$,
$\overline{z_1}, \ldots, \overline{z_k}$ die Eigenvektoren zu den Eigenwerten $\overline{\mu_1}, \ldots, \overline{\mu_k}$.

Da A reell ist, liegen die Eigenvektoren $v_1, \ldots, v_p, w_1, \ldots, w_q$ in $\mathbb{R}^n$.

Zu einem Paar $z, \overline{z}$ von Eigenvektoren zu $\mu, \overline{\mu}$ konstruieren wir nun einen unter A invarianten Unterraum $W \subset \mathbb{R}^n$. Dazu sei

$$z = x + iy, \quad x, y \in \mathbb{R}^n$$

und

$$W := \operatorname{Span}\{x, y\} \subset \mathbb{R}^n.$$

Behauptung $A(W) = W$

Beweis. Es gilt

$$\begin{aligned} 0 &= \langle z, \overline{z} \rangle = \langle x + iy, x - iy \rangle = \langle x, x \rangle - \langle y, y \rangle + 2i\langle x, y \rangle \\ 1 &= \langle z, z \rangle = \langle x + iy, x + iy \rangle = \langle x, x \rangle + \langle y, y \rangle. \end{aligned}$$

Daraus folgt $\langle x, x \rangle = \langle y, y \rangle = \frac{1}{2}$ und $\langle x, y \rangle = 0$. Aus $\mu = \cos\alpha + i \sin\alpha$, $x = \frac{1}{2}(z + \overline{z})$, $y = \frac{1}{2i}(z - \overline{z})$ folgt

$$\begin{aligned} Ax &= \frac{1}{2}(Az + A\overline{z}) = \frac{1}{2}(\mu z + \overline{\mu}\,\overline{z}) = \cos\alpha x - \sin\alpha y, \\ Ay &= \frac{1}{2i}(Az - A\overline{z}) = \frac{1}{2i}(\mu z - \overline{\mu}\,\overline{z}) = \sin\alpha x + \cos\alpha y. \end{aligned}$$

□

Nun setzen wir

$$x' := \sqrt{2}x, \quad y' := -\sqrt{2}y.$$

Bezüglich der ON-Basis $\{x', y'\}$ von W wird die Einschränkung von A auf W beschrieben durch die Matrix

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

Damit haben wir eine Orthonormalbasis

$$B'' := \{v_1, \ldots, v_p, w_1, \ldots, w_q, x'_1, y'_1, \ldots, x'_k, y'_k\}$$

von $\mathbb{R}^n$ gefunden, bezüglich der die Abbildung $A : \mathbb{R}^n \to \mathbb{R}^n$ die in Theorem 3.2 angegebene Gestalt hat. Die Transformationsmatrix, die die Standardbasis des $\mathbb{R}^n$ in die Basis B'' des $\mathbb{R}^n$ transformiert, transformiert dann die Basis B' von V in eine ON-Basis B von V mit den gewünschten Eigenschaften. □

4 Selbstadjungierte Endomorphismen

Es sei V ein euklidischer (unitärer) Vektorraum.

Definition Ein Endomorphismus $f : V \to V$ heißt *selbstadjungiert*, falls

$$\langle f(v), w\rangle = \langle v, f(w)\rangle \text{ für alle } v, w \in V.$$

Bemerkung 4.1 Ist $f : V \to V$ ein Endomorphismus und V endlich dimensional, so kann man zeigen, dass es genau eine lineare Abbildung $f^{\mathrm{adj}} : V \to V$ mit der Eigenschaft

$$\langle f(v), w\rangle = \langle v, f^{\mathrm{adj}}(w)\rangle \text{ für alle } v, w \in V$$

gibt. Die Abbildung f^{adj} heißt die zu f *adjungierte Abbildung*. Daher rührt die Bezeichnung selbstadjungiert.

Satz 4.1 *Es sei* $\dim V < \infty$*,* $f : V \to V$ *ein Endomorphismus und* B *eine ON-Basis von* V*. Dann ist* f *genau dann selbstadjungiert, wenn die Darstellungsmatrix* $M_B^B(f)$ *bezüglich der Basis* B *symmetrisch (hermitesch) ist.*

Beweis. Es seien $v, w \in V$ und x bzw. y die Koordinatenvektoren von v bzw. w bezüglich der Basis B. Es sei $A := M_B^B(f)$. Dann gilt:

$$\langle f(v), w\rangle = \langle v, f(w)\rangle \Leftrightarrow (Ax)^T\overline{y} = x^T\overline{(Ay)} \Leftrightarrow A^T = \overline{A}.$$

□

Lemma 4.1 *Ist* f *selbstadjungiert, so sind (auch im komplexen Fall) alle Eigenwerte reell. Insbesondere hat eine hermitesche Matrix nur reelle Eigenwerte.*

Beweis. Ist $f(v) = \lambda v$ mit $v \neq 0$, so gilt

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

Wegen $v \neq 0$ folgt daraus $\lambda = \overline{\lambda}$. □

Theorem 4.1 *Es sei* V *ein euklidischer (unitärer) Vektorraum und* $f : V \to V$ *ein selbstadjungierter Endomorphismus. Dann besitzt* V *eine ON-Basis, die aus Eigenvektoren von* f *besteht.*

Beweis. Ist $K = \mathbb{C}$, so zerfällt das charakteristische Polynom in Linearfaktoren

$$p_f(t) = \pm(t - \lambda_1) \cdots (t - \lambda_n)$$

und nach Lemma 4.1 sind die Eigenwerte $\lambda_1, \ldots, \lambda_n$ reell.

Auch im Fall $K = \mathbb{R}$ zerfällt das charakteristische Polynom in Linearfaktoren: Dazu betrachten wir die Matrixdarstellung $A = M_B^B(f)$ von f bezüglich irgendeiner ON-Basis B von V. Nach Satz 4.1 ist A symmetrisch. Wir können A auch als komplexe Matrix auffassen. Dann ist A hermitesch. Dann zerfällt das charakteristische Polynom $p_A(t) = p_f(t)$ über $\mathbb{C}$ in Linearfaktoren und nach Lemma 4.1 sind alle Eigenwerte reell.

Nun wenden wir Induktion nach $n = \dim V$ an: Es sei v_1 ein Eigenvektor von f zum Eigenwert λ_1. O. B. d. A. sei $\|v_1\| = 1$. Es sei

$$W := \{w \in V \mid \langle v_1, w\rangle = 0\}.$$

Behauptung (a) W ist ein Unterraum von V der Dimension $n-1$.

(b) Es gilt $f(W) \subset W$.

Beweis. (a) Dass W ein Unterraum von V ist, lässt sich leicht zeigen. Wegen $v_1 \notin W$ gilt $\dim W \leq n-1$. Nach I, Satz 20.2, lässt sich $\{v_1\}$ zu einer ON-Basis $\{v_1, v_2', \ldots, v_n'\}$ von V ergänzen. Da $v_i' \in W$ für $i = 2, \ldots, n$, folgt $\dim W \geq n-1$, also $\dim W = n-1$.

(b) Es gilt für $w \in W$:

$$\langle v_1, f(w)\rangle = \langle f(v_1), w\rangle = \langle \lambda_1 v_1, w\rangle = \lambda_1\langle v_1, w\rangle = 0.$$

Daraus folgt $f(W) \subset W$. □

Aufgrund der Behauptung definiert $f|_W$ einen Endomorphismus $f|_W : W \to W$. Da $f|_W$ die Einschränkung eines selbstadjungierten Endomorphismus ist, ist auch $f|_W$ wieder selbstadjungiert. Da W ein Unterraum von V der Dimension $n-1$ ist, können wir die Induktionsannahme auf den Endomorphismus $f|_W : W \to W$ anwenden. Nach Induktionsannahme gibt es eine ON-Basis $\{v_2, \ldots, v_n\}$ von Eigenvektoren für $f|_W : W \to W$. Dann ist $B = \{v_1, \ldots, v_n\}$ eine ON-Basis von Eigenvektoren für f. □

Korollar 4.1 *Ist $A \in \mathrm{Mat}(n,n;K)$ eine symmetrische bzw. hermitesche Matrix, so gibt es eine orthogonale bzw. unitäre Matrix S, so dass*

$$\overline{S}^T A S = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

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

Wir halten noch das folgende Korollar fest.

Korollar 4.2 *Es sei V ein euklidischer (unitärer) Vektorraum, $f : V \to V$ ein selbstadjungierter Endomorphismus und $\lambda_1, \ldots, \lambda_k$ die paarweise verschiedenen Eigenwerte. Dann ist*

$$V = \mathrm{Eig}(f, \lambda_1) \perp \ldots \perp \mathrm{Eig}(f, \lambda_k).$$

Beweis. Aus dem Theorem folgt, dass V die direkte Summe der Eigenräume ist. Es bleibt zu zeigen: $\mathrm{Eig}(f, \lambda_i) \perp \mathrm{Eig}(f, \lambda_j)$ für alle $i \neq j$, $i, j = 1, \ldots, k$.

Dazu sei $v \in \mathrm{Eig}(f, \lambda_i)$, $w \in \mathrm{Eig}(f, \lambda_j)$, $i \neq j$. Dann gilt

$$\lambda_i \langle v, w \rangle = \langle \lambda_i v, w \rangle = \langle f(v), w \rangle = \langle v, f(w) \rangle = \langle v, \lambda_j w \rangle = \lambda_j \langle v, w \rangle.$$

Daraus folgt

$$(\lambda_i - \lambda_j) \langle v, w \rangle = 0,$$

also $\langle v, w \rangle = 0$. □

5 Hauptachsentransformation

Es sei

$$\begin{array}{rccc} s: & \mathbb{R}^n \times \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ & (v, w) & \longmapsto & s(v, w) \end{array}$$

eine symmetrische Bilinearform auf dem $\mathbb{R}^n$. Nach I, Lemma 20.2, wird sie durch eine symmetrische Matrix $A = (s(e_i, e_j))$ gegeben. Für $x, y \in \mathbb{R}^n$ gilt

$$s(x, y) = x^T A y = \langle Ax, y \rangle = \langle x, Ay \rangle,$$

wobei $\langle\ ,\ \rangle$ das euklidische Skalarprodukt des $\mathbb{R}^n$ bezeichnet. Die Matrix A definiert daher einen selbstadjungierten Endomorphismus

$$\begin{array}{rccc} f: & \mathbb{R}^n & \longrightarrow & \mathbb{R}^n \\ & x & \longmapsto & Ax \end{array}.$$

Aus Theorem 4.1 und Korollar 4.1 folgt daher:

Satz 5.1 (Hauptachsentransformation) *Es sei $A \in \mathrm{Mat}(n, n; \mathbb{R})$ eine symmetrische Matrix und s die durch A gegebene Bilinearform auf dem $\mathbb{R}^n$.*

(i) *Die Matrix A hat nur reelle Eigenwerte $\lambda_1, \ldots, \lambda_n$ und sie ist diagonalisierbar. Genauer gilt: Es gibt eine orthogonale Matrix S, so dass*

$$S^T A S = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

(ii) *Es gibt eine Orthonormalbasis $B = (w_1, \dots, w_n)$ des $\mathbb{R}^n$, so dass gilt*

$$(s(w_i, w_j)) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Beispiel 5.1 Es sei $n = 2$ und

$$A = \begin{pmatrix} a & b \\ b & a \end{pmatrix}.$$

Für $v = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ gilt

$$s(v, v) = v^T A v = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b \\ b & a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = ax^2 + 2bxy + ay^2.$$

Im $\mathbb{R}^2$ mit den Koordinaten x, y betrachten wir die Kurve

$$\begin{aligned} C &= \{v \in \mathbb{R}^2 \,|\, s(v, v) = 1\} \\ &= \{(x, y) \in \mathbb{R}^2 \,|\, ax^2 + 2bxy + ay^2 = 1\}. \end{aligned}$$

Wir untersuchen die Eigenwerte und Eigenvektoren der Matrix A. Das charakteristische Polynom lautet

$$P_A(\lambda) = \lambda^2 - 2a\lambda + a^2 - b^2 = (\lambda - a + b)(\lambda - a - b).$$

Die Eigenwerte sind also $\lambda_1 = a + b$ und $\lambda_2 = a - b$. Als Eigenvektoren berechnet man

$$v_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Sind x', y' die Koordinaten des Vektors v bezüglich der Basis $B = \{v_1, v_2\}$, so gilt

$$\begin{aligned} \begin{pmatrix} x \\ y \end{pmatrix} &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}, \\ s(v, v) &= (a + b)x'^2 + (a - b)y'^2. \end{aligned}$$

In den neuen Koordinaten x', y' ist also der gemischte Term mit $x'y'$ verschwunden. Wir nehmen an, dass $\lambda_1 > 0$ und $\lambda_2 \neq 0$. Dann setzen wir

$$\alpha = \frac{1}{\sqrt{\lambda_1}} \text{ und } \beta = \frac{1}{\sqrt{|\lambda_2|}}.$$

Damit lautet die Gleichung von C in den Koordinaten x', y':

$$\frac{x'^2}{\alpha^2} \pm \frac{y'^2}{\beta^2} = 1.$$

Im Fall $+$ beschreibt diese Gleichung eine Ellipse, α und β sind gerade die *Hauptachsen* (vgl. Abbildung 1).

Im Fall $-$ können wir für die linke Seite der Gleichung auch schreiben

$$\frac{x'^2}{\alpha^2} - \frac{y'^2}{\beta^2} = \left(\frac{x'}{\alpha} + \frac{y'}{\beta}\right)\left(\frac{x'}{\alpha} - \frac{y'}{\beta}\right).$$

Mit der Koordinatentransformation

$$\begin{aligned} \tilde{x} &= \frac{x'}{\alpha} + \frac{y'}{\beta}, \\ \tilde{y} &= \frac{x'}{\alpha} - \frac{y'}{\beta} \end{aligned}$$

schreibt sich die Gleichung von C als

$$\tilde{x}\tilde{y} = 1.$$

Diese Gleichung beschreibt eine Hyperbel, α und β sind wieder die *Hauptachsen* (vgl. Abbildung 2). Deswegen nennt man die obige Transformation auch *Hauptachsentransformation*.

Die Matrix

$$S = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

ist eine orthogonale Matrix. Wegen $\cos 45° = \frac{1}{\sqrt{2}}$ beschreibt S eine Drehung um 45°. Das bedeutet, dass die durch

$$ax^2 + 2bxy + ay^2 = 1$$

gegebene Kurve nach einer Drehung des Koordinatensystems um 45° in Hauptachsenform vorliegt.

Beispiel 5.2 Wir betrachten die Matrix

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

Das charakteristische Polynom dieser Matrix lautet

$$\begin{aligned} P_A(\lambda) &= \begin{vmatrix} 1-\lambda & 1 & -2\sqrt{2} \\ 1 & 1-\lambda & 2\sqrt{2} \\ -2\sqrt{2} & 2\sqrt{2} & -\lambda \end{vmatrix} \\ &= -\lambda^3 + 2\lambda^2 + 16\lambda - 32 = (-\lambda+2)(\lambda+4)(\lambda-4). \end{aligned}$$

Die Eigenwerte von A sind also $\lambda_1 = 2$, $\lambda_2 = 4$ und $\lambda_3 = -4$. Eine Lösung von

$$\begin{pmatrix} -1 & 1 & -2\sqrt{2} \\ 1 & -1 & 2\sqrt{2} \\ -2\sqrt{2} & 2\sqrt{2} & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

ist der Vektor

$$v_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},$$

v_1 ist also ein Eigenvektor zum Eigenwert $\lambda_1 = 2$. Eine Lösung des Gleichungssystems

$$\begin{pmatrix} -3 & 1 & -2\sqrt{2} \\ 1 & -3 & 2\sqrt{2} \\ -2\sqrt{2} & 2\sqrt{2} & -4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

ist der Vektor

$$v_2 = \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ -\sqrt{2} \end{pmatrix},$$

v_2 ist also ein Eigenvektor zum Eigenwert $\lambda_2 = 4$. Entsprechend sieht man, dass der Vektor

$$v_3 = \frac{1}{2} \begin{pmatrix} -1 \\ 1 \\ -\sqrt{2} \end{pmatrix}$$

ein Eigenvektor zum Eigenwert $\lambda_3 = -4$ ist. Also diagonalisiert die Matrix

$$S = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{\sqrt{2}} & -\frac{1}{2} & \frac{1}{2} \\ 0 & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

die Matrix A, d.h. es gilt

$$S^T A S = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -4 \end{pmatrix}.$$

Sind x', y', z' die Koordinaten von v bezüglich der Basis $B = \{v_1, v_2, v_3\}$, so lautet die Gleichung $s(v, v) = 1$ in diesen Koordinaten

$$2x'^2 + 4y'^2 - 4z'^2 = 1.$$

Diese Gleichung beschreibt ein einschaliges Hyperboloid (vgl. Abbildung 6).

6 Symmetrische Bilinearformen

Es sei nun V ein endlich dimensionaler reeller Vektorraum mit einer symmetrischen Bilinearform

$$\langle\ ,\ \rangle : V \times V \to \mathbb{R}.$$

(Diese Bilinearform braucht nicht positiv definit zu sein.) Wie wir bereits gesehen haben, entspricht dieser Bilinearform eine quadratische Form

$$q : V \to \mathbb{R}, \quad q(v) = \langle v, v \rangle.$$

Bezüglich einer Basis $B = \{v_1, \ldots, v_n\}$ von V wird $\langle\ ,\ \rangle$ dargestellt durch die Matrix

$$A = (a_{ij}), \quad a_{ij} = \langle v_i, v_j \rangle.$$

Nach der Transformationsformel ändert sich bei einem Basiswechsel mit Transformationsmatrix $S \in \mathrm{GL}(n; \mathbb{R})$ die Darstellungsmatrix wie folgt:

$$A \mapsto S^T A S.$$

Nach dem Satz über die Hauptachsentransformation gibt es eine orthogonale Matrix S mit

$$S^T A S = S^{-1} A S = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}, \quad (\lambda_1, \ldots, \lambda_n \in \mathbb{R}).$$

Wir fragen nun nach einer *Normalform*, wenn wir allgemeiner $S \in \mathrm{GL}(n; \mathbb{R})$ zulassen.

Satz 6.1 *Es sei $\langle\ ,\ \rangle : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform. Dann gibt es eine Basis B von V, bezüglich der $\langle\ ,\ \rangle$ dargestellt wird durch*

$$\begin{pmatrix} E_k & & 0 \\ & -E_l & \\ 0 & & 0 \end{pmatrix}.$$

Beweis. Nach dem Satz über die Hauptachsentransformation gibt es eine Basis $B' = \{w_1, \ldots, w_n\}$, bezüglich der die darstellende Matrix wie folgt aussieht:

$$\begin{pmatrix} \lambda_1 & & & & & & & & \\ & \ddots & & & & & 0 & & \\ & & \lambda_k & & & & & & \\ & & & \lambda_{k+1} & & & & & \\ & & & & \ddots & & & & \\ & & & & & \lambda_{k+l} & & & \\ & & & & & & 0 & & \\ & 0 & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix}, \quad \lambda_i \begin{cases} > 0 & \text{für } i \leq k, \\ < 0 & \text{für } k < i \leq k+l. \end{cases}$$

Wir setzen

$$v_i := \begin{cases} \frac{1}{\sqrt{|\lambda_i|}} w_i & \text{für } 1 \leq i \leq k+l, \\ w_i & \text{sonst.} \end{cases}$$

Dann gilt:

$$\langle v_i, v_i \rangle = \begin{cases} +1 & \text{für } 1 \leq i \leq k, \\ -1 & \text{für } k < i \leq k+l, \\ 0 & \text{sonst.} \end{cases}$$

Also hat $\langle\ ,\ \rangle$ bezüglich der Basis $B = \{v_1, \ldots, v_n\}$ die gewünschte Gestalt.
□

Definition Es sei $\langle\ ,\ \rangle : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform. Die Menge

$$V_0 := \{v \in V \mid \langle v, w \rangle = 0 \text{ für alle } w \in V\}$$

heißt das *Radikal* (oder der *Nullraum*) von $\langle\ ,\ \rangle$. Es ist ein Unterraum von V.

Korollar 6.1 *Es sei $\langle\ ,\ \rangle : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform. Dann gibt es eine Zerlegung*

$$V = V_+ \oplus V_- \oplus V_0,$$

in Unterräume, so dass gilt:

(i) *Die Zerlegung ist orthogonal bezüglich $\langle\ ,\ \rangle$.*

(ii) $\langle v, v \rangle > 0$ *für* $0 \neq v \in V_+$, $\langle v, v \rangle < 0$ *für* $0 \neq v \in V_-$.

Beweis. Es sei $B = \{v_1, \dots, v_n\}$ eine Basis wie in Satz 6.1. Setze

$$V_+ := \mathrm{Span}\{v_1, \dots, v_k\}, \quad V_- := \mathrm{Span}\{v_{k+1}, \dots, v_{k+l}\}.$$

Dann bleibt zu zeigen, dass $V_0 = \mathrm{Span}\{v_{k+l+1}, \dots, v_n\}$. Die Inklusion

$$\mathrm{Span}\{v_{k+l+1}, \dots, v_n\} \subset V_0$$

ist klar. Es sei umgekehrt $v \in V_0$. Dann gilt

$$v = \mu_1 v_1 + \cdots + \mu_{k+l} v_{k+l} + \mu_{k+l+1} v_{k+l+1} + \cdots + \mu_n v_n.$$

Für $i \in \{1, \dots, k+l\}$ gilt aber $\langle v, v_i \rangle = \pm\mu_i = 0$. Daraus folgt $\mu_i = 0$. □

Wie der Beweis zeigt, hängt die Zerlegung von einer Basis B von V ab. Ist A die Darstellungsmatrix von $\langle\ ,\ \rangle$ bezüglich dieser Basis, so gilt

$$\begin{aligned} \dim V_+ &= \text{Anzahl der positiven Eigenwerte von } A, \\ \dim V_- &= \text{Anzahl der negativen Eigenwerte von } A. \end{aligned}$$

Der Trägheitssatz von Sylvester besagt, dass diese Zahlen tatsächlich unabhängig von der Wahl von B sind.

Satz 6.2 (Trägheitssatz von Sylvester) *Es sei $\langle\ ,\ \rangle : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform, B eine Basis von V und A die Darstellungsmatrix von $\langle\ ,\ \rangle$ bezüglich B. Dann sind die Zahlen*

$$\begin{aligned} k &:= \textit{Anzahl der positiven Eigenwerte von } A, \\ l &:= \textit{Anzahl der negativen Eigenwerte von } A \end{aligned}$$

unabhängig von der Auswahl von B.

Beweis. Es sei B' eine andere Basis, k', l' die entsprechenden Anzahlen und

$$V = V_+ \oplus V_- \oplus V_0 = V'_+ \oplus V'_- \oplus V_0$$

die entsprechenden zugehörigen Zerlegungen nach Korollar 6.1. Da die Anzahl der von Null verschiedenen Eigenwerte gleich $\dim V - \dim V_0$ ist und damit nicht von der Auswahl der Basis abhängt, gilt $k + l = k' + l'$. Daher reicht es, $l = l'$ zu zeigen.

Angenommen, es gibt

$$0 \neq v \in V_+ \cap (V'_- \oplus V_0).$$

Dann gilt $\langle v, v \rangle > 0$ und $v = v'_- + v_0$ mit $v'_- \neq 0$. Dann folgt aber

$$\langle v, v \rangle = \langle v'_-, v'_- \rangle + \langle v_0, v_0 \rangle = \langle v'_-, v'_- \rangle < 0,$$

ein Widerspruch. Also gilt $V_+ \cap (V'_- \oplus V_0) = \{0\}$ und aus Satz 1.2 folgt

$$k + l' + \dim V_0 \leq \dim V, \text{ also } k + l' \leq k + l, \text{ d.h. } l' \leq l.$$

Durch Vertauschen der Rollen von l und l' folgt $l \leq l'$, also $l = l'$ und $k = k'$.
□

Definition Die Zahl k nennt man auch den *Index*, die Zahl $k-l$ die *Signatur* der symmetrischen Bilinearform $\langle\ ,\ \rangle$.

Korollar 6.2 *Eine symmetrische Bilinearform ist genau dann positiv definit, wenn alle Eigenwerte einer Darstellungsmatrix positiv sind.*

Definition Eine symmetrische Matrix $A \in \mathrm{Mat}(n, n; \mathbb{R})$ heißt *positiv definit*, in Zeichen $A > 0$, falls die zugehörige Form $\langle\ ,\ \rangle_A$ positiv definit ist, d.h.

$$x^T A x > 0 \text{ für alle } x \in \mathbb{R}^n, x \neq 0.$$

Korollar 6.3 *Eine symmetrische Matrix $A \in \mathrm{Mat}(n, n; \mathbb{R})$ ist genau dann positiv definit, wenn alle Eigenwerte positiv sind.*

Wir wollen zum Abschluss noch ein anderes Kriterium dafür angeben, dass eine symmetrische Matrix $A = (a_{ij})$ positiv definit ist. Dazu bezeichnen wir mit

$$A_k := \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}$$

die linke obere $k \times k$-Teilmatrix von A. Die Determinante $\det A_k$ bezeichnet man auch als *Hauptminor* von A.

Satz 6.3 (Hurwitz-Kriterium) *Es sei $A \in \mathrm{Mat}(n, n; \mathbb{R})$ eine symmetrische Matrix. Dann gilt:*

$$A \textit{ positiv definit } \Leftrightarrow \det A_k > 0 \textit{ für } 1 \leq k \leq n.$$

Beweis. „$\Rightarrow$": Wir zeigen zunächst

$$\det A = \det A_n > 0.$$

Da A positiv definit ist, gibt es ein $S \in \mathrm{GL}(n;\mathbb{R})$ mit

$$S^T A S = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}.$$

Also folgt

$$1 = \det(S^T A S) = \det A (\det S)^2, \text{ also } \det A > 0.$$

Um nun $\det A_k > 0$ für $1 \leq k < n$ zu zeigen, betrachten wir

$$U_k := \{x \in \mathbb{R}^n \mid x_{k+1} = \ldots = x_n = 0\} \subset \mathbb{R}^n.$$

Die Form $\langle\ ,\ \rangle_A$ definiert durch Einschränkung eine Form $\langle\ ,\ \rangle_k : U_k \times U_k \to \mathbb{R}$ mit Darstellungsmatrix A_k. Da auch $\langle\ ,\ \rangle_k$ positiv definit ist, folgt $\det A_k > 0$.

"$\Leftarrow$": Wir führen Induktion über n durch. Der Induktionsanfang $n = 1$ ist klar. Nach Induktionsvoraussetzung ist A_{n-1} positiv definit. Also gibt es ein $S' \in \mathrm{GL}(n-1;\mathbb{R})$ mit

$$(S')^T A_{n-1} S' = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} = E_{n-1}.$$

Es sei

$$S := \left(\begin{array}{ccc|c} & & & 0 \\ & S' & & \vdots \\ & & & 0 \\ \hline 0 & \cdots & 0 & 1 \end{array}\right) \in \mathrm{GL}(n;\mathbb{R}).$$

Es gilt

$$S^T A S = \left(\begin{array}{ccc|c} 1 & & & b_1 \\ & \ddots & & \vdots \\ & & 1 & b_{n-1} \\ \hline b_1 & \cdots & b_{n-1} & b_n \end{array}\right) =: B$$

und

$$\det B = (\det S)^2 \det A > 0.$$

Es genügt zu zeigen, dass B positiv definit ist. Dazu setze

$$T := \left(\begin{array}{ccc|c} 1 & & & -b_1 \\ & \ddots & & \vdots \\ & & 1 & -b_{n-1} \\ \hline 0 & \cdots & 0 & 1 \end{array}\right) \in \mathrm{GL}(n;\mathbb{R}).$$

Dann ist

$$T^T BT = \left(\begin{array}{ccc|c} 1 & & & 0 \\ & \ddots & & \vdots \\ & & 1 & 0 \\ \hline 0 & \cdots & 0 & c_n \end{array} \right) =: C.$$

Nun ist

$$\det C = (\det T)^2 \det B = \det B > 0$$

und damit $c_n > 0$. Also ist C positiv definit und damit auch B und A. □

7 Das Minimalpolynom

Wir kommen nun auf das schon in LA I betrachtete Problem zurück, für die Darstellungsmatrix eines Endomorphismus $f : V \to V$ eines K-Vektorraums eine Normalform zu finden. Dabei spielt neben dem charakteristischen Polynom ein anderes Polynom eine Rolle, das wir nun einführen wollen. Dazu machen wir zunächst einen Exkurs über den Polynomring.

Es sei K ein beliebiger Körper. Dann betrachten wir den Polynomring in einer Variablen über K:

$$K[x] := \{a_0 + a_1 x + \cdots + a_n x^n \mid a_i \in K\}.$$

Auf $K[x]$ ist eine Addition und eine Multiplikation erklärt. Es sei $P(x), Q(x) \in K[x]$,

$$P(x) = a_0 + a_1 x + \cdots + a_n x^n, \quad Q(x) = b_0 + b_1 x + \cdots + b_m x^m.$$

O. B. d. A. sei $n \leq m$. Zur Definition der *Addition* setzen wir $a_{n+1} = \ldots a_m = 0$. Dann definieren wir

$$P(x) + Q(x) := (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_m + b_m)x^m.$$

Die *Multiplikation* ist dadurch erklärt, dass man formal ausmultipliziert:

$$\begin{aligned} P(x) \cdot Q(x) &= (a_0 + a_1 x + \cdots + a_n x^n) \cdot (b_0 + b_1 x + \cdots + b_m x^m) \\ &:= a_0 b_0 + (a_0 b_1 + a_1 b_0)x + \cdots + a_n b_m x^{n+m} \\ &= c_0 + c_1 x + \cdots + c_{n+m} x^{n+m} \end{aligned}$$

mit

$$c_k := \sum_{i=0}^{k} a_i b_{k-i}.$$

Statt $P(x)$ und $Q(x)$ schreiben wir von nun an auch P und Q.

Satz 7.1 *Mit dieser Addition und Multiplikation wird $K[x]$ zu einem kommutativen Ring mit Einselement.*

Beweis. Die Axiome sind leicht nachzuprüfen. Was ist das Einselement? □

Das Nullpolynom (alle Koeffizienten $a_i = 0$) bezeichnen wir mit 0.

Definition Der *Grad* eines Polynoms

$$P(x) = a_0 + a_1 x + \cdots + a_n x^n \text{ mit } a_n \neq 0$$

ist die Zahl n ($n = \deg P$). Den Grad des Nullpolynoms definieren wir als $\deg(0) := -\infty$. Das Polynom P heißt *normiert*, falls $a_n = 1$ ist.

Satz 7.2 (Gradformel) *Für $P, Q \in K[x]$ gilt:*

$$\deg(P \cdot Q) = \deg P + \deg Q.$$

Dabei soll formal $n - \infty = m - \infty = -\infty - \infty = -\infty$ gelten.

Beweis. Dies folgt aus $c_{n+m} = a_n b_m \neq 0$ falls $a_n \neq 0$ und $b_m \neq 0$. □

Satz 7.3 (Division mit Rest) *Es seien $P, Q \in K[x]$ mit $P, Q \neq 0$. Dann gibt es eindeutig bestimmte Polynome q, r mit*

(i) $P = Qq + r$,

(ii) $\deg r < \deg Q$.

Beweis. Wir zeigen zunächst die Eindeutigkeit. Es seien $q, q', r, r' \in K[x]$ mit

$$P = Qq + r = Qq' + r', \quad \deg r, \deg r' < \deg Q.$$

Dann folgt

$$Q(q - q') = r' - r.$$

1.Fall: $q = q' \Rightarrow r = r'$.
2.Fall: $q \neq q'$. Dann ist

$$\deg(r' - r) = \deg Q + \deg(q - q') \geq \deg Q$$

im Widerspruch zu $\deg(r' - r) \leq \max\{\deg r, \deg r'\} < \deg Q$.

Nun beweisen wir die Existenz von q und r. Wenn es ein $q \in K[x]$ gibt mit

$$P = Qq,$$

so können wir $r = 0$ setzen und die Behauptung ist bewiesen. Andernfalls betrachten wir die Menge

$$M := \{\deg(P - Qp) \mid p \in K[x]\} \subset \mathbb{N} = \{0, 1, 2, \ldots\}.$$

Diese Menge besitzt ein Minimum in $\mathbb{N}$. Es sei $q \in K[x]$ mit

$$\deg(P - Qq) \leq \deg(P - Qp) \text{ für alle } p \in K[x].$$

Es sei ferner

$$r := P - Qq,$$

d.h.

$$P = Qq + r.$$

Es bleibt zu zeigen: $\deg r < \deg Q$. Angenommen, $\deg r \geq \deg Q$. Es sei

$$\begin{aligned} Q &= b_0 + b_1 x + \cdots + b_m x^m \quad (b_m \neq 0), \\ r &= c_0 + c_1 x + \cdots + c_k x^k \quad (c_k \neq 0). \end{aligned}$$

Dann ist nach Annahme $k \geq m$. Es sei

$$p := q + \frac{c_k}{b_m} x^{k-m} \in K[x].$$

Dann ist

$$P - Qp = P - Qq - Q\frac{c_k}{b_m} x^{k-m} = r - Q\frac{c_k}{b_m} x^{k-m}.$$

Es ist also

$$P - Qp = c_k x^k - b_m \frac{c_k}{b_m} x^k + \text{ Terme der Ordnung } < k.$$

Daher folgt

$$\deg(P - Qp) < k = \deg r = \deg(P - Qq).$$

im Widerspruch zur Wahl von q. □

Dies führt zu dem schon aus der Schule bekannten Verfahren zur Polynomdivision (siehe Vorlesung und Übungen).

In Polynome können wir nun Körperelemente einsetzen. Ist

$$P(x) = a_0 + a_1 x + \cdots + a_n x^n \in K[x]$$

und $\lambda \in K$, so setzen wir

$$P(\lambda) := a_0 + a_1 \lambda + \cdots + a_n \lambda^n \in K.$$

Auf diese Weise definiert $P(x)$ eine Funktion

$$\begin{array}{rccc} \widetilde{P}: & K & \longrightarrow & K \\ & \lambda & \longmapsto & P(\lambda) \end{array}.$$

Damit erhalten wir eine Abbildung

$$\begin{array}{rccc} \widetilde{\ }: & K[x] & \longrightarrow & \mathrm{Abb}(K,K) \\ & P & \longmapsto & \widetilde{P} \end{array}.$$

Warnung Wir müssen zwischen Polynomen P und Polynomfunktionen $\widetilde{P}$ unterscheiden, denn bei endlichen Körpern können Unterschiede auftreten: Ist z.B. $K = \mathbb{F}_2$ und $P(x) = x^2 + x$, so ist $\widetilde{P}$ die Nullfunktion, da $P(0) = P(1) = 1 + 1 = 0$. Also ist in diesem Fall die Abbildung $\widetilde{\ }$ nicht injektiv.

Definition Ist $0 \neq P \in K[x]$ und $\lambda \in K$, so dass $P(\lambda) = 0$ gilt, so heißt λ eine *Nullstelle* von P.

Korollar 7.1 *Es sei* $0 \neq P \in K[x]$ *und* λ *eine Nullstelle von* P*. Dann gibt es genau ein Polynom* $Q \in K[x]$ *mit*

$$P = Q(x - \lambda).$$

Es ist $\deg Q = \deg P - 1$*.*

Beweis. Wir dividieren P durch $(x - \lambda)$ mit Rest: Nach Satz 7.3 gibt es eindeutig bestimmte $Q, r \in K[x]$ mit

$$P = (x - \lambda)Q + r, \quad \deg r < 1 = \deg(x - \lambda).$$

Also ist $r(x) = a_0 \in K$. Setzen wir λ in diese Gleichung ein, so folgt

$$0 = P(\lambda) = (\lambda - \lambda)Q(\lambda) + a_0 = a_0,$$

also $r = 0$. □

Korollar 7.2 *Es sei* $0 \neq P \in K[x]$*. Dann ist die Anzahl der Nullstellen von* P *höchstens gleich dem Grad von* P*.*

Beweis. Wir führen Induktion über den Grad $n := \deg P$. Für $n = 0$ ist P ein konstantes Polynom $P(x) = a_0 \neq 0$ und das hat gar keine Nullstelle. Damit ist die Behauptung für $n = 0$ bewiesen.

Nun sei $\deg P = n \geq 1$ und die Behauptung sei schon für alle Polynome $Q \in K[x]$ mit $\deg Q \leq n - 1$ bewiesen. Hat P keine Nullstelle, so ist die Behauptung richtig. Andernfalls sei $\lambda \in K$ eine Nullstelle von P. Nach Korollar 7.1 gibt es dann ein $Q \in K[x]$ mit

$$P = (x - \lambda)Q \text{ und } \deg Q = n - 1.$$

Alle von λ verschiedenen Nullstellen von P müssen auch Nullstellen von Q sein. Nach Induktionsannahme hat Q höchstens $n-1$ verschiedene Nullstellen, also P höchstens n verschiedene Nullstellen. $\square$

Korollar 7.3 *Hat K unendlich viele Elemente, so ist die Abbildung*

$$\widetilde{\ }: K[x] \to \mathrm{Abb}(K, K), \quad P \mapsto \widetilde{P},$$

injektiv.

Beweis. Es seien $P_1, P_2 \in K[x]$ mit $\widetilde{P_1} = \widetilde{P_2}$. Betrachte $Q := P_1 - P_2$. Dann ist $\widetilde{Q} = 0$, also hat Q unendlich viele Nullstellen. Aus Korollar 7.2 folgt damit $Q = 0$, also $P_1 = P_2$. $\square$

Satz 7.4 *Es sei K ein unendlicher Körper. Jedes Polynom $0 \neq P \in K[x]$ besitzt eine Darstellung*

$$P = (x - \lambda_1)^{\nu_1} \cdots (x - \lambda_r)^{\nu_r} \cdot Q,$$

wobei $\lambda_1, \ldots, \lambda_r$ paarweise verschieden sind und Q ein Polynom ohne Nullstellen ist. Diese Darstellung ist bis auf die Reihenfolge der Faktoren eindeutig.

Definition Man nennt ν_i die *Ordnung* oder *Vielfachheit* der Nullstelle λ_i.

Für den Beweis des Satzes brauchen wir ein Lemma.

Lemma 7.1 *Es sei K ein unendlicher Körper und $P, Q \in K[x]$ Polynome mit $P(\lambda) \neq 0$, $Q(\lambda) \neq 0$. Gilt*

$$(x - \lambda)^{\nu} P(x) = (x - \lambda)^{\mu} Q(x)$$

für alle $x \in K$, so ist $\nu = \mu$.

Beweis. O. B. d. A. sei $\nu \geq \mu$. Dann gilt

$$(x - \lambda)^{\nu - \mu} P(x) - Q(x) = 0 \quad (\text{für } x \neq \lambda).$$

Da K unendlich viele Elemente enthält, gilt für die Polynome

$$(x - \lambda)^{\nu - \mu} P - Q = 0.$$

Falls $\nu > \mu$, wäre $Q(\lambda) = 0$, ein Widerspruch. □

Beweis von Satz 7.4. Die Existenz der Darstellung folgt aus Korollar 7.1.

Zum Beweis der Eindeutigkeit: Die λ_i sind genau die Nullstellen von P, liegen also eindeutig fest. Die Eindeutigkeit der ν_i folgt aus Lemma 7.1. Es bleibt die Eindeutigkeit von Q zu zeigen. Dazu sei

$$(x-\lambda_1)^{\nu_1}\cdots(x-\lambda_r)^{\nu_r}Q = (x-\lambda_1)^{\nu_1}\cdots(x-\lambda_r)^{\nu_r}Q'.$$

Dann gilt für alle $x \neq \lambda_1, \ldots, \lambda_r$

$$Q(x) = Q'(x).$$

Also hat $Q - Q'$ unendlich viele Nullstellen und es folgt $Q = Q'$. □

Definition Man sagt, das Polynom $0 \neq P \in K[x]$ *zerfällt* über K, falls es eine Darstellung

$$P(x) = a(x-\lambda_1)^{\nu_1}\cdots(x-\lambda_r)^{\nu_r} \quad (a \in K)$$

gibt.

Beispiel 7.1 Das Polynom $P(x) = 1 + x^2$ zerfällt nicht über $\mathbb{R}$, aber über $\mathbb{C}$:

$$P(x) = (x-i)(x+i).$$

Satz 7.5 (Fundamentalsatz der Algebra) *Jedes nicht konstante Polynom $P \in \mathbb{C}[x]$ besitzt eine Nullstelle.*

Diesen Satz hat erstmals C. F. Gauß 1799 bewiesen. Heutzutage wird er meist in der Vorlesung Funktionentheorie bewiesen, da man mit Methoden dieser Vorlesung einen sehr knappen und eleganten Beweis geben kann.

Korollar 7.4 *Jedes Polynom $P \in \mathbb{C}[x]$ zerfällt.*

Definition Eine Teilmenge I eines Ringes R heißt ein *Ideal*, falls gilt:

(I1) $I \subset R$ ist eine Untergruppe bezüglich der Addition.

(I2) Ist $r \in R$ und $s \in I$, so ist auch $r \cdot s \in I$.

Beispiel 7.2 $I := \langle P_1, \ldots, P_n \rangle := \{Q_1P_1 + \cdots + Q_nP_n \mid Q_i \in R\}$.

Definition $I = \langle P_1, \ldots, P_n \rangle$ heißt das von $P_1, \ldots, P_n$ *erzeugte* Ideal.

Definition Ein Ideal I heißt *Hauptideal*, falls es von einem Element erzeugt, wird, d.h. $I = \langle P \rangle = \{QP \,|\, Q \in R\}$ für ein $P \in R$.

Lemma 7.2 *Es seien R, R' Ringe und $f : R \to R'$ ein Ringhomomorphismus. Dann ist der Kern von f ein Ideal in R.*

Beweis. Nach I, Satz 8.4, ist $\operatorname{Ker} f$ eine Untergruppe von R. Es sei $r \in R$ und $s \in \operatorname{Ker} f$. Dann gilt

$$f(r \cdot s) = f(r) \cdot f(s) = f(r) \cdot 0 = 0,$$

also $r \cdot s \in \operatorname{Ker} f$. □

Satz 7.6 (i) *$K[x]$ ist ein* Hauptidealring, *d.h. jedes Ideal von $K[x]$ ist ein Hauptideal.*

(ii) *Zu jedem Ideal $I \neq \{0\}$ in $K[x]$ gibt es genau ein normiertes Polynom P mit $I = \langle P \rangle$.*

Beweis. (i): Im Fall $I = \{0\}$ ist $I = \langle 0 \rangle$. Es sei also $I \neq \{0\}$. Dann hat die Menge

$$M := \{\deg P \,|\, 0 \neq P \in I\} \subset \mathbb{N}$$

ein Minimum $m := \min M$. Es sei $P \in I$ mit $\deg P = m$.

Behauptung $I = \langle P \rangle = K[x] \cdot P$.

Beweis. Es gilt $\langle P \rangle \subset I$ nach Definition eines Ideals.

Es bleibt zu zeigen: $I \subset \langle P \rangle$. Dazu sei $Q \in I$ beliebig. Nach Satz 7.3 gibt es eine Darstellung

$$Q = qP + r, \quad \deg r < \deg P.$$

Ist $r = 0$, so ist $Q \in \langle P \rangle$. Andernfalls folgt mit (I1) und (I2)

$$r = Q - qP \in I.$$

Wegen $0 \leq \deg r < \deg P = m$ ist dies ein Widerspruch zur Wahl von P. □

(ii): Mit P liegt auch aP in I für $a \in K$. Also kann man P als normiert annehmen. Es sei

$$I = \langle P \rangle = \langle P' \rangle \text{ mit } P, P' \text{ normiert.}$$

Dann gilt

$$P' = Q'P, \quad P = QP' \quad \text{für geeignete } Q, Q' \in K[x].$$

Es folgt

$$P = QQ'P \Leftrightarrow P(1 - QQ') = 0.$$

Nach der Gradformel folgt daraus

$$\deg QQ' = 0, \text{ also } QQ' = 1,$$

und damit sind Q und Q' ebenfalls nach der Gradformel konstant. Da P, P' normiert sind, folgt $Q = Q' = 1$, also $P = P'$. □

Nun wollen wir die Theorie auf Matrizen anwenden. Die Menge $\mathrm{Mat}(n, n; K)$ ist ein Vektorraum der Dimension n^2 und obendrein ein Ring mit der Matrizenaddition und Matrizenmultiplikation. Wie üblich setzen wir für $A \in \mathrm{Mat}(n, n; K)$:

$$A^m = A \cdots A \quad (m\text{-mal}), \quad A^0 = E.$$

Wir setzen nun Matrizen in Polynome ein, d.h. wir betrachten die Einsetzungsabbildung

$$\begin{array}{cccc} \varphi_A : & K[x] & \longrightarrow & \mathrm{Mat}(n, n; K) \\ & P(x) = \sum_{i=0}^m a_i x^i & \longmapsto & P(A) := \sum_{i=0}^m a_i A^i \end{array}.$$

Die Abbildung φ_A ist linear, also ein Homomorphismus von K-Vektorräumen und sogar ein Ringhomomorphismus:

$$\begin{aligned} (P + Q)(A) &= P(A) + Q(A), \\ (\lambda P)(A) &= \lambda P(A), \\ (P \cdot Q)(A) &= P(A)Q(A). \end{aligned}$$

Das Bild von φ_A ist der Untervektorraum

$$K[A] := \mathrm{Span}\{E, A, A^2, \ldots\}$$

von $\mathrm{Mat}(n, n; K)$. Wir betrachten nun die Menge

$$I_A := \{P \in K[x] \mid P(A) = 0\} = \mathrm{Ker}\, \varphi_A.$$

Aus Lemma 7.2 folgt, dass I_A ein Ideal in $K[x]$ ist. Da $K[x]$ ein Hauptidealring ist, gibt es (falls $I_A \neq \{0\}$) genau ein normiertes Polynom $\mu_A \in K[x]$ mit

$$I_A = \langle \mu_A \rangle.$$

Satz 7.7 *Es gibt genau ein normiertes Polynom* $0 \neq \mu_A \in K[x]$ *mit folgenden Eigenschaften:*

(i) $\mu_A(A) = 0$.

(ii) *Ist* $P \in K[x]$ *ein Polynom mit* $P(A) = 0$, *so ist* $P = Q \cdot \mu_A$.

(iii) *Unter allen normierten Polynomen* $P \in K[x]$ *mit* $P(A) = 0$ *hat* μ_A *minimalen Grad.*

Definition Das Polynom $\mu_A \in K[x]$ heißt das *Minimalpolynom* von A.

Beweis. Wir zeigen zunächst $I_A \neq \{0\}$. Es ist $K[A] \subset \mathrm{Mat}(n, n; K)$, also gilt

$$\dim K[A] \leq n^2 =: N.$$

Daher sind die Matrizen

$$E, A, A^2, \ldots, A^N$$

linear abhängig, d.h. es gibt $a_0, a_1, \ldots, a_N \in K$, nicht alle gleich 0, mit

$$a_0 E + a_1 A + \cdots + a_N A^N = 0.$$

Also ist

$$0 \neq P(x) := a_0 + a_1 x + \cdots + a_N x^N \in I_A.$$

Nach Satz 7.6 ist dann

$$I_A = \langle \mu_A \rangle$$

für ein eindeutig bestimmtes normiertes Polynom $\mu_A \neq 0$. □

Beispiel 7.3 Es sei

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

Man rechnet leicht aus:

$$A^2 = \begin{pmatrix} 0 & 0 & 1 & & 0 \\ & \ddots & \ddots & \ddots & \\ & & 0 & 0 & 1 \\ & & & 0 & 0 \\ 0 & & & & 0 \end{pmatrix}, \ldots, A^n = 0.$$

Daraus folgt

$$\mu_A(x) = x^n.$$

Satz 7.8 *Es gilt*
$$\deg \mu_A = \dim K[A].$$

Beweis. Es sei $m := \dim K[A]$. Dann sind

$$E, A, A^2, \ldots, A^m$$

linear abhängig. Wie im Beweis von Satz 7.7 zeigt man, dass es ein (normiertes) Polynom $P \in K[x]$ gibt mit $\deg P \leq m$ und $P(A) = 0$. Also ist

$$\deg \mu_A \leq m = \dim K[A].$$

Es sei umgekehrt $m' = \deg \mu_A$. Zum Beweis von $\dim K[A] \leq m'$ betrachten wir
$$U := \mathrm{Span}\{E, A, \ldots, A^{m'-1}\}.$$
Dann sind $E, A, \ldots, A^{m'-1}$ linear unabhängig, denn andernfalls wäre $m' = \deg \mu_A$ nicht minimal. Also ist $\dim U = m'$. Es genügt zu zeigen:

$$K[A] \subset U.$$

Dazu ist zu zeigen, dass

$$A^s \in U \quad \text{für } s \geq m'.$$

Es sei
$$\mu_A(x) = a_0 + a_1 x + \cdots + a_{m'-1} x^{m'-1} + x^{m'}.$$
Da $\mu_A(A) = 0$ folgt

$$A^{m'} = -a_0 E - \cdots - a_{m'-1} A^{m'-1} \in U.$$

Durch Multiplikation mit A auf beiden Seiten folgt dann aber auch

$$A^{m'+1} = -a_0 A - \cdots - a_{m'-1} A^{m'} \in U.$$

Die Behauptung folgt dann durch Induktion. □

Korollar 7.5 *Ist* $m = \deg \mu_A$, *so ist* $\{E, A, \ldots, A^{m-1}\}$ *eine Basis von* $K[A]$.

Korollar 7.6 *Eine Matrix* $A \in \mathrm{Mat}(n, n; K)$ *ist genau dann invertierbar, wenn* $\mu_A(0) \neq 0$ *gilt. In diesem Fall liegt* $A^{-1} \in K[A]$.

Beweis. Es sei

$$\mu_A(x) = a_0 + a_1 x + \cdots + a_{m-1}x^{m-1} + x^m$$

das Minimalpolynom von A. Dann gilt

$$a_1 A + \cdots + a_{m-1}A^{m-1} + A^m = -a_0 E.$$

Setzen wir

$$B := a_1 E + \cdots + a_{m-1}A^{m-2} + A^{m-1},$$

so folgt

$$AB = -a_0 E.$$

Nach Korollar 7.5 ist $B \neq 0$. Ist A nicht invertierbar, so gilt $\det A = 0$, also $\mu_A(0) = a_0 = 0$. Ist A invertierbar, so ist $B = -a_0 A^{-1}$ und $\mu_A(0) = a_0 \neq 0$. □

Lemma 7.3 (Invarianz) *Sind die Matrizen $A, B \in \mathrm{Mat}(n,n;K)$ ähnlich, dann stimmen die Minimalpolynome μ_A und μ_B überein.*

Beweis. Es sei $P \in K[x]$ und $B = S^{-1}AS$. Dann gilt

$$\begin{aligned} P(S^{-1}AS) &= a_0 E + a_1 S^{-1}AS + \cdots + a_m (S^{-1}AS)^m \\ &= a_0 S^{-1}ES + a_1 S^{-1}AS + \cdots + a_m S^{-1}A^m S \\ &= S^{-1}P(A)S. \end{aligned}$$

Daraus folgt

$$P(A) = 0 \Leftrightarrow P(B) = P(S^{-1}AS) = 0.$$

□

Damit können wir definieren:

Definition Es sei V ein endlich dimensionaler K-Vektorraum und $f : V \to V$ ein Endomorphismus. Dann ist das *Minimalpolynom* μ_f von f das Minimalpolynom einer Darstellungsmatrix von f.

Nun zeigen wir, dass auch das charakteristische Polynom einer Matrix A in I_A liegt.

Satz 7.9 (Satz von Cayley-Hamilton) *Es sei P_A das charakteristische Polynom einer Matrix $A \in \mathrm{Mat}(n,n;K)$. Dann gilt*

$$P_A(A) = 0.$$

Daraus folgt unmittelbar:

Korollar 7.7 *Das Minimalpolynom* μ_A *teilt das charakteristische Polynom* P_A *einer Matrix* $A \in \mathrm{Mat}(n, n; K)$.

Beweis von Satz 7.9. Wir wenden einen Trick an. Wir setzen

$$B(x) := (A - xE)^T \in \mathrm{Mat}(n, n; K[x]).$$

Dann gilt

$$\det B(x) = P_A(x) \in K[x].$$

Nun ersetzen wir die Unbestimmte x durch die Matrix A und jeden Eintrag a_{ij} durch die Matrix $a_{ij}E$. Das ergibt

$$B(A) = \begin{pmatrix} a_{11}E - A & a_{21}E & \cdots & a_{n1}E \\ a_{12}E & a_{22}E - A & \cdots & a_{n2}E \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n}E & a_{2n}E & \cdots & a_{nn}E - A \end{pmatrix} \in \mathrm{Mat}(n, n; K[A]).$$

Diese Matrix kann mit einem Spaltenvektor des K^{n^2} multipliziert werden, d.h. einem Spaltenvektor, dessen Einträge wiederum Spaltenvektoren des K^n sind. Insbesondere gilt

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

Nun sei $B^*(x) \in \mathrm{Mat}(n, n; K[x])$ die zu $B(x)$ adjungierte Matrix, die wir in LA I definiert haben. Ihre Einträge sind entsprechend der Definition Polynome vom Grad $\leq n - 1$, und es gilt

$$B^*(x)B(x) = (\det B(x))E = P_A(x)E.$$

Setzen wir nun A für x ein, so folgt

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = B^*(A)B(A) \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} P_A(A)e_1 \\ \vdots \\ P_A(A)e_n \end{pmatrix}.$$

Also ist $P_A(A) = 0$. □

8 Diagonalisierbarkeit

Nun kommen wir zurück auf das Problem, für die Darstellungsmatrix eines Endomorphismus eines endlich dimensionalen K-Vektorraums eine Normalform zu finden. Zunächst betrachten wir noch einmal die Diagonalisierbarkeit.

Es sei V ein K-Vektorraum der Dimension n und $f : V \to V$ ein Endomorphismus.

Definition Es sei λ ein Eigenwert von f.

(i) Die *algebraische Vielfachheit* von λ, in Zeichen $\nu_{\mathrm{alg}}(f, \lambda)$, ist die Vielfachheit von λ als Nullstelle des charakteristischen Polynoms.

(ii) Die *geometrische Vielfachheit* von λ, in Zeichen $\nu_{\mathrm{geom}}(f, \lambda)$, ist die Dimension des Eigenraums $\mathrm{Eig}(f, \lambda)$.

Lemma 8.1 *Ist λ Eigenwert von f, so gilt*

$$1 \leq \nu_{\mathrm{geom}}(f, \lambda) \leq \nu_{\mathrm{alg}}(f, \lambda).$$

Beweis. Es sei $(v_1, \ldots, v_s)$ eine Basis von $\mathrm{Eig}(f, \lambda)$. Da λ Eigenwert von f ist, gilt $s \geq 1$. Wir ergänzen diese Basis zu einer Basis

$$B = (v_1, \ldots, v_s, v_{s+1}, \ldots, v_n)$$

von V. Dann ist

$$A := M_B^B(f) = \left(\begin{array}{ccc|c} \lambda & & 0 & \\ & \ddots & & * \\ 0 & & \lambda & \\ \hline & 0 & & A' \end{array} \right).$$

Daraus folgt

$$P_f(x) = (x - \lambda)^s P_{A'}(x)$$

und damit

$$\nu_{\mathrm{geom}}(f, \lambda) = \dim \mathrm{Eig}(f, \lambda) = s \leq \nu_{\mathrm{alg}}(f, \lambda).$$

□

Theorem 8.1 *Es sei V ein n-dimensionaler K-Vektorraum und $f : V \to V$ ein Endomorphismus von V. Dann sind die folgenden Bedingungen äquivalent:*

(i) *f ist diagonalisierbar.*

(ii) *Das charakteristische Polynom zerfällt in Linearfaktoren und es gilt $\nu_{\text{geom}}(f, \lambda) = \nu_{\text{alg}}(f, \lambda)$ für alle Eigenwerte λ von f.*

(iii) *Sind $\lambda_1, \ldots, \lambda_k$ die paarweise verschiedenen Eigenwerte von f, so ist*

$$V = \mathrm{Eig}(f, \lambda_1) \oplus \cdots \oplus \mathrm{Eig}(f, \lambda_k).$$

Beweis. (i) $\Rightarrow$ (ii): Es sei f diagonalisierbar und $\lambda_1, \ldots, \lambda_k$ die paarweise verschiedenen Eigenwerte von f. Zu λ_i $(i = 1, \ldots, k)$ betrachten wir eine Basis

$$(v_1^{(i)}, \ldots, v_{s_i}^{(i)}) \text{ von } \mathrm{Eig}(f, \lambda_i).$$

Setzen wir $r_i := \nu_{\text{alg}}(f, \lambda_i)$, so gilt

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

Daraus folgt aber $s_i = r_i$ für alle $i = 1, \ldots, k$.

(ii) $\Rightarrow$ (iii): Es sei

$$W := \mathrm{Eig}(f, \lambda_1) + \cdots + \mathrm{Eig}(f, \lambda_k).$$

Nach I, Satz 19.3, und der Bedingung (iii) in Satz 1.5 folgt

$$W = \mathrm{Eig}(f, \lambda_1) \oplus \cdots \oplus \mathrm{Eig}(f, \lambda_k).$$

Aus (ii) und Satz 1.6 (iii) folgt dann $W = V$.

(iii) $\Rightarrow$ (i): Für jedes $i = 1, \ldots, k$ sei

$$(v_1^{(i)}, \ldots, v_{s_i}^{(i)}) \text{ eine Basis von } \mathrm{Eig}(f, \lambda_i).$$

Nach Satz 1.6 (ii) ist dann

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

eine Basis von V. Da sie nach Definition aus Eigenvektoren von f besteht, ist f diagonalisierbar. □

Als Anwendung von Theorem 8.1 betrachten wir das Problem, zwei Endomorphismen mit einer gemeinsamen Basis zu diagonalisieren (*simultane Diagonalisierung*).

Bemerkung 8.1 Angenommen, die Matrizen $A, B \in \mathrm{Mat}(n, n; K)$ lassen sich simultan diagonalisieren. Das bedeutet, dass es eine Matrix $S \in \mathrm{GL}(n; K)$ gibt mit

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

wobei D und $\widetilde{D}$ Diagonalmatrizen sind. Dann gilt

$$BA = S^{-1}\widetilde{D}SS^{-1}DS = S^{-1}\widetilde{D}DS = S^{-1}D\widetilde{D}S = S^{-1}DSS^{-1}\widetilde{D}S = AB.$$

Das bedeutet, dass A und B kommutieren müssen.

Satz 8.1 *Sind f, g diagonalisierbare Endomorphismen von V und gilt $f \circ g = g \circ f$, so sind f und g simultan diagonalisierbar.*

Beweis. Nach Theorem 8.1 gilt

$$\begin{aligned} V &= \mathrm{Eig}(f, \lambda_1) \oplus \cdots \oplus \mathrm{Eig}(f, \lambda_k) \\ &= \mathrm{Eig}(g, \mu_1) \oplus \cdots \oplus \mathrm{Eig}(g, \mu_\ell), \end{aligned}$$

wobei $\lambda_1, \ldots, \lambda_k$ bzw. $\mu_1, \ldots, \mu_\ell$ die verschiedenen Eigenwerte von f bzw. g sind. Es sei λ einer der Eigenwerte von f und

$$W := \mathrm{Eig}(f, \lambda).$$

Es sei $w \in W$. Dann gilt

$$f(g(w)) = g(f(w)) = g(\lambda w) = \lambda g(w).$$

Also ist auch $g(w)$ ein Eigenvektor von f zum Eigenwert λ, also liegt auch $g(w)$ in W. Damit gilt $g(W) \subset W$. Setze

$$W_j := W \cap \mathrm{Eig}(g, \mu_j) \quad \text{für } j = 1, \ldots, \ell.$$

Behauptung $W = W_1 \oplus \cdots \oplus W_\ell$.

Beweis. Wegen I, Satz 19.3, und der Bedingung (iii) in Satz 1.5 reicht es zu zeigen:

$$W = W_1 + \cdots + W_\ell.$$

Es sei $w \in W$. Dann gibt es $w_j \in \mathrm{Eig}(g, \mu_j)$, so dass $w = w_1 + \cdots + w_\ell$. Dann gilt

$$f(w) = f(w_1) + \cdots + f(w_\ell) = \lambda w_1 + \cdots + \lambda w_\ell = \lambda w.$$

Da $f(w_j) \in \mathrm{Eig}(g, \mu_j)$ und $\lambda w_j \in \mathrm{Eig}(g, \mu_j)$ und die Darstellung von w in

$$\mathrm{Eig}(g, \mu_1) \oplus \cdots \oplus \mathrm{Eig}(g, \mu_\ell)$$

eindeutig ist, folgt

$$f(w_j) = \lambda w_j,$$

also $w_j \in W$ und somit $w_j \in W_j$. □

Da die Behauptung für alle Eigenwerte λ von f gilt, folgt die Aussage des Satzes. □

9 Nilpotente Endomorphismen

Wie wir in Theorem 8.1 gesehen haben, gibt es zwei Bedingungen für die Diagonalisierbarkeit:

(a) Das charakteristische Polynom muss in Linearfaktoren zerfallen, und

(b) die geometrische Vielfachheit muss gleich der algebraischen Vielfachheit der Eigenwerte sein.

Wir untersuchen nun, welche Aussage man noch treffen kann, wenn nur die Bedingung (a) erfüllt ist.

Definition Eine Matrix $A = (a_{ij}) \in \mathrm{Mat}(n, n; K)$ heißt *obere Dreiecksmatrix*, wenn $a_{ij} = 0$ für $i > j$ gilt.

Satz 9.1 *Für einen Endomorphismus f eines n-dimensionalen Vektorraums sind die folgenden Bedingungen äquivalent:*

(i) *Es gibt eine Basis B, so dass $M_B^B(f)$ eine obere Dreiecksmatrix ist.*

(ii) *Das charakteristische Polynom P_f zerfällt in Linearfaktoren, d.h.*

$$P_f(x) = \pm(x - \lambda_1) \cdots (x - \lambda_n) \text{ mit } \lambda_1, \ldots, \lambda_n \in K.$$

Beweis. (i) $\Rightarrow$ (ii): Dies folgt aus I, Satz 18.1.

(ii) $\Rightarrow$ (i): Wir führen den Beweis durch Induktion über n. Für $n = 0, 1$ ist die Behauptung klar. Es sei $n \geq 2$ und v_1 ein Eigenvektor zu dem Eigenwert λ_1. Wir ergänzen ihn zu einer Basis

$$B = (v_1, w_2, \ldots, w_n)$$

von V. Dann gilt

$$V = U_1 \oplus W \text{ mit } U_1 := \mathrm{Span}\{v_1\} \text{ und } W := \mathrm{Span}\{w_2, \dots, w_n\}$$

und

$$M_B^B(f) = \left(\begin{array}{c|ccc} \lambda_1 & a_{12} & \cdots & a_{1n} \\ \hline 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{array}\right).$$

Wir definieren nun lineare Abbildungen $h : W \to U_1$ und $g : W \to W$ durch

$$h(w_j) = a_{1j} v_1 \text{ und } g(w_j) = a_{2j} w_2 + \cdots + a_{nj} w_n$$

für $j = 2, \dots, n$. Dann gilt

$$f(w) = h(w) + g(w) \text{ für alle } w \in W.$$

Für die charakteristischen Polynome gilt

$$P_f(x) = (x - \lambda_1) P_g(x), \text{ also } P_g(x) = \pm(x - \lambda_2) \cdots (x - \lambda_n).$$

Deswegen können wir die Induktionsvoraussetzung auf $g : W \to W$ anwenden. Demnach gibt es eine Basis $(v_2, \dots, v_n)$ von W, bezüglich der g durch eine obere Dreiecksmatrix dargestellt wird. Für f gilt dann

$$f(v_j) = h(v_j) + g(v_j) \in \mathrm{Span}\{v_1, \dots, v_j\} \text{ für } j = 2, \dots, n.$$

Also ist auch die Darstellungsmatrix von f bezüglich der Basis $B = (v_1, \dots, v_n)$ eine obere Dreiecksmatrix. □

Wir betrachten nun eine Anwendung dieses Satzes.

Definition (i) Ein Endomorphismus $f : V \to V$ heißt *nilpotent*, wenn $f^k = 0$ für ein $k \geq 1$ ist.

(ii) Eine Matrix $A \in \mathrm{Mat}(n, n; K)$ heißt *nilpotent*, wenn $A^k = 0$ für ein $k \geq 1$ ist.

Lemma 9.1 *Es sei A nilpotent.*

(i) *Ist B ähnlich zu A, dann ist auch B nilpotent.*

(ii) 0 *ist der einzige Eigenwert von A.*

Beweis.

(i): Es sei $B = S^{-1}AS$ und $A^k = 0$. Dann gilt

$$B^k = (S^{-1}AS)(S^{-1}AS)\cdots(S^{-1}AS) = S^{-1}A^kS = 0.$$

(ii): Es sei $A^k = 0$. Aus $\det A^k = 0$ folgt $\det A = 0$. Deshalb ist 0 ein Eigenwert von A. Dies ist auch der einzige Eigenwert: Ist $\lambda \in K$ ein Eigenwert von A mit Eigenvektor $x \neq 0$, dann gilt

$$A^k x = \lambda^k x = 0$$

und daraus folgt $\lambda = 0$. □

Satz 9.2 *Für einen Endomorphismus f eines n-dimensionalen Vektorraums V sind die folgenden Bedingungen äquivalent:*

(i) *f ist nilpotent.*

(ii) *Es sei B eine Basis von V. Dann ist die Darstellungsmatrix $M_B^B(f)$ von f bezüglich der Basis B nilpotent.*

(iii) *Es gilt $P_f(x) = \pm x^n$.*

(iv) *Es gilt $f^d = 0$ für ein d mit $1 \leq d \leq n$.*

(v) *Es gibt eine Basis B von V, so dass*

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

Beweis. (i) $\Leftrightarrow$ (ii) ist klar.

(ii) $\Rightarrow$ (iii): Es sei $A := M_B^B(f)$. Nach Lemma 9.1 (ii) ist 0 der einzige Eigenwert von A. Deswegen hat das charakteristische Polynom die Gestalt $P_f(x) = \pm x^n$.

(iii) $\Rightarrow$ (iv): Aus dem Satz von Cayley-Hamilton folgt $\mu_f(x) = x^d$ für ein d mit $1 \leq d \leq n$. Das bedeutet $f^d = 0$.

(iv) $\Rightarrow$ (i) ist klar.

(iii) $\Rightarrow$ (v) folgt aus Satz 9.1.

(v) $\Rightarrow$ (ii): Es sei $A = M_B^B(f) = (a_{ij})$ mit $a_{ij} = 0$ für $i \geq j$.

Behauptung $A^r = (a_{ij}^{(r)})$ mit $a_{ij}^{(r)} = 0$ für $i \geq j + 1 - r$.

Beweis. Wir führen Induktion nach r durch.

Induktionsanfang $r = 1$: Dies gilt nach Voraussetzung.

Induktionsschritt $r - 1 \to r$: Es gilt

$$a_{ij}^{(r)} = \sum_{\ell=1}^{n} a_{i\ell} a_{\ell j}^{(r-1)}.$$

Es sei nun $i \geq j + 1 - r$ und $1 \leq \ell \leq n$. Wir unterscheiden zwei Fälle:

Fall 1: $j + 1 - r \geq \ell$. Dann ist $i \geq \ell$, also $a_{i\ell} = 0$ nach Voraussetzung.

Fall 2: $\ell > j+1-r$. Dann ist $\ell \geq j+1-(r-1)$. Dann gilt aber $a_{\ell j}^{(r-1)} = 0$ nach Induktionsvoraussetzung.

Also gilt

$$a_{ij}^{(r)} = 0.$$

□

Aus der Behauptung folgt $A^n = 0$. Ist A' die Darstellungsmatrix von f bezüglich einer anderen Basis, so ist A' ähnlich zu A. Die Behauptung folgt damit aus Lemma 9.1 (i). □

Wir wollen nun zeigen, dass wir die Matrix in Satz 9.2 (v) noch auf eine einfachere Gestalt bringen können.

Definition Die Matrix

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

heißt *Jordanmatrix* von der Ordnung k.

Nach Beispiel 7.3 ist J_k nilpotent, genauer gilt $J_k^k = 0$ und k ist die minimale Potenz mit dieser Eigenschaft. Wie sieht J_1 aus?

Theorem 9.1 (Jordannormalform nilpotenter Endomorphismen) *Es sei f ein nilpotenter Endomorphismus eines K-Vektorraums V und $d := \min\{\ell \mid f^\ell = 0\}$ (d heißt auch der* Nilpotenzindex *von f). Dann gibt es eindeutig bestimmte Zahlen $s_1, \ldots, s_d \in \mathbb{N}$ mit*

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

und eine Basis B von V, so dass

$$M_B^B(f) = \begin{pmatrix} \boxed{J_d} & & & & & & & & \\ & \ddots & & & & & & & \\ & & \boxed{J_d} & & & & & & \\ & & & \boxed{J_{d-1}} & & & 0 & & \\ & & & & \ddots & & & & \\ & & & & & \boxed{J_{d-1}} & & & \\ & & & 0 & & & \ddots & & \\ & & & & & & & \boxed{J_1} & \\ & & & & & & & & \ddots \\ & & & & & & & & & \boxed{J_1} \end{pmatrix}$$

Beweis. Wir definieren $U_\ell := \operatorname{Ker} f^\ell$ und betrachten die Kette von Unterräumen

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

Dabei gilt $U_d = V$ nach Definition von d, und da d die minimale Potenz mit $f^d = 0$ ist, sind alle Inklusionen echt.

Behauptung

(1) Für $1 \leq \ell \leq d$ ist $f^{-1}(U_{\ell-1}) = U_\ell$, insbesondere $f(U_\ell) = U_{\ell-1}$.

(2) Ist W ein Unterraum von V mit $W \cap U_\ell = \{0\}$ für ein $1 \leq \ell \leq d$, so ist $f|_W$ injektiv.

Beweis.
(1): Es gilt

$$v \in f^{-1}(U_{\ell-1}) \Leftrightarrow f(v) \in U_{\ell-1} \Leftrightarrow 0 = f^{\ell-1}(f(v)) = f^\ell(v) \Leftrightarrow v \in U_\ell.$$

(2) Es gilt $\operatorname{Ker} f = U_1 \subset U_\ell$ für jedes $1 \leq \ell \leq d$, also $W \cap \operatorname{Ker} f = \{0\}$. □

Nun konstruieren wir schrittweise eine direkte Summenzerlegung von V. Zunächst wählen wir ein Komplement $W_d \subset V$ von U_{d-1} in $V = U_d$:

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

Aus Behauptung (1) folgt dann

(a) $f(W_d) \subset U_{d-1}$ und

(b) $f(W_d) \cap U_{d-2} = \{0\}$.

Denn aus $W_d \subset U_d$ und $f(U_d) = U_{d-1}$ folgt (a). Aus $f^{-1}(U_{d-2}) = U_{d-1}$ und $W_d \cap U_{d-1} = \{0\}$ folgt (b). Also gibt es eine Zerlegung

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

Fahren wir so fort, so erhalten wir folgendes Schema:

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

Dabei zeigen die Pfeile an, wie f die entsprechenden Unterräume abbildet. Jede Zeile ist eine Zerlegung von V, wegen $U_0 = \{0\}$ ist insbesondere

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

Da die Abbildungen

$$W_d \xrightarrow{f|_{W_d}} W_{d-1} \xrightarrow{f|_{W_{d-1}}} \ldots \xrightarrow{f|_{W_2}} W_1$$

nach Behauptung (2) alle injektiv sind, können wir mit einer Basis von W_d anfangen, das Bild dieser Basis unter $f|_{W_d}$ zu einer Basis von W_{d-1} ergänzen, usw., bis wir zu einer Basis von V gelangen:

$$\begin{array}{cccc}
w_1^{(d)}, & f(w_1^{(d)}), & \ldots, & f^{d-1}(w_1^{(d)}), \\
\vdots & \vdots & & \vdots \\
w_{s_d}^{(d)}, & f(w_{s_d}^{(d)}), & \ldots, & f^{d-1}(w_{s_d}^{(d)}), \\
 & w_1^{(d-1)}, & \ldots, & f^{d-2}(w_1^{(d-1)}), \\
 & \vdots & & \vdots \\
 & w_{s_{d-1}}^{(d-1)}, & \ldots, & f^{d-2}(w_{s_{d-1}}^{(d-1)}), \\
 & & & \vdots \\
 & & & w_1^{(1)}, \\
 & & & \vdots \\
 & & & w_{s_1}^{(1)}.
\end{array}$$

Dabei ist die erste Spalte eine Basis von W_d, die zweite Spalte eine Basis von W_{d-1}, und schließlich die letzte Spalte eine Basis von $W_1 = U_1 = \operatorname{Ker} f$. Ordnen wir die Basis nun so an, dass wir die Zeilen von oben nach unten lesen, aber in jeder Zeile umgekehrt, also von rechts nach links, laufen, so erhalten wir eine Basis von V, bezüglich der die Darstellungsmatrix von f die angebene Gestalt hat.

Wir müssen nun noch zeigen, dass die Zahlen $s_1, \ldots, s_d$ eindeutig bestimmt sind. Dazu sei $\widetilde{W}_\ell$ ein Komplement von $f(W_{\ell+1})$ in W_ℓ, $\ell = 1, \ldots, d$ (hier setzen wir $W_{d+1} = \{0\}$). Dann gilt wegen

$$U_\ell = U_{\ell-1} \oplus f(W_{\ell+1}) \oplus \widetilde{W}_\ell$$

und da $f|_{W_{\ell+1}}$ injektiv ist

$$s_\ell = \dim \widetilde{W}_\ell = \dim U_\ell - \dim U_{\ell-1} - \dim W_{\ell+1}.$$

Damit sind diese Zahlen rekursiv aus den Dimensionen der Kerne von f^ℓ berechenbar. □

10 Die Jordansche Normalform

Es sei K ein beliebiger Körper, V ein endlich dimensionaler K-Vektorraum mit $\dim V \geq 1$ und $f : V \to V$ ein Endomorphismus mit zerfallendem charakteristischen Polynom

$$P_f(x) = \pm(x - \lambda_1)^{r_1} \cdots (x - \lambda_k)^{r_k}, \quad \lambda_1, \ldots, \lambda_k \in K \text{ paarweise verschieden.}$$

Wir haben bereits gesehen, dass dann f durch eine obere Dreiecksmatrix dargestellt werden kann. Dieses Ergebnis soll nun noch präzisiert werden.

Im Allgemeinen gilt

$$\dim \operatorname{Eig}(f, \lambda_i) = \nu_{\text{geom}}(f, \lambda_i) \leq r_i.$$

Gilt hier nicht die Gleichheit, so betrachtet man anstelle des Eigenraums einen größeren Unterraum.

Definition Für einen Eigenwert λ der Vielfachheit $r \geq 1$ nennt man

$$\operatorname{Hau}(f, \lambda) := \operatorname{Ker}(f - \lambda \mathrm{id})^r$$

den *Hauptraum* (oder *verallgemeinerten Eigenraum*) von f zum Eigenwert λ.

Satz 10.1 (Hauptraumzerlegung) *Es sei f ein Endomorphismus von V und*

$$P_f(x) = \pm(x - \lambda_1)^{r_1} \cdots (x - \lambda_k)^{r_k}$$

mit paarweise verschiedenen $\lambda_1, \ldots, \lambda_k \in K$. Es sei

$$V_i := \mathrm{Hau}(f, \lambda_i) \subset V$$

der Hauptraum zum Eigenwert λ_i. Dann gilt:

(1) $\dim V_i = r_i$ *und* $f(V_i) \subset V_i$ *für* $i = 1, \ldots, k$.

(2) $V = V_1 \oplus \cdots \oplus V_k$.

(3) f *hat eine Zerlegung* $f = f_D + f_N$ *mit*

   (a) f_D *ist diagonalisierbar.*

   (b) f_N *ist nilpotent.*

   (c) $f_D \circ f_N = f_N \circ f_D$.

Durch Kombination dieses Satzes mit der Klassifikation nilpotenter Endomorphismen (Theorem 9.1) erhält man das Hauptresultat dieses Abschnitts.

Theorem 10.1 (Jordansche Normalform) *Es sei f ein Endomorphismus von V und*

$$P_f(x) = \pm(x - \lambda_1)^{r_1} \cdots (x - \lambda_k)^{r_k}$$

mit paarweise verschiedenen $\lambda_1, \ldots, \lambda_k \in K$. Dann gibt es eine Basis B von V, so dass

$$M_B^B(f) = \begin{pmatrix} \boxed{\lambda_1 E_{r_1} + N_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{\lambda_k E_{r_k} + N_k} \end{pmatrix},$$

wobei N_i für $i = 1, \ldots, k$ in der Normalform von Theorem 9.1 ist. Ausgeschrieben bedeutet das:

$$\lambda_i E_{r_i} + N_i = \begin{pmatrix} \boxed{\begin{matrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{matrix}} & & & & & 0 \\ & \ddots & & & & \\ & & \boxed{\begin{matrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{matrix}} & & & \\ & & & \lambda_i & & \\ & & & & \ddots & \\ 0 & & & & & \lambda_i \end{pmatrix}.$$

Die Eigenwerte $\lambda_1, \ldots, \lambda_k$, die Zahlen $r_1, \ldots, r_k$ sowie die Zahlen $s_j^{(i)}$ mit

$$d_i s_{d_i}^{(i)} + (d_i - 1) s_{d_i - 1}^{(i)} + \cdots + s_1^{(i)} = r_i, \quad i = 1, \ldots k$$

nach Theorem 9.1 sind durch f eindeutig bestimmt. Man nennt sie *Invarianten* von f.

Beweis von Theorem 10.1. Setze für $i = 1, \ldots, k$

$$V_i := \mathrm{Hau}(f, \lambda_i) \text{ und } g_i := (f - \lambda_i \mathrm{id})|_{V_i}.$$

Anwendung von Theorem 9.1 auf g_i ergibt eine Basis B_i von V_i. Nach Satz 10.1 setzen sich die Basen $B_1, \ldots, B_k$ zu einer Basis B mit der gewünschten Eigenschaft zusammen. □

Korollar 10.1 *Für einen Endomorphismus f von V sind die folgenden Bedingungen äquivalent:*

(i) *f ist diagonalisierbar.*

(ii) *$\mu_f(x) = (x - \lambda_1) \cdots (x - \lambda_k)$, wobei $\lambda_1, \ldots, \lambda_k$ die verschiedenen Eigenwerte von f sind.*

Beweis. Es sei wie oben

$$d_i := \min\{\ell_i \mid g_i^{\ell_i} = 0\}.$$

Dann gilt für das Minimalpolynom von g_i

$$\mu_{g_i}(x) = x^{d_i}.$$

Die Abbildung g_i ist aber genau dann diagonalisierbar, wenn $d_i = 1$. $\square$

Nun wollen wir den Satz über die Hauptraumzerlegung beweisen. Dazu betrachten wir für einen Eigenwert λ von f die Abbildung

$$g := f - \lambda \mathrm{id}.$$

Die folgenden Überlegungen gelten für einen beliebigen Endomorphismus g und seine Potenzen. Man hat zwei Ketten von Unterräumen:

$$\begin{array}{ccccccccccc} \{0\} & \subset & \operatorname{Ker} g & \subset & \operatorname{Ker} g^2 & \subset & \dots & \subset & \operatorname{Ker} g^\ell & \subset & \dots \\ V & \supset & \operatorname{Im} g & \supset & \operatorname{Im} g^2 & \supset & \dots & \supset & \operatorname{Im} g^\ell & \supset & \dots \end{array}$$

Da V endlich dimensional ist, müssen die beiden Ketten stationär werden, d.h. irgendwann sind die Inklusionen nicht mehr echt. Genauer bedeutet das, dass es d und d' geben muss mit

$$\begin{array}{ccccccccccccc} \{0\} & \subset & \operatorname{Ker} g & \subset & \operatorname{Ker} g^2 & \subset & \dots & \subset & \operatorname{Ker} g^d & = & \operatorname{Ker} g^{d+1} & = & \dots \\ V & \supset & \operatorname{Im} g & \supset & \operatorname{Im} g^2 & \supset & \dots & \supset & \operatorname{Im} g^{d'} & = & \operatorname{Im} g^{d'+1} & = & \dots \end{array}$$

Genauer gilt Folgendes:

Lemma 10.1 (Fitting) *Zu einem Endomorphismus g von V betrachten wir die Zahlen*

$$\begin{array}{rcl} d & := & \min\{\ell \mid \operatorname{Ker} g^\ell = \operatorname{Ker} g^{\ell+1}\}, \\ d' & := & \min\{\ell \mid \operatorname{Im} g^\ell = \operatorname{Im} g^{\ell+1}\}, \\ r & := & \nu_{\mathrm{alg}}(g, 0). \end{array}$$

Dann gilt:

(i) $d = d'$.

(ii) $\operatorname{Ker} g^{d+i} = \operatorname{Ker} g^d$, $\operatorname{Im} g^{d+i} = \operatorname{Im} g^d$ *für alle* $i \in \mathbb{N}$.

(iii) *Die Räume* $U := \operatorname{Ker} g^d$ *und* $W := \operatorname{Im} g^d$ *sind unter* g *invariant.*

(iv) $(g|_U)^d = 0$ *und* $g|_W$ *ist ein Isomorphismus.*

(v) *Für das Minimalpolynom von* $g|_U$ *gilt* $\mu_{g|_U}(x) = x^d$.

(vi) $V = U \oplus W$, $\dim U = r \geq d$, $\dim W = n - r$.

Beweis. Wir betrachten das Diagramm

$$\begin{array}{ccccc} \operatorname{Ker} g^\ell & \subset & V & \xrightarrow{g^\ell} & \operatorname{Im} g^\ell \\ \cap & & \| & & \cup \\ \operatorname{Ker} g^{\ell+1} & \subset & V & \xrightarrow{g^{\ell+1}} & \operatorname{Im} g^{\ell+1} \end{array}$$

Nach der Dimensionsformel gilt

$$\dim V = \dim \operatorname{Ker} g^\ell + \dim \operatorname{Im} g^\ell = \dim \operatorname{Ker} g^{\ell+1} + \dim \operatorname{Im} g^{\ell+1}.$$

Daraus folgt

$$\begin{aligned} \operatorname{Im} g^{\ell+1} = \operatorname{Im} g^\ell \quad & \Leftrightarrow \quad \dim \operatorname{Im} g^{\ell+1} = \dim \operatorname{Im} g^\ell \\ & \Leftrightarrow \quad \dim \operatorname{Ker} g^{\ell+1} = \dim \operatorname{Ker} g^\ell \\ & \Leftrightarrow \quad \operatorname{Ker} g^{\ell+1} = \operatorname{Ker} g^\ell. \end{aligned}$$

Daraus folgt zunächst einmal die Aussage (i).

Weiterhin ist die Aussage $\operatorname{Im} g^{\ell+1} = \operatorname{Im} g^\ell$ gleichbedeutend damit, dass $g|_{\operatorname{Im} g^\ell} : \operatorname{Im} g^\ell \to \operatorname{Im} g^{\ell+1}$ ein Isomorphismus ist. Daraus folgt (ii), (iii) und (iv).

(v): Es genügt zu zeigen, dass $(g|_U)^{d-1} \neq 0$. Angenommen, $(g|_U)^{d-1} = 0$. Dann folgt

$$\operatorname{Ker} g^d = U \subset \operatorname{Ker} g^{d-1}.$$

Da aber $\operatorname{Ker} g^{d-1} \subset \operatorname{Ker} g^d$, erhalten wir $\operatorname{Ker} g^{d-1} = \operatorname{Ker} g^d$ im Widerspruch zur Definition von d.

(vi): Wir zeigen zunächst $V = U \oplus W$. Es sei $v \in U \cap W$. Dann ist $g^d(v) = 0$ und $v = g^d(w)$ für ein $w \in V$. Daraus folgt $g^{2d}(w) = 0$, also $w \in \operatorname{Ker} g^{2d}$. Nach (ii) gilt $\operatorname{Ker} g^{2d} = \operatorname{Ker} g^d$. Damit folgt $w \in \operatorname{Ker} g^d$ und somit

$$v = g^d(w) = 0.$$

Nach Definition von U ist $\dim U \geq d$, Denn es gilt

$$\{0\} \subsetneq \operatorname{Ker} g \subsetneq \ldots \subsetneq \operatorname{Ker} g^{d-1} \subsetneq \operatorname{Ker} g^d,$$

und in jedem Schritt erhöht sich die Dimension mindestens um 1. Nun gilt

$$P_g(x) = x^r \cdot Q(x) = P_{g|_U}(x) \cdot P_{g|_W}(x) \text{ mit } Q(0) \neq 0.$$

Auf der anderen Seite gilt

$$\begin{aligned} P_{g|_U}(x) &= \pm x^m \text{ mit } m = \dim U, \\ P_{g|_W}(0) &\neq 0 \quad (\text{da } g|_W \text{ ein Isomorphismus nach (iv)}). \end{aligned}$$

Daraus folgt $m = r$, was zu zeigen war. □

Beweis von Satz 10.1. Wir führen Induktion über die Zahl k der verschiedenen Eigenwerte durch. Für $k = 1$ ist der Satz trivial. Es sei nun $k \geq 2$. Zu λ_1 definieren wir

$$g := f - \lambda_1 \cdot \mathrm{id}.$$

Dann gilt

$$P_g(x + \lambda_1) = P_f(x), \quad \text{also } \nu_{\mathrm{alg}}(g, 0) = \nu_{\mathrm{alg}}(f, \lambda_1) = r_1.$$

Nach Lemma 10.1 gilt

$$V = \mathrm{Hau}(f, \lambda_1) \oplus W, \quad W = \mathrm{Im}\, g^d,$$

und die beiden Summanden werden von g und damit auch von $f = g + \lambda_1\mathrm{id}$ invariant gelassen. Außerdem gilt

$$P_{f|_W}(x) = \pm(x - \lambda_2)^{r_2} \cdots (x - \lambda_k)^{r_k}.$$

Damit können wir auf $f|_W$ die Induktionsannahme anwenden und erhalten (1) und (2).

Zum Beweis von (3) bemerken wir zunächst, dass es nach Satz 9.1 eine Basis B von V gibt, so dass die Darstellungsmatrix $M_B^B(f)$ eine obere Dreiecksmatrix ist. Diese Matrix schreiben wir als

$$M_B^B(f) = \begin{pmatrix} \boxed{\lambda_1 E_{r_1} + N_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{\lambda_k E_{r_k} + N_k} \end{pmatrix}.$$

Dann ist jedes N_i eine obere Dreiecksmatrix mit Nullen auf der Diagonalen, nach Satz 9.2 also nilpotent. Setze

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

Dann rechnet man aus:

$$DN = \begin{pmatrix} \boxed{\lambda_1 N_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{\lambda_k N_k} \end{pmatrix} = ND.$$

□

Bemerkung 10.1 Man kann zeigen, dass die Zerlegung $f = f_D + f_N$ in Satz 10.1 sogar eindeutig ist, wenn man (a), (b) und (c) verlangt.

Bemerkung 10.2 Es sei f ein Endomorphismus mit charakteristischem Polynom

$$P_f(x) = \pm(x-\lambda_1)^{r_1} \cdots (x-\lambda_k)^{r_k}$$

und Minimalpolynom

$$\mu_f(x) = \pm(x-\lambda_1)^{d_1} \cdots (x-\lambda_k)^{d_k}$$

mit paarweise verschiedenen $\lambda_1, \dots, \lambda_k \in K$. Die Matrix

$$\begin{pmatrix} \lambda_i & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{pmatrix} \in \mathrm{Mat}(m, m; K)$$

nennt man auch einen *Jordanblock* der Länge m zum Eigenwert λ_i. Dann folgt aus Theorem 10.1, dass r_i die Summe der Längen aller Jordanblocks zum Eigenwert λ_i und d_i die Länge des größten Jordanblocks zum Eigenwert λ_i ist.

Beispiel 10.1 Gegeben sei die Matrix

$$A := \begin{pmatrix} -5 & 15 & 11 \\ -5 & 11 & 5 \\ 3 & -6 & -2 \end{pmatrix}.$$

Dann gilt

$$P_A(x) = -x^3 + 4x^2 - 5x + 2 = -(x-1)^2(x-2).$$

Damit ist $k = 2$, $\lambda_1 = 1$, $r_1 = 2$, $\lambda_2 = 2$, $r_2 = 1$. Wir setzen

$$B_1 := A - E = \begin{pmatrix} -6 & 15 & 11 \\ -5 & 10 & 5 \\ 3 & -6 & -3 \end{pmatrix}, \quad B_2 := A - 2E = \begin{pmatrix} -7 & 15 & 11 \\ -5 & 9 & 5 \\ 3 & -6 & -4 \end{pmatrix}.$$

Dann gilt

$$B_1^2 = \begin{pmatrix} -6 & -6 & -24 \\ -5 & -5 & -20 \\ 3 & 3 & 12 \end{pmatrix}.$$

Daraus folgt

$$1 = \dim \operatorname{Eig}(A,1) = \dim \operatorname{Ker} B_1 < \dim \operatorname{Ker} B_1^2 = \dim \operatorname{Hau}(A,1) = 2.$$

Damit ist A nicht diagonalisierbar. Durch Lösen der Gleichungssysteme $B_1^2 x = 0$ und $B_2 x = 0$ erhält man Basen

$$\begin{aligned} &\{(4,0,-1)^T, (0,4,-1)^T\} \text{ von } \operatorname{Hau}(A,1), \\ &\{(6,5,-3)^T\} \text{ von } \operatorname{Hau}(A,2) = \operatorname{Eig}(A,2). \end{aligned}$$

Daraus bilden wir die Matrix

$$T := \begin{pmatrix} 4 & 0 & 6 \\ 0 & 4 & 5 \\ -1 & -1 & -3 \end{pmatrix} \text{ mit } T^{-1} = \frac{1}{4} \begin{pmatrix} 7 & 6 & 24 \\ 5 & 6 & 20 \\ -4 & -4 & -16 \end{pmatrix}.$$

Dann gilt

$$B := T^{-1} A T = \begin{pmatrix} \frac{-31}{4} & \frac{49}{4} & 0 \\ \frac{-25}{4} & \frac{39}{4} & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Nun transformieren wir die Basis von $\operatorname{Hau}(A,1)$ so, dass ein Basisvektor v ein Eigenvektor von A zum Eigenwert 1 ist. Er hat die Form

$$v = \alpha(4,0,-1)^T + \beta(0,4,-1)^T.$$

Dann müssen α und β der Bedingung

$$\begin{pmatrix} \frac{-35}{4} & \frac{49}{4} \\ \frac{-25}{4} & \frac{35}{4} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

genügen. Daraus folgt, dass man $\alpha = \frac{7}{4}$ und $\beta = \frac{5}{4}$ wählen kann. Mit der neuen Transformationsmatrix

$$\widetilde{T} := \begin{pmatrix} 7 & 0 & 6 \\ 5 & 4 & 5 \\ -3 & -1 & -3 \end{pmatrix} \text{ mit } \widetilde{T}^{-1} := \frac{1}{7} \begin{pmatrix} 7 & 6 & 24 \\ 0 & 3 & 5 \\ -7 & -7 & -28 \end{pmatrix}$$

ergibt sich

$$\widetilde{T}^{-1} A \widetilde{T} = \begin{pmatrix} 1 & 7 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Um nun zur Jordannormalform zu kommen, suchen wir einen Vektor

$$w = \gamma(4, 0, -1)^T + \delta(0, 4, -1)^T$$

mit der Eigenschaft $(B - E)w = v$. Also müssen γ und δ der Bedingung

$$\begin{pmatrix} \frac{-35}{4} & \frac{49}{4} \\ \frac{-25}{4} & \frac{35}{4} \end{pmatrix} \begin{pmatrix} \gamma \\ \delta \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

genügen. Daraus folgt $\gamma = \delta = \frac{1}{2}$. Mit der neuen Transformationsmatrix

$$S := \begin{pmatrix} 7 & 2 & 6 \\ 5 & 2 & 5 \\ -3 & -1 & -3 \end{pmatrix} \text{ mit } S^{-1} := \begin{pmatrix} 1 & 0 & 2 \\ 0 & 3 & 5 \\ -1 & -1 & -4 \end{pmatrix}$$

ergibt sich schließlich

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

11 Affine Quadriken

Wir betrachten nun Quadriken. Die Literatur für diesen Abschnitt ist

- G. Fischer: Analytische Geometrie. Vieweg 1978.

Es sei K im Folgenden immer ein Körper, für den $1 + 1 \neq 0$ ist.

Definition Unter einem *quadratischen Polynom* über K in den Unbestimmten $x_1, \ldots, x_n$ versteht man einen Ausdruck der Gestalt

$$P(x_1, \ldots, x_n) = \sum_{i=1}^{n} a_{ii} x_i^2 + \sum_{1 \leq i < j \leq n} 2a_{ij} x_i x_j + \sum_{i=1}^{n} 2a_{0i} x_i + a_{00},$$

wobei $a_{ij} \in K$ für $0 \leq i \leq j \leq n$.

Definition Eine Teilmenge $Q \subset K^n$ heißt *(affine) Quadrik* (oder *(affine) Hyperfläche zweiter Ordnung*), wenn es ein quadratisches Polynom P gibt, so dass

$$Q = \{(x_1, \ldots, x_n) \in K^n \mid P(x_1, \ldots, x_n) = 0\}.$$

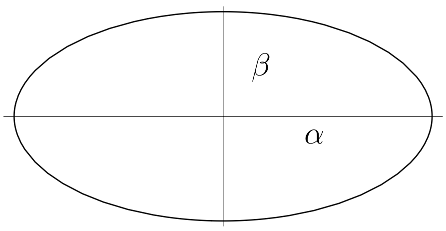

Abbildung 1: Ellipse $\frac{x_1^2}{\alpha^2} + \frac{x_2^2}{\beta^2} = 1$

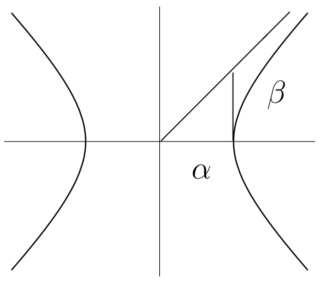

Abbildung 2: Hyperbel $\frac{x_1^2}{\alpha^2} - \frac{x_2^2}{\beta^2} = 1$

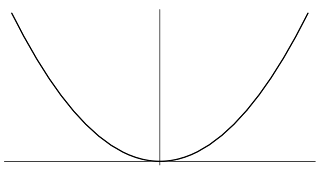

Abbildung 3: Parabel $x_1^2 - x_2 = 0$

Beispiel 11.1 (a) $\frac{x_1^2}{\alpha^2} + \frac{x_2^2}{\beta^2} = 1$, $\alpha, \beta > 0$. Diese Gleichung beschreibt eine *Ellipse* (vgl. Abbildung 1).

(b) $\frac{x_1^2}{\alpha^2} - \frac{x_2^2}{\beta^2} = 1$, $\alpha, \beta > 0$. Diese Gleichung beschreibt eine *Hyperbel* (vgl. Abbildung 2).

(c) $x_1^2 - x_2 = 0$. Diese Gleichung beschreibt eine *Parabel* (vgl. Abbildung 3).

Es ist vorteilhaft, die Gleichung für eine Quadrik durch Matrizen auszudrücken. Dazu setzen wir $a_{ji} := a_{ij}$ für $0 \leq i < j \leq n$ und

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix},$$

$$\mathbf{x} = \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{A} = \left(\begin{array}{c|ccc} a_{00} & a_{01} & \cdots & a_{0n} \\ \hline a_{10} & & & \\ \vdots & & A & \\ a_{n0} & & & \end{array}\right).$$

Dann sind A und $\mathbf{A}$ symmetrisch und es gilt

$$P(x_1, \ldots, x_n) = \mathbf{x}^T \mathbf{A} \mathbf{x}$$

und

$$Q = \{(x_1, \ldots, x_n) \in K^n \mid \mathbf{x}^T \mathbf{A} \mathbf{x} = 0\}.$$

Man nennt $\mathbf{A}$ die *erweiterte Matrix* zu A und $\mathbf{x}$ den *erweiterten Spaltenvektor* zu x.

Definition Es seien V und W K-Vektorräume. Eine Abbildung $f : V \to W$ heißt *affin*, wenn es eine lineare Abbildung $F : V \to W$ und einen Vektor $w_0 \in W$ gibt mit

$$f(v) = F(v) + w_0 \text{ für alle } v \in V.$$

Ist $W = V$ und $F : V \to V$ ein Automorphismus, so nennt man die Abbildung $f : V \to V$ eine *Affinität*.

Eine affine Abbildung ensteht also durch die Hintereinanderschaltung einer linearen Abbildung $F : V \to W$ und einer Translation $t : W \to W$, $w \mapsto w + w_0$.

Beispiel 11.2 Für $V = K^n$ lässt sich eine Affinität wie folgt beschreiben:

$$\begin{array}{rccc} f: & K^n & \longrightarrow & K^n \\ & x & \longmapsto & Ax + b \end{array}$$

mit $A \in \mathrm{GL}(n; K)$ und $b \in K^n$.

Satz 11.1 (i) *Sind f, g Affinitäten, so auch $f \circ g$.*

(ii) *Ist f eine Affinität, so ist f bijektiv und die Umkehrabbildung f^{-1} ist wieder eine Affinität. Es gilt*

$$f^{-1}(v) = F^{-1}(v) - F^{-1}(w_0).$$

Beweis.

(i) Es sei

$$\begin{aligned} f(v) &= F(v) + w_0, \\ g(v) &= G(v) + u_0. \end{aligned}$$

Dann gilt

$$\begin{aligned} (f \circ g)(v) &= f(g(v)) = f(G(v) + u_0) = F(G(v) + u_0) + w_0 \\ &= F(G(v)) + F(u_0) + w_0 = (F \circ G)(v) + (F(u_0) + w_0). \end{aligned}$$

(ii)

$$\begin{aligned} (f^{-1} \circ f)(v) &= F^{-1}(f(v)) - F^{-1}(w_0) \\ &= F^{-1}(F(v) + w_0) - F^{-1}(w_0) \\ &= v + F^{-1}(w_0) - F^{-1}(w_0) = v, \\ (f \circ f^{-1})(v) &= F(f^{-1}(v)) + w_0 \\ &= F(F^{-1}(v) + F^{-1}(w_0)) + w_0 \\ &= v - w_0 + w_0 = v. \end{aligned}$$

□

Wir wollen nun untersuchen, wie sich die Gleichung einer Quadrik bei einer Affinität von K^n ändert. Dafür erweitern wir die obige Schreibweise auf Affinitäten. Es sei $f : K^n \to K^n$, $x \mapsto y$ mit

$$y = Sx + b \quad (S \in \mathrm{GL}(n; K), b \in K^n).$$

Dann definieren wir

$$\mathbf{S} = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline b_1 & & & \\ \vdots & & S & \\ b_n & & & \end{array} \right), \quad \mathbf{y} = \begin{pmatrix} 1 \\ y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

In dieser Schreibweise wird f gegeben durch

$$\mathbf{y} = \mathbf{S}\mathbf{x}.$$

Satz 11.2 *Ist $Q \subset K^n$ eine Quadrik und $f : K^n \to K^n$ eine Affinität, so ist auch $f(Q) \subset K^n$ eine Quadrik.*

Beweis. Es sei Q gegeben durch

$$\mathbf{x}^T \mathbf{A}\mathbf{x} = 0$$

und f durch

$$\mathbf{y} = \mathbf{S}\mathbf{x}.$$

Nach Satz 11.1 wird die Abbildung $f^{-1} : K^n \to K^n$ beschrieben durch

$$f^{-1}(y) = S^{-1}y - S^{-1}b.$$

Ist $T := S^{-1}$ und

$$\mathbf{T} := \left(\begin{array}{c|c} 1 & 0 \ \cdots \ 0 \\ \hline -S^{-1}(b) & S^{-1} \end{array}\right),$$

so ist $\mathbf{x} = \mathbf{T}\mathbf{y}$. Damit gilt

$$y = f(x) \in f(Q) \Leftrightarrow x \in Q \Leftrightarrow 0 = \mathbf{x}^T \mathbf{A}\mathbf{x} = \mathbf{y}^T(\mathbf{T}^T A\mathbf{T})\mathbf{y},$$

also

$$f(Q) = \{y \in K^n \,|\, \mathbf{y}^T \mathbf{B}\mathbf{y} = 0\} \text{ mit } \mathbf{B} = \mathbf{T}^T \mathbf{A}\mathbf{T}.$$

□

Wir wollen nun versuchen, die Affinität so zu wählen, dass die neue Gleichung so einfach wie möglich wird.

Beispiel 11.3 Es sei $Q \subset \mathbb{R}^2$ gegeben durch

$$x_1^2 + 4x_2^2 - 4x_1x_2 - 6x_1 + 14x_2 + 13 = 0.$$

Im ersten Schritt eliminieren wir den *gemischten* Term x_1x_2. Es ist

$$x_1^2 + 4x_2^2 - 4x_1x_2 = (x_1 - 2x_2)^2.$$

Also wird Q nach der Koordinatentransformation

$$\begin{aligned} z_1 &= x_1 - 2x_2, \\ z_2 &= x_2 \end{aligned}$$

gegeben durch

$$z_1^2 - 6z_1 + 2z_2 + 13 = 0.$$

Nun reduzieren wir die linearen Terme durch *quadratische Ergänzung*. Die Gleichung ist äquivalent zu

$$(z_1^2 - 6z_1) + 2z_2 + 13 = 0.$$

Durch quadratische Ergänzung der Klammer ergibt sich

$$(z_1^2 - 6z_1 + 9) + 2z_2 + 13 - 9 = 0$$

oder

$$(z_1 - 3)^2 + 2(z_2 + 2) = 0.$$

Nach der Transformation

$$\begin{aligned} y_1 &= z_1 - 3, \\ y_2 &= z_2 + 2 \end{aligned}$$

erhält man die Gleichung

$$y_1^2 + 2y_2 = 0.$$

Wir verallgemeinern nun die Methode aus diesem Beispiel. Wir betrachten den Spezialfall $K = \mathbb{R}$.

Es sei also die Quadrik $Q \subset \mathbb{R}^n$ beschrieben durch

$$\mathbf{A} = \left(\begin{array}{c|ccc} a_{00} & a_{01} & \cdots & a_{0n} \\ \hline a_{10} & & & \\ \vdots & & A & \\ a_{n0} & & & \end{array} \right).$$

Im ersten Schritt bringen wir die symmetrische Teilmatrix A auf Normalform. Nach Satz 6.1 gibt es eine Matrix $T_1 \in \mathrm{GL}(n; \mathbb{R})$ mit

$$T_1^T A T_1 = \begin{pmatrix} E_k & 0 & 0 \\ 0 & -E_{m-k} & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

wobei m der Rang von A und k der Index von A ist. Ist

$$\mathbf{T}_1 = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & T_1 & \\ 0 & & & \end{array} \right)$$

so wird

$$\mathbf{B}_1 = \mathbf{T}_1^T \mathbf{A} \mathbf{T}_1 = \left(\begin{array}{c|c|c|c} c_{00} & c_{01} & \cdots & c_{0n} \\ \hline c_{10} & E_k & 0 & 0 \\ \cline{2-4} \vdots & 0 & -E_{m-k} & 0 \\ \cline{2-4} c_{n0} & 0 & 0 & 0 \end{array}\right).$$

In den neuen Koordinaten lautet die Gleichung

$$z_1^2 + \cdots + z_k^2 - z_{k+1}^2 - \cdots - z_m^2 + 2(c_{01}z_1 + \cdots + c_{0n}z_n) + c_{00} = 0,$$

hat also keine gemischten Terme mehr.

Im zweiten Schritt reduzieren wir durch eine Translation die linearen Terme. Wir setzen

$$\mathbf{T}_2 = \left(\begin{array}{c|ccccccccc} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \hline c_{10} & 1 & & & & & & & & \\ \vdots & & \ddots & & & & & & & \\ c_{k0} & & & 1 & & & & 0 & & \\ -c_{k+1,0} & & & & 1 & & & & & \\ \vdots & & & & & \ddots & & & & \\ -c_{m0} & & & & & & 1 & & & \\ 0 & & & 0 & & & & 1 & & \\ \vdots & & & & & & & & \ddots & \\ 0 & & & & & & & & & 1 \end{array}\right).$$

Damit ergibt sich

$$\begin{aligned} \mathbf{B}_2 &:= \mathbf{T}_2^T \mathbf{B}_1 \mathbf{T}_2 = \mathbf{T}_2^T \mathbf{T}_1^T \mathbf{A} \mathbf{T}_1 \mathbf{T}_2 \\ &= \left(\begin{array}{c|cccccc|ccc} d_{00} & 0 & \cdots & 0 & 0 & \cdots & 0 & c_{0,m+1} & \cdots & c_{0n} \\ \hline 0 & +1 & & & & & 0 & & & \\ \vdots & & \ddots & & & & & & & \\ 0 & & & +1 & & & & & 0 & \\ 0 & & & & -1 & & & & & \\ \vdots & & & & & \ddots & & & & \\ 0 & 0 & & & & & -1 & & & \\ \cline{1-7} c_{m+1,0} & & & & & & & 0 & & \\ \vdots & & & 0 & & & & & \ddots & \\ c_{n0} & & & & & & & & & 0 \end{array}\right). \end{aligned}$$

Das bedeutet, dass Q in den neuen Koordinaten beschrieben werden kann durch die Gleichung

$$u_1^2 + \cdots + u_k^2 - u_{k+1}^2 - \cdots - u_m^2 + 2(c_{m+1,0}u_{m+1} + \cdots + c_{n0}u_n) + d_{00} = 0.$$

Nun unterscheiden wir drei Fälle:

(1) $d_{00} = c_{m+1,0} = \ldots = c_{n0} = 0$.

(2) $d_{00} \neq 0$, $c_{m+1,0} = \ldots = c_{n0} = 0$.

(3) $c_{r0} \neq 0$ für mindestens ein $r \in \{m+1, \ldots, n\}$.

Fall (1): Dann reduziert sich die obige Gleichung auf

$$u_1^2 + \cdots + u_k^2 - u_{k+1}^2 - \cdots - u_m^2$$

und wir sind fertig.

Fall (2): O.B.d.A. $d_{00} < 0$ (sonst multipliziere man die Gleichung mit -1 und ordne durch eine weitere Transformation $u_1, \ldots, u_m$ um). Setze

$$(u_1, \ldots, u_n) = \rho(y_1, \ldots, y_n) \text{ mit } \rho = \sqrt{-d_{00}},$$

d.h. betrachte

$$\mathbf{T}_3 := \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & \frac{1}{\rho}E_n \end{array}\right).$$

Dividiert man die entstehende Gleichung durch ρ^2, so erhält man

$$y_1^2 + \cdots + y_k^2 - y_{k+1}^2 - \cdots - y_m^2 = 1.$$

Fall (3): O.B.d.A. $r = m+1$ (sonst ordne man in einer weiteren Transformation $u_{m+1}, \ldots, u_n$ um). Setzt man

$$\begin{aligned} y_i &= u_i \text{ für } i \neq m+1, \\ 2y_{m+1} &= 2(c_{m+1,0}u_{m+1} + \cdots + c_{n0}u_n) + d_{00}, \end{aligned}$$

so erhält man als neue Gleichung für Q

$$y_1^2 + \cdots + y_k^2 - y_{k+1}^2 - \cdots - y_m^2 + 2y_{m+1} = 0.$$

Die Transformation kann man so beschreiben: Durch simultane Zeilen- und Spaltenumformungen der Matrix $\mathbf{B}_2$ beseitigt man mit Hilfe von $c_{m+1,0} = c_{0,m+1}$ nacheinander die Einträge

$$d_{00}, c_{m+2,0} = c_{0,m+2}, \ldots, c_{n0} = c_{0n}.$$

Insgesamt ergibt dies eine affine Transformation $\mathbf{T}_3$.

Damit haben wir folgendes Ergebnis bewiesen:

Theorem 11.1 (Affine Klassifikation von Quadriken) *Gegeben sei eine Quadrik*

$$Q = \{x \in \mathbb{R}^n \mid \mathbf{x}^T \mathbf{A} \mathbf{x} = 0\},$$

wobei $\mathbf{A}$ *eine symmetrische* $(n+1) \times (n+1)$*-Matrix ist. Es sei*

$$m := \operatorname{Rang} A, \quad \mathbf{m} := \operatorname{Rang} \mathbf{A}.$$

Dann gibt es eine Affinität $f : \mathbb{R}^n \to \mathbb{R}^n$*, so dass* $f(Q)$ *beschrieben wird durch eine der folgenden Gleichungen*

(1) $y_1^2 + \cdots + y_k^2 - y_{k+1}^2 - \cdots - y_m^2 = 0$, *falls* $m = \mathbf{m}$,

(2) $y_1^2 + \cdots + y_k^2 - y_{k+1}^2 - \cdots - y_m^2 = 1$, *falls* $m + 1 = \mathbf{m}$,

(3) $y_1^2 + \cdots + y_k^2 - y_{k+1}^2 - \cdots - y_m^2 + 2y_{m+1} = 0$, *falls* $m + 2 = \mathbf{m}$.

Speziell für $n = 2, 3$ erhalten wir die Tabellen 1 und 2.

Typ	m	$\mathbf{m}$	k	Gleichung	Beschreibung
(1)	0	0	0	$0 = 0$	Ebene $\mathbb{R}^2$
	1	1	1	$y_1^2 = 0$	(Doppel-)Gerade
	2	2	1	$y_1^2 - y_2^2 = 0$	Geradenpaar (mit Schnittpunkt)
	2	2	2	$y_1^2 + y_2^2 = 0$	Punkt
(2)	1	2	1	$y_1^2 = 1$	Zwei parallele Geraden
	2	3	1	$y_1^2 - y_2^2 = 1$	Hyperbel
	2	3	2	$y_1^2 + y_2^2 = 1$	Kreis
(3)	1	3	1	$y_1^2 + 2y_2 = 0$	Parabel

Tabelle 1: Normalformen von nicht leeren Quadriken im $\mathbb{R}^2$

Nun wollen wir statt allgemeinen Affinitäten nur solche Affinitäten zulassen, bei denen die lineare Abbildung orthogonal ist.

Definition Es sei V ein euklidischer Vektorraum. Eine Affinität $f : V \to V$ heißt *Kongruenz* (oder *Bewegung*), falls es eine orthogonale Abbildung $F : V \to V$ und ein $w_0 \in V$ gibt, so dass

$$f(v) = F(v) + w_0 \text{ für alle } v \in V.$$

Eine Kongruenz ist also die Hintereinanderschaltung einer orthogonalen Abbildung und einer Translation. Ist insbesondere $V = \mathbb{R}^n$ und

$$f(x) = Ax + b \text{ für alle } x \in \mathbb{R}^n$$

Typ	m	$\mathbf{m}$	k	Gleichung	Beschreibung
(1)	0	0	0	$0 = 0$	Raum $\mathbb{R}^3$
	1	1	1	$y_1^2 = 0$	(Doppel-)Ebene
	2	2	1	$y_1^2 - y_2^2 = 0$	Ebenenpaar (mit Schnittgerade)
	2	2	2	$y_1^2 + y_2^2 = 0$	Gerade
	3	3	2	$y_1^2 + y_2^2 - y_3^2 = 0$	Kreiskegel
	3	3	3	$y_1^2 + y_2^2 + y_3^2 = 0$	Punkt
(2)	1	2	1	$y_1^2 = 1$	Zwei parallele Ebenen
	2	3	1	$y_1^2 - y_2^2 = 1$	hyperbolischer Zylinder
	2	3	2	$y_1^2 + y_2^2 = 1$	Kreiszylinder
	3	4	1	$y_1^2 - y_2^2 - y_3^2 = 1$	zweischaliges Hyperboloid
	3	4	2	$y_1^2 + y_2^2 - y_3^2 = 1$	einschaliges Hyperboloid
	3	4	3	$y_1^2 + y_2^2 + y_3^2 = 1$	Kugel
(3)	1	3	1	$y_1^2 + 2y_2 = 0$	parabolischer Zylinder
	2	4	1	$y_1^2 - y_2^2 + 2y_3 = 0$	hyperbolisches Paraboloid
	2	4	2	$y_1^2 + y_2^2 + 2y_3 = 0$	elliptisches Paraboloid

Tabelle 2: Normalformen von nicht leeren Quadriken im $\mathbb{R}^3$

mit $A \in \mathrm{GL}(n;\mathbb{R})$, so ist f genau dann eine Kongruenz, wenn A orthogonal ist.

Nun modifizieren wir unsere Überlegungen, die zum Beweis von Theorem 11.1 führten, so, dass wir statt beliebiger Affinitäten nur Kongruenzen zulassen.

Gegeben sei wieder die Quadrik

$$Q = \{x \in \mathbb{R}^n \,|\, \mathbf{x}^T \mathbf{A} \mathbf{x} = 0\}.$$

Nach dem Satz über die Hauptachsentransformation gibt es eine orthogonale Matrix T_1 mit

$$T_1^T A T_1 = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}, \quad (\lambda_1, \ldots, \lambda_n \in \mathbb{R}).$$

Wir können $\lambda_1, \ldots, \lambda_m \neq 0$ und $\lambda_{m+1} = \ldots = \lambda_n = 0$ annehmen. Damit sind die gemischten Terme beseitigt. Mit der erweiterten Matrix $\mathbf{T}_1$ wie oben gilt dann

$$\mathbf{B}_1 := \mathbf{T}_1^T \mathbf{A} \mathbf{T}_1 = \left(\begin{array}{c|ccc|ccc} d_{00} & c_{01} & \cdots & c_{0m} & c_{0,m+1} & \cdots & c_{0n} \\ \hline c_{10} & \lambda_1 & & 0 & & & \\ \vdots & & \ddots & & & 0 & \\ c_{m0} & 0 & & \lambda_m & & & \\ \cline{1-4} c_{m+1,0} & & & & 0 & & \\ \vdots & & 0 & & & \ddots & \\ c_{n0} & & & & & & 0 \end{array}\right).$$

Wie oben führt eine Translation zu

$$\mathbf{B}_2 := \mathbf{T}_2^T \mathbf{B}_1 \mathbf{T}_2 = \left(\begin{array}{c|ccc|ccc} d_{00} & 0 & \cdots & 0 & c_{0,m+1} & \cdots & c_{0n} \\ \hline 0 & \lambda_1 & & 0 & & & \\ \vdots & & \ddots & & & 0 & \\ 0 & 0 & & \lambda_m & & & \\ \cline{1-4} c_{m+1,0} & & & & 0 & & \\ \vdots & & 0 & & & \ddots & \\ c_{n0} & & & & & & 0 \end{array}\right).$$

Nun unterscheiden wir wieder drei Fälle:

Fall (1): $d_{00} = c_{m+1,0} = \ldots = c_{n0} = 0$.
Hier sind wir bereits fertig: O.B.d.A. $\lambda_1, \ldots, \lambda_k > 0$, $\lambda_{k+1}, \ldots, \lambda_m < 0$. Wir setzen für $i = 1, \ldots, m$

$$\alpha_i := \frac{1}{\sqrt{|\lambda_i|}}.$$

Damit erhalten wir die Gleichung

$$\frac{y_1^2}{\alpha_1^2} + \cdots + \frac{y_k^2}{\alpha_k^2} - \frac{y_{k+1}^2}{\alpha_{k+1}^2} - \cdots - \frac{y_m^2}{\alpha_m^2} = 0.$$

Fall (2): $d_{00} \neq 0$, $c_{m+1,0} = \ldots = c_{n0} = 0$.
O.B.d.A. $d_{00} < 0$, $\lambda_1, \ldots, \lambda_k > 0$, $\lambda_{k+1}, \ldots, \lambda_m < 0$. Wir dividieren die Gleichung durch $|d_{00}|$ und setzen

$$\alpha_i := \frac{\sqrt{|d_{00}|}}{\sqrt{|\lambda_i|}}.$$

Dann erhalten wir die Gleichung

$$\frac{y_1^2}{\alpha_1^2} + \cdots + \frac{y_k^2}{\alpha_k^2} - \frac{y_{k+1}^2}{\alpha_{k+1}^2} - \cdots - \frac{y_m^2}{\alpha_m^2} = 1.$$

Fall (3): $c_{r0} \neq 0$ für mindestens ein $r \in \{m+1, \dots, n\}$.
Hier ist die Transformation, mit der wir $d_{00}, c_{m+2,0}, \dots, c_{n0}$ beseitigt haben, keine Kongruenz. Deswegen brauchen wir hier zusätzliche Überlegungen. Wir setzen

$$v := (c_{m+1,0}, \dots, c_{n0})^T \in \mathbb{R}^{n-m}, \quad v_1 = \frac{1}{\|v\|} v,$$

und konstruieren nach dem E. Schmidtschen Orthonormalisierungsverfahren eine Orthonormalbasis $v_1, \dots, v_{n-m}$ von $\mathbb{R}^{n-m}$. Dann beschreibt die Matrix

$$\mathbf{T}_3 := \left(\begin{array}{c|ccc|ccc} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \hline 0 & 1 & & 0 & & & \\ \vdots & & \ddots & & & 0 & \\ 0 & 0 & & 1 & & & \\ \hline & & & & & & \\ \mu v_1 & & 0 & & v_1 & \cdots & v_{n-m} \\ & & & & & & \end{array}\right), \quad \text{mit } \mu := \frac{-d_{00}}{2\|v\|},$$

eine Kongruenz und man rechnet leicht aus:

$$\mathbf{B}_3 := \mathbf{T}_3^T \mathbf{B}_2 \mathbf{T}_3 = \left(\begin{array}{c|ccc|c|ccc} 0 & 0 & \cdots & 0 & \|v\| & 0 & \cdots & 0 \\ \hline 0 & \lambda_1 & & 0 & 0 & & & \\ \vdots & & \ddots & & \vdots & & 0 & \\ 0 & 0 & & \lambda_m & 0 & & & \\ \hline \|v\| & 0 & \cdots & 0 & 0 & & & \\ \hline 0 & & & & & & & \\ \vdots & & 0 & & & & 0 & \\ 0 & & & & & & & \end{array}\right).$$

Wieder nehmen wir $\lambda_1, \dots, \lambda_k > 0$ und $\lambda_{k+1}, \dots, \lambda_m < 0$ an und setzen

$$\alpha_i := \frac{\sqrt{\|v\|}}{\sqrt{|\lambda_i|}}.$$

Damit ergibt sich die Gleichung

$$\frac{y_1^2}{\alpha_1^2} + \cdots + \frac{y_k^2}{\alpha_k^2} - \frac{y_{k+1}^2}{\alpha_{k+1}^2} - \cdots - \frac{y_m^2}{\alpha_m^2} + 2y_{m+1} = 0.$$

Damit haben wir bewiesen:

Theorem 11.2 (Metrische Klassifikation von Quadriken) *Gegeben sei eine Quadrik*

$$Q = \{x \in \mathbb{R}^n \,|\, \mathbf{x}^T \mathbf{A} \mathbf{x} = 0\},$$

wobei $\mathbf{A}$ *eine symmetrische* $(n+1)\times(n+1)$*-Matrix ist. Es sei*

$$m := \operatorname{Rang} A, \quad \mathbf{m} := \operatorname{Rang} \mathbf{A}.$$

Dann gibt es eine Kongruenz $f : \mathbb{R}^n \to \mathbb{R}^n$ *und* $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$*, so dass* $f(Q)$ *beschrieben wird durch eine der folgenden Gleichungen:*

$$(1) \quad \frac{y_1^2}{\alpha_1^2} + \cdots + \frac{y_k^2}{\alpha_k^2} - \frac{y_{k+1}^2}{\alpha_{k+1}^2} - \cdots - \frac{y_m^2}{\alpha_m^2} = 0, \ \textit{falls } m = \mathbf{m},$$

$$(2) \quad \frac{y_1^2}{\alpha_1^2} + \cdots + \frac{y_k^2}{\alpha_k^2} - \frac{y_{k+1}^2}{\alpha_{k+1}^2} - \cdots - \frac{y_m^2}{\alpha_m^2} = 1, \ \textit{falls } m+1 = \mathbf{m},$$

$$(3) \quad \frac{y_1^2}{\alpha_1^2} + \cdots + \frac{y_k^2}{\alpha_k^2} - \frac{y_{k+1}^2}{\alpha_{k+1}^2} - \cdots - \frac{y_m^2}{\alpha_m^2} + 2y_{m+1} = 0, \ \textit{falls } m+2 = \mathbf{m}.$$

Speziell für $n = 2$ erhalten wir folgende Tabelle:

Typ	m	$\mathbf{m}$	k	Gleichung	Beschreibung
(1)	0	0	0	$0 = 0$	Ebene $\mathbb{R}^2$
	1	1	1	$x^2 = 0$	(Doppel-)Gerade
	2	2	1	$x^2 - \alpha y^2 = 0$, $\alpha > 0$	Geradenpaar
	2	2	2	$x^2 + \alpha y^2 = 0$, $\alpha > 0$	Punkt
(2)	1	2	1	$\frac{x^2}{\alpha^2} = 1$, $\alpha > 0$	Zwei parallele Geraden
	2	3	1	$\frac{x^2}{\alpha^2} - \frac{y^2}{\beta^2} = 1$, $\alpha, \beta > 0$	Hyperbel
	2	3	2	$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} = 1$, $\alpha, \beta > 0$	Ellipse, Kreis
(3)	1	3	1	$x^2 + \alpha y = 0$, $\alpha \neq 0$	Parabel

Tabelle 3: Metrische Klassifikation von nicht leeren Quadriken im $\mathbb{R}^2$

Im Fall $n = 3$ illustrieren wir im Folgenden den Klassifikationssatz durch Bilder.

Fall (1): Der interessanteste Fall ist $m = 3$, $k = 2$. Die Gleichung

$$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} - \frac{z^2}{\gamma^2} = 0$$

beschreibt einen *elliptischen Kegel* (Bild 4).

Fall (2): In diesem Fall beschränken wir uns auf $m = 3$. Für $k = 3$ ergibt sich die Gleichung

$$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} + \frac{z^2}{\gamma^2} = 1,$$

die ein *Ellipsoid* beschreibt (Bild 5).

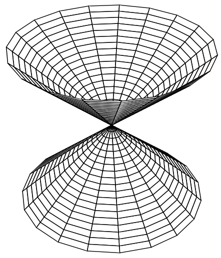

Abbildung 4: Elliptischer Kegel $\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} - \frac{z^2}{\gamma^2} = 0$

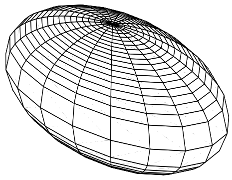

Abbildung 5: Ellipsoid $\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} + \frac{z^2}{\gamma^2} = 1$

Für $k = 2$ beschreibt

$$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} - \frac{z^2}{\gamma^2} = 1$$

ein *einschaliges Hyperboloid* (Bild 6).

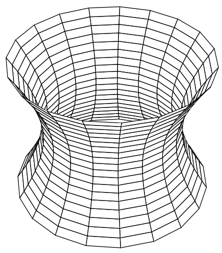

Abbildung 6: Einschaliges Hyperboloid $\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} - \frac{z^2}{\gamma^2} = 1$

Für $k = 1$ ergibt sich die Gleichung

$$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} - \frac{z^2}{\gamma^2} = -1$$

für ein *zweischaliges Hyperboloid* (Bild 7).

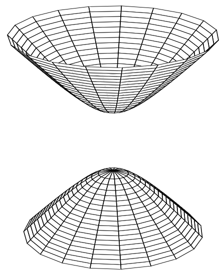

Abbildung 7: Zweischaliges Hyperboloid $\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} - \frac{z^2}{\gamma^2} = -1$

Fall (3): Hier beschränken wir uns auf den Fall $m = 2$. Für $k = 0$ erhalten wir die Gleichung

$$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} - 2z = 0,$$

die ein *elliptisches Paraboloid* beschreibt (Bild 8).

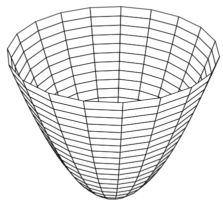

Abbildung 8: Elliptisches Paraboloid $\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} - 2z = 0$

Für $k = 1$ beschreibt

$$\frac{y^2}{\beta^2} - \frac{x^2}{\alpha^2} - 2z = 0$$

ein *hyperbolisches Paraboloid* (Bild 9).

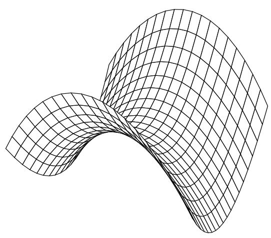

Abbildung 9: Hyperbolisches Paraboloid $\frac{y^2}{\beta^2} - \frac{x^2}{\alpha^2} - 2z = 0$

12 Der Dualraum

Ein grundlegender Begriff in der Linearen Algebra ist der Begriff des Dualraums, den wir nun einführen wollen. Im Folgenden sei K wieder ein beliebiger Körper.

Definition Es seien V, W K-Vektorräume. Die Menge aller linearen Abbildungen von V nach W bezeichnen wir mit

$$\mathrm{Hom}_K(V,W) := \{f : V \to W \mid f \text{ linear}\}.$$

Auf der Menge $\mathrm{Hom}_K(V,W)$ erklären wir eine Addition und skalare Multiplikation wie folgt ($f, g \in \mathrm{Hom}_K(V,W)$, $\lambda \in K$):

$$\begin{aligned} (f+g)(x) &= f(x) + g(x), \\ (\lambda f)(x) &= \lambda f(x) \quad \text{für alle } x \in V. \end{aligned}$$

Mit dieser Addition und skalaren Multiplikation wird $\mathrm{Hom}_K(V,W)$ zu einem K-Vektorraum. Nun betrachten wir den Spezialfall $W = K$.

Definition Der Vektorraum

$$V^* := \mathrm{Hom}_K(V,K) = \{\varphi : V \to K \mid \varphi \text{ linear}\}$$

heißt der *Dualraum* von V. Die Elemente von V^* heißen *Linearformen* auf V.

Beispiel 12.1 Wir betrachten eine lineare Gleichung

$$a_1x_1 + \cdots + a_nx_n = 0.$$

Setzt man $a = (a_1, \ldots, a_n)$ (Zeilenvektor!), so gilt

$$a \cdot x = \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = a_1x_1 + \cdots + a_nx_n.$$

Deswegen können wir a als eine lineare Abbildung

$$\begin{aligned} a : K^n &\longrightarrow K \\ x &\longmapsto a \cdot x \end{aligned}$$

auffassen, d.h. als ein Element von $(K^n)^*$.

Beispiel 12.2 Es sei $I = [a, b] \subset \mathbb{R}$ ein Intervall und $V = C[a, b]$ der Vektorraum der auf I stetigen Funktionen. Dann ist

$$\int_a^b : V \to \mathbb{R}, \quad f \mapsto \int_a^b f(x)\,dx,$$

eine Linearform auf V. Linearformen auf unendlich dimensionalen Vektorräumen nennt man auch *lineare Funktionale*. Mit ihnen beschäftigt sich die *Funktionalanalysis*.

Wir setzen von nun an voraus, dass V ein endlich dimensionaler K-Vektorraum ist und $n = \dim V > 0$. Es sei $B = \{v_1, \ldots, v_n\}$ eine Basis von V. Dann gibt es zu jedem $i = 1, \ldots, n$ genau eine lineare Abbildung

$$v_i^* : V \to K$$

mit

$$v_i^*(v_j) := \delta_{ij} := \begin{cases} 1 & \text{falls } j = i, \\ 0 & \text{falls } j \neq i \end{cases}$$

(δ_{ij} heißt das *Kroneckersymbol*).

Satz 12.1 *Die Menge* $B^* := \{v_1^*, \ldots, v_n^*\}$ *ist eine Basis von* V^*. *Insbesondere gilt* $\dim V^* = \dim V = n$.

Definition Man nennt B^* die zu der Basis B *duale Basis*.

Beweis.
(a) B^* ist linear unabhängig:

$$\sum_{i=1}^n \alpha_i v_i^* = 0 \Rightarrow \sum_{i=1}^n \alpha_i v_i^*(v_j) = 0 \Rightarrow \alpha_j = 0 \quad (j = 1, \ldots, n).$$

(b) B^* ist ein Erzeugendensystem: Es sei $v^* \in V^*$. Wir setzen

$$\alpha_i := v^*(v_i), \quad i = 1, \ldots, n.$$

Behauptung $v^* = \sum_{i=1}^n \alpha_i v_i^*$.

Beweis. Da eine lineare Abbildung durch die Bilder einer Basis eindeutig festgelegt ist, reicht es zu zeigen, dass die Bilder der Basisvektoren v_j, $j = 1, \ldots, n$, unter den beiden linearen Abbildungen übereinstimmen:

$$\left(\sum_{i=1}^n \alpha_i v_i^* \right)(v_j) = \sum_{i=1}^n \alpha_i v_i^*(v_j) = \alpha_j = v^*(v_j).$$

□

□

Korollar 12.1 *Zu jedem $v \in V$ mit $v \neq 0$ gibt es ein $\varphi \in V^*$ mit $\varphi(v) \neq 0$.*

Beweis. Dies folgt daraus, dass man jeden Vektor $v \neq 0$ zu einer Basis von V ergänzen kann. □

Korollar 12.2 *Zu jeder Basis $B = \{v_1, \ldots, v_n\}$ von V gibt es einen Isomorphismus*

$$\Psi_B : V \to V^* \text{ mit } \Psi_B(v_i) = v_i^*.$$

Warnung Dieser Isomorphismus hängt von der Wahl der Basis ab!

Beweis. Dies folgt aus der Tatsache, dass eine lineare Abbildung durch die Bilder der Vektoren einer Basis bestimmt ist. □

Beispiel 12.3 $V = K^n$, $B = \{e_1, \ldots, e_n\}$ kanonische Basis. Duale Basis:

$$B^* = \{e_1^*, \ldots, e_n^*\} \quad \text{(kanonische Basis von } V^*\text{)}.$$

Wir machen die Konvention, Vektoren als Spalten und Linearformen als Zeilen zu schreiben. Dann gilt

$$e_i^* = (0, \ldots, 0, 1, 0, \ldots, 0), \quad 1 \text{ an der } i\text{-ten Stelle}.$$

Definition Ist V ein K-Vektorraum und $U \subset V$ ein Unterraum, so heißt

$$U^0 := \{\varphi \in V^* \,|\, \varphi(u) = 0 \text{ für alle } u \in U\} \subset V^*$$

der zu U *orthogonale Raum* (oder der *Annullator* von U).

Der zu U orthogonale Raum U^0 ist ein Unterraum von V^*.

Warnung Der zu einem Unterraum $U \subset V$ orthogonale Raum U^0 liegt in V^* und ist nicht zu verwechseln mit dem orthogonalen Komplement $U^\perp$ von U, das nur in Räumen mit Skalarprodukt definiert ist und dann in V liegt!

Satz 12.2 *Für jeden Unterraum $U \subset V$ gilt*

$$\dim U^0 = \dim V - \dim U.$$

Genauer gilt: Ist $\{u_1, \ldots, u_k\}$ eine Basis von U und $B = \{u_1, \ldots, u_k, w_1, \ldots, w_\ell\}$ eine Basis von V, so bilden die Linearformen $\{w_1^*, \ldots, w_\ell^*\}$ aus B^* eine Basis von U^0.

Beweis.

(a) Lineare Unabhängigkeit: Da $w_1^*, \ldots, w_\ell^*$ Elemente der dualen Basis B^* sind, sind sie linear unabhängig.

(b) Zu zeigen: $U^0 = \mathrm{Span}\{w_1^*, \ldots, w_\ell^*\}$.

"$\supset$": klar, da $w_j^*(u_i) = 0$.

"$\subset$": Es sei $\varphi \in U^0$,

$$\varphi = \mu_1 u_1^* + \cdots + \mu_k u_k^* + \lambda_1 w_1^* + \cdots + \lambda_\ell w_\ell^*.$$

Wendet man diese Abbildung auf u_i $(i = 1, \ldots, k)$ an, so folgt

$$0 = \varphi(u_i) = \mu_i.$$

□

Nun wollen wir auch lineare Abbildungen dualisieren. Es sei $f : V \to W$ eine lineare Abbildung und $\psi \in W^*$. Dann betrachten wir dazu das Diagramm

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[dr, "\psi\circ f"'] & W \arrow[d, "\psi"] \\ & K \end{tikzcd}$$

Dann gilt $\psi \circ f \in V^*$. Damit können wir definieren:

Definition Die Abbildung

$$f^* : W^* \to V^*, \quad \psi \mapsto f^*(\psi) := \psi \circ f,$$

heißt die zu f *duale Abbildung.*

Bemerkung 12.1 Die Abbildung f^* ist linear:

$$\begin{aligned} f^*(\psi_1 + \psi_2) &= (\psi_1 + \psi_2) \circ f = \psi_1 \circ f + \psi_2 \circ f = f^*(\psi_1) + f^*(\psi_2), \\ f^*(\lambda\psi) &= (\lambda\psi) \circ f = \lambda\psi \circ f = \lambda f^*(\psi). \end{aligned}$$

Satz 12.3 *Es seien V und W K-Vektorräume mit Basen B und C und $f : V \to W$ eine lineare Abbildung. Dann gilt*

$$M_{B^*}^{C^*}(f^*) = \left(M_C^B(f)\right)^T,$$

d.h. die duale Abbildung wird bezüglich der dualen Basen durch die transponierte *Matrix beschrieben.*

Beweis. Es sei $B = \{v_1, \ldots, v_n\}$, $C = \{w_1, \ldots, w_m\}$, $M_C^B(f) = (a_{ij})$, $M_{B^*}^{C^*}(f^*) = (b_{ij})$. Dann gilt

$$\begin{aligned} f(v_j) &= \sum_{i=1}^{m} a_{ij} w_i, \quad \text{also } a_{ij} = w_i^*(f(v_j)) = f^*(w_i^*)(v_j), \\ f^*(w_i^*) &= \sum_{j=1}^{n} b_{ji} v_j^*, \quad \text{also } b_{ji} = f^*(w_i^*)(v_j). \end{aligned}$$

Also gilt $a_{ij} = b_{ji}$. □

Definition Es sei

$$\cdots \longrightarrow V_{i-1} \xrightarrow{f_{i-1}} V_i \xrightarrow{f_i} V_{i+1} \longrightarrow \cdots$$

eine (endliche oder unendliche) Sequenz von K-Vektorräumen und linearen Abbildungen. Die Sequenz heißt *exakt*, wenn für jedes i gilt:

$$\operatorname{Ker} f_i = \operatorname{Im} f_{i-1}.$$

Unter einer *kurzen exakten Sequenz* versteht man eine exakte Sequenz der Gestalt

$$0 \longrightarrow U \xrightarrow{f} V \xrightarrow{g} W \longrightarrow 0.$$

Es sei

$$0 \longrightarrow U \xrightarrow{f} V \xrightarrow{g} W \longrightarrow 0$$

eine kurze exakte Sequenz. Dann gilt

$$\begin{aligned} \text{Exaktheit an der Stelle } U &\Leftrightarrow f \text{ injektiv}, \\ \text{Exaktheit an der Stelle } V &\Leftrightarrow \operatorname{Im} f = \operatorname{Ker} g, \\ \text{Exaktheit an der Stelle } W &\Leftrightarrow g \text{ surjektiv}. \end{aligned}$$

Ist $f : V \to W$ eine lineare Abbildung zwischen K-Vektorräumen V und W, so hat man immer eine kurze exakte Sequenz

$$0 \longrightarrow \operatorname{Ker} f \hookrightarrow V \xrightarrow{f} \operatorname{Im} f \longrightarrow 0,$$

wobei $\operatorname{Ker} f \hookrightarrow V$ die Inklusionsabbildung ist. Dazu gehört eine duale kurze exakte Sequenz

$$0 \longleftarrow \operatorname{im} f^* \xleftarrow{f^*} W^* \hookleftarrow \operatorname{Ker} f^* \longleftarrow 0.$$

Der Zusammenhang zwischen diesen beiden kurzen exakten Sequenzen ist der Folgende:

Satz 12.4 *Für eine lineare Abbildung $f : V \to W$ zwischen endlich dimensionalen Vektorräumen gilt*

$$\operatorname{Im} f^* = (\operatorname{Ker} f)^0 \text{ und } \operatorname{Ker} f^* = (\operatorname{Im} f)^0.$$

Korollar 12.3 *Unter den obigen Voraussetzungen gilt*

$$\operatorname{Rang} f^* = \operatorname{Rang} f.$$

Beweis.

$$\begin{aligned} \operatorname{Rang} f^* &= \dim \operatorname{Im} f^* \\ &= \dim(\operatorname{Ker} f)^0 \quad \text{(nach Satz 12.4)} \\ &= \dim V - \dim \operatorname{Ker} f \quad \text{(nach Satz 12.2)} \\ &= \dim \operatorname{Im} f = \operatorname{Rang} f. \end{aligned}$$

□

Mit Hilfe von Korollar 12.3 erhalten wir damit einen neuen Beweis des folgenden Resultats.

Korollar 12.4 *Für jede Matrix $A \in \operatorname{Mat}(m \times n; K)$ gilt*

$$\textit{Zeilenrang}\, A = \textit{Spaltenrang}\, A.$$

Beweis von Satz 12.4. Wir zeigen $\operatorname{Im} f^* = (\operatorname{Ker} f)^0$. Der Beweis der zweiten Gleichung geht analog.

"$\subset$": Es sei $\varphi \in \operatorname{Im} f^*$. Dann ist $\varphi = f^*(\psi) = \psi \circ f$ für ein $\psi \in V^*$. Für $x \in \operatorname{Ker} f$ gilt dann

$$\varphi(x) = \psi(f(x)) = \psi(0) = 0,$$

also $\varphi \in (\operatorname{Ker} f)^0$.

"$\supset$": Es sei umgekehrt $\varphi \in V^*$ mit $\varphi|_{\operatorname{Ker} f} = 0$ gegeben. Wir müssen ein $\psi \in W^*$ mit $\varphi = \psi \circ f$ konstruieren. Dazu sei

$$\begin{aligned} B &= \{u_1, \dots, u_r, v_1, \dots, v_k\} \text{ Basis von } V, \\ B' &= \{w_1, \dots, w_r, w_{r+1}, \dots, w_m\} \text{ Basis von } W \end{aligned}$$

mit $\operatorname{Ker} f = \operatorname{Span}\{v_1, \dots, v_k\}$, $\operatorname{Im} f = \operatorname{Span}\{w_1, \dots, w_r\}$ und $f(u_i) = w_i$ für $i = 1, \dots, r$. Dann gibt es genau eine lineare Abbildung ψ mit

$$\psi(w_i) := \begin{cases} \varphi(u_i) & \text{für } i = 1, \dots, r, \\ 0 & \text{sonst.} \end{cases}$$

Nach Konstruktion von ψ ist $\varphi = \psi \circ f$. □

Es sei nun

$$\begin{aligned} \langle\ ,\ \rangle : \quad V \times V & \longrightarrow K \\ (v, w) & \longmapsto \langle v, w\rangle \end{aligned}$$

eine Bilinearform. Dann können wir die beiden folgenden Abbildungen betrachten:

$$\begin{aligned} \langle v, \cdot\rangle : V \to K, &\qquad w \mapsto \langle v, w\rangle, \\ \langle \cdot, w\rangle : V \to K, &\qquad v \mapsto \langle v, w\rangle. \end{aligned}$$

Damit erhalten wir Abbildungen

$$\begin{aligned} b_1 : V \to V^*, &\qquad v \mapsto \langle v, \cdot\rangle, \\ b_2 : V \to V^*, &\qquad w \mapsto \langle \cdot, w\rangle. \end{aligned}$$

Definition Eine Bilinearform $\langle\ ,\ \rangle : V \times V \to K$ heißt *nicht ausgeartet*, wenn die beiden Abbildungen b_1 und b_2 injektiv sind.

Beispiel 12.4 Ein Skalarprodukt $\langle\ ,\ \rangle$ auf einem $\mathbb{R}$-Vektorraum ist nicht ausgeartet. Denn

$$b_1(v) = 0 \Leftrightarrow \langle v, w\rangle = 0 \text{ für alle } w \in V \Leftrightarrow v = 0,$$

da $\langle\ ,\ \rangle$ positiv definit ist. Da $\langle\ ,\ \rangle$ symmetrisch ist, folgt, dass auch b_2 injektiv ist.

Daraus folgt unmittelbar:

Satz 12.5 *In einem endlich dimensionalen euklidischen Vektorraum V ist die Abbildung*

$$\Psi : V \to V^*, \quad v \mapsto \langle v, \cdot\rangle,$$

ein Isomorphismus.

Bemerkung 12.2 Im Gegensatz zu den Isomorphismen $\Psi_B : V \to V^*$ in Korollar 12.2, die von der Wahl der Basis abhängen, ist dieser Isomorphismus *kanonisch*, d.h. er hängt nicht von der Wahl einer Basis ab. Er existiert aber nur, wenn ein Skalarprodukt gegeben ist.

Satz 12.6 *Es sei V ein euklidischer Vektorraum und $\Psi : V \to V^*$ der kanonische Isomorphismus. Dann gilt:*

(i) *Für jeden Unterraum $U \subset V$ ist $\Psi(U^\perp) = U^0$.*

(ii) *Ist* $B = \{v_1, \ldots, v_n\}$ *eine Orthonormalbasis von* V *und* $B^* = \{v_1^*, \ldots, v_n^*\}$ *die duale Basis, so ist* $\Psi(v_i) = v_i^*$.

Beweis.
(i): Nach den Dimensionsformeln gilt

$$\dim U^\perp = \dim V - \dim U = \dim U^0.$$

Daher reicht es zu zeigen: $\Psi(U^\perp) \subset U^0$. Dies folgt aus

$$\Psi(v)(u) = \langle v, u\rangle = 0 \text{ für } v \in U^\perp \text{ und } u \in U.$$

(ii): Dies folgt aus

$$\Psi(v_i)(v_j) = \langle v_i, v_j\rangle = \delta_{ij} = v_i^*(v_j).$$

□

13 Multilineare Abbildungen

Literatur für diesen Abschnitt:

- M. Spivak: Calculus on Manifolds. W. A. Benjamin 1965.
- W. Greub: Multilinear Algebra, 2nd Edition. Springer-Verlag 1978.

Definition Es seien $V_1, \ldots, V_p, W$ K-Vektorräume. Eine Abbildung $\varphi : V_1 \times \cdots \times V_p \to W$ heißt *multilinear* oder genauer *p-linear*, wenn für jedes $i = 1, \ldots, p$ gilt

$$\begin{aligned} \varphi(v_1, \ldots, v_i + v_i', \ldots, v_p) &= \varphi(v_1, \ldots, v_i, \ldots, v_p) + \varphi(v_1, \ldots, v_i', \ldots, v_p), \\ \varphi(v_1, \ldots, \lambda v_i, \ldots, v_p) &= \lambda \varphi(v_1, \ldots, v_i, \ldots, v_p), \end{aligned}$$

für $v_i, v_i' \in V_i$, $\lambda \in K$. Gilt $W = K$, so heißt φ eine *multilineare* (oder *p-lineare*) *Funktion.*

Beispiel 13.1 (a) Der Fall $p = 1$ und $W = K$ ist der Spezialfall der Linearformen.

(b) Im Fall $p = 2$, $W = K$, $V_1 = V_2 = V$, erhalten wir gerade die am Anfang der Vorlesung betrachteten Bilinearformen.

(c) Es sei $\varphi : V \times V^* \to K$ die durch $(v, f) \mapsto f(v)$ definierte Abbildung. Dies ist eine bilineare Funktion.

Es sei $\mathrm{Hom}(V_1, \ldots, V_p; W)$ die Menge aller p-linearen Abbildungen $\varphi : V_1 \times \cdots \times V_p \to W$. Wir erklären eine Addition durch

$$(\varphi + \psi)(v_1, \ldots, v_p) = \varphi(v_1, \ldots, v_p) + \psi(v_1, \ldots, v_p)$$

und eine skalare Multiplikation mit $\lambda \in K$ durch

$$(\lambda\varphi)(v_1, \ldots, v_p) = \lambda\varphi(v_1, \ldots, v_p).$$

Damit wird $\mathrm{Hom}(V_1, \ldots, V_p; W)$ zu einem K-Vektorraum.

Definition Ist $V_1 = \ldots = V_p = V$ und $W = K$, so nennt man eine multilineare Funktion $\varphi : V \times \cdots \times V \to K$ auch eine *Multilinearform* oder einen p-*Tensor* auf V. Die Menge aller p-Tensoren auf V bezeichnen wir mit $T^p(V)$.

Nach den Bemerkungen vor der Definition ist $T^p(V)$ ein K-Vektorraum.

Beispiel 13.2 Es sei $V = K^n$ und det die Funktion

$$\begin{array}{rccc} \det : & \underbrace{K^n \times \cdots \times K^n}_{n} & \longrightarrow & K \\ & (a_1, \ldots, a_n) & \longmapsto & \det(a_1 \cdots a_n) \end{array},$$

wobei $(a_1 \cdots a_n)$ die $n \times n$-Matrix mit den Spalten $a_1, \ldots, a_n$ ist. Aus den Eigenschaften der Determinante folgt, dass det eine Multilinearform ist, also $\det \in T^n(K^n)$.

Definition Ist $\varphi \in T^p(V)$ und $\psi \in T^q(V)$, so definieren wir das *Tensorprodukt* $\varphi \otimes \psi \in T^{p+q}(V)$ durch

$$(\varphi \otimes \psi)(v_1, \ldots, v_p, v_{p+1}, \ldots, v_{p+q}) = \varphi(v_1, \ldots, v_p) \cdot \psi(v_{p+1}, \ldots, v_{p+q}).$$

Warnung Man beachte, dass hier die Reihenfolge wichtig ist, da $\varphi \otimes \psi$ und $\psi \otimes \varphi$ ganz unterschiedlich sind.

Satz 13.1 *Das Tensorprodukt hat die folgenden Eigenschaften:*

(i) $(\varphi \otimes \psi) \otimes \theta = \varphi \otimes (\psi \otimes \theta)$.

(ii) $(\varphi_1 + \varphi_2) \otimes \psi = \varphi_1 \otimes \theta + \varphi_2 \otimes \psi$.

(iii) $\varphi \otimes (\psi_1 + \psi_2) = \varphi \otimes \psi_1 + \varphi \otimes \psi_2$.

(vi) $(\lambda\varphi) \otimes \psi = \varphi \otimes (\lambda\psi) = \lambda(\varphi \otimes \psi)$.

Beweis. Der Nachweis dieser Eigenschaften ist einfach. Siehe Vorlesung. □

Der Vektorraum $T^1(V)$ ist gerade der Dualraum V^*. Mit dem Tensorprodukt können wir nun die anderen Vektorräume $T^p(V)$ durch $T^1(V)$ ausdrücken:

Satz 13.2 *Es sei $(v_1, \ldots, v_n)$ eine Basis von V und $(v_1^*, \ldots, v_n^*)$ die duale Basis. Dann ist die Menge aller p-fachen Tensorprodukte*

$$v_{i_1}^* \otimes \cdots \otimes v_{i_p}^*, \quad 1 \leq i_1, \ldots, i_p \leq n,$$

eine Basis von $T^p(V)$. Insbesondere hat $T^p(V)$ die Dimension n^p.

Beweis.
(a) Man beachte zunächst, dass

$$\begin{aligned} v_{i_1}^* \otimes \cdots \otimes v_{i_p}^*(v_{j_1}, \ldots, v_{j_p}) &= \delta_{i_1,j_1} \cdots \delta_{i_p,j_p} \\ &= \begin{cases} 1 & \text{falls } j_1 = i_1, \ldots, j_p = i_p, \\ 0 & \text{sonst.} \end{cases} \end{aligned}$$

Gilt

$$(w_1, \ldots, w_p) = \left(\sum_{j=1}^n a_{1j} v_j, \ldots, \sum_{j=1}^n a_{pj} v_j \right),$$

so ist

$$\begin{aligned} \varphi(w_1, \ldots, w_p) &= \sum_{j_1, \ldots, j_p = 1}^n a_{1,j_1} \cdots a_{p,j_p} \varphi(v_{j_1}, \ldots, v_{j_p}) \\ &= \sum_{i_1, \ldots, i_p = 1}^n \varphi(v_{i_1}, \ldots, v_{i_p}) \cdot v_{i_1}^* \otimes \cdots \otimes v_{i_p}^*(w_1, \ldots, w_p). \end{aligned}$$

Also gilt

$$\varphi = \sum_{i_1, \ldots, i_p = 1}^n \varphi(v_{i_1}, \ldots, v_{i_p}) \cdot v_{i_1}^* \otimes \cdots \otimes v_{i_p}^*.$$

Also bilden die $v_{i_1}^* \otimes \cdots \otimes v_{i_p}^*$ ein Erzeugendensystem von $T^p(V)$.

(b) Zum Beweis der linearen Unabhängigkeit nehmen wir an

$$\sum_{i_1, \ldots, i_p = 1}^n a_{i_1, \ldots, i_p} v_{i_1}^* \otimes \cdots \otimes v_{i_p}^* = 0.$$

Indem wir beide Seiten dieser Gleichung auf $(v_{j_1}, \ldots, v_{j_p})$ anwenden, erhalten wir

$$a_{j_1,\ldots,j_p} = 0.$$

Also sind die $v_{i_1}^* \otimes \cdots \otimes v_{i_p}^*$ linear unabhängig. □

Wie im Fall des Dualraums können wir einer linearen Abbildung $f : V \to W$ eine lineare Abbildung $f^* : T^p(W) \to T^p(V)$ zuordnen: Diese Abbildung ist definiert durch

$$f^*\varphi(v_1, \ldots, v_p) = \varphi(f(v_1), \ldots, f(v_p)) \text{ für } \varphi \in T^p(W), v_1, \ldots, v_p \in V.$$

Man kann leicht zeigen:

Satz 13.3 *Für eine lineare Abbildung $f : V \to W$ und $\varphi \in T^p(W)$, $\psi \in T^q(W)$ gilt*

$$f^*(\varphi \otimes \psi) = f^*\varphi \otimes f^*\psi.$$

14 Alternierende Multilinearformen

Für den Grundkörper K setzen wir in diesem Abschnitt voraus:

$$n_K := \underbrace{1 + \cdots + 1}_{n} \neq 0 \text{ für alle } n \geq 1.$$

Wir identifizieren dann n_K mit n.

Die Multilinearform $\det \in T^n(K^n)$ hat die folgende wichtige Eigenschaft:

$$\det(a_1, \ldots, a_i, \ldots, a_j, \ldots, a_n) = -\det(a_1, \ldots, a_j, \ldots, a_i, \ldots, a_n).$$

Solche Multilinearformen nennt man alternierend.

Definition Eine Multilinearform $\omega \in T^p(V)$ heißt *alternierend*, wenn für jede Permutation $\sigma \in \mathcal{S}_p$ gilt

$$\omega(v_1, \ldots, v_p) = \operatorname{sign}\sigma \cdot \omega(v_{\sigma(1)}, \ldots, v_{\sigma(p)}) \text{ für alle } v_1, \ldots, v_p \in V.$$

Die Menge aller alternierenden p-Tensoren bezeichnen wir mit $\bigwedge^p(V)$.

Offensichtlich ist $\bigwedge^p(V)$ ein Unterraum von $T^p(V)$.

Lemma 14.1 *Es sei $\omega \in T^p(V)$ eine Multilinearform. Dann sind die folgenden Bedingungen äquivalent:*

(i) *ω ist alternierend.*

(ii) $\omega(v_1, \ldots, v_p) = -\omega(v_{\tau(1)}, \ldots, v_{\tau(p)})$ *für jede Transposition* $\tau \in \mathcal{S}_p$.

(iii) *Ist* $v_i = v_j$ *für ein* $i \neq j$, *so ist* $\omega(v_1, \ldots, v_p) = 0$.

Beweis.
(i) $\Rightarrow$ (ii): Klar.
(ii) $\Rightarrow$ (i): Es sei $\sigma \in \mathcal{S}_p$. Dann gibt es nach I, Lemma 17.1, eine Darstellung

$$\sigma = \tau_1 \circ \cdots \circ \tau_m,$$

wobei die τ_i Transpositionen sind. Es gilt

$$m = \begin{cases} \text{gerade} & \text{falls } \operatorname{sign} \sigma = 1, \\ \text{ungerade} & \text{falls } \operatorname{sign} \sigma = -1. \end{cases}$$

Daraus folgt die Behauptung.
(ii) $\Rightarrow$ (iii): Es sei $v_i = v_j$ mit $i \neq j$ und τ sei die Transposition, die i und j vertauscht. Dann gilt

$$\omega(v_1, \ldots, v_p) = \omega(v_{\tau(1)}, \ldots, v_{\tau(p)}) = -\omega(v_1, \ldots, v_p).$$

Also folgt mit $1 + 1 \neq 0$

$$\omega(v_1, \ldots, v_p) = 0.$$

(iii) $\Rightarrow$ (ii): Es sei τ die Transposition, die i und j vertauscht ($i \neq j$). Dann gilt

$$\begin{aligned} 0 &= \omega(v_1, \ldots, \overset{i}{v_i + v_j}, \ldots, \overset{j}{v_i + v_j}, \ldots, v_p) \\ &= \omega(v_1, \ldots, v_i, \ldots, v_i, \ldots, v_p) + \omega(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_p) \\ &\quad + \omega(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_p) + \omega(v_1, \ldots, v_j, \ldots, v_j, \ldots, v_p). \end{aligned}$$

Daraus folgt

$$\omega(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_p) = -\omega(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_p).$$

□

Definition Es sei $\varphi \in T^p(V)$. Dann definieren wir

$$\operatorname{Alt}(\varphi)(v_1, \ldots, v_p) := \frac{1}{p!} \sum_{\sigma \in \mathcal{S}_p} \operatorname{sign} \sigma \cdot \varphi(v_{\sigma(1)}, \ldots, v_{\sigma(p)}).$$

Satz 14.1 (i) *Für* $\varphi \in T^p(V)$ *ist* $\operatorname{Alt}(\varphi) \in \bigwedge^p(V)$.

(ii) *Für* $\omega \in \bigwedge^p(V)$ *gilt* $\mathrm{Alt}(\omega) = \omega$.

(iii) *Für* $\varphi \in T^p(V)$ *gilt* $\mathrm{Alt}(\mathrm{Alt}(\varphi)) = \mathrm{Alt}(\varphi)$.

Beweis.
(i) Es sei $\tau \in \mathcal{S}_p$ beliebig. Dann gilt

$$\begin{aligned}
&\mathrm{Alt}(\varphi)(v_{\tau(1)}, \ldots, v_{\tau(p)}) \\
&= \frac{1}{p!} \sum_{\sigma \in \mathcal{S}_p} \operatorname{sign} \sigma \cdot \varphi(v_{\sigma(\tau(1))}, \ldots, v_{\sigma(\tau(p))}) \\
&= \operatorname{sign} \tau \cdot \frac{1}{p!} \sum_{\sigma \circ \tau \in \mathcal{S}_p} \operatorname{sign}(\sigma \circ \tau) \cdot \varphi(v_{(\sigma\circ\tau)(1)}, \ldots, v_{(\sigma\circ\tau)(p)}) \\
&= \operatorname{sign} \tau \cdot \mathrm{Alt}(\varphi)(v_1, \ldots, v_p).
\end{aligned}$$

(ii) Es gilt

$$\begin{aligned}
\mathrm{Alt}(\omega)(v_1, \ldots, v_p) &= \frac{1}{p!} \sum_{\sigma \in \mathcal{S}_p} \operatorname{sign} \sigma \cdot \omega(v_{\sigma(1)}, \ldots, v_{\sigma(p)}) \\
&= \frac{1}{p!} \sum_{\sigma \in \mathcal{S}_p} (\operatorname{sign} \sigma)^2 \omega(v_1, \ldots, v_p) \\
&= \omega(v_1, \ldots, v_p).
\end{aligned}$$

(iii) folgt aus (i) und (ii). □

Wir möchten eine Basis von $\bigwedge^p(V)$ bestimmen. Dazu bemerken wir, dass mit $\omega \in \bigwedge^p(V)$ und $\eta \in \bigwedge^q(V)$ das Tensorprodukt $\omega \otimes \eta$ im Allgemeinen nicht in $\bigwedge^{p+q}(V)$ liegt. Daher definieren wir ein neues Produkt:

Definition Ist $\omega \in \bigwedge^p(V)$ und $\eta \in \bigwedge^q(V)$, so definieren wir das *Dachprodukt* $\omega \wedge \eta \in \bigwedge^{p+q}(V)$ durch

$$\omega \wedge \eta := \frac{(p+q)!}{p!q!} \mathrm{Alt}(\omega \otimes \eta).$$

Satz 14.2 *Das Dachprodukt hat folgende Eigenschaften:*

(i) $(\omega_1 + \omega_2) \wedge \eta = \omega_1 \wedge \eta + \omega_2 \wedge \eta$.

(ii) $\omega \wedge (\eta_1 + \eta_2) = \omega \wedge \eta_1 + \omega \wedge \eta_2$.

(iii) $(\lambda\omega) \wedge \eta = \omega \wedge (\lambda\eta) = \lambda(\omega \wedge \eta)$.

(iv) $\omega \wedge \eta = (-1)^{pq} \eta \wedge \omega$.

(v) $f^*(\omega \wedge \eta) = f^*(\omega) \wedge f^*(\eta)$.

Beweis. Diese Eigenschaften sind leicht nachzuweisen. □

Mehr Arbeit erfordert der Nachweis der Gleichung $(\omega\wedge\eta)\wedge\theta = \omega\wedge(\eta\wedge\theta)$. Dazu brauchen wir zwei Hilfssätze:

Lemma 14.2 *Für* $\varphi \in T^p(V)$ *mit* $\mathrm{Alt}(\varphi) = 0$ *und* $\psi \in T^q(V)$ *gilt*

$$\mathrm{Alt}(\varphi \otimes \psi) = \mathrm{Alt}(\psi \otimes \varphi) = 0.$$

Beweis. Es gilt

$$\begin{aligned}
&(p+q)!\mathrm{Alt}(\varphi \otimes \psi)(v_1, \ldots, v_{p+q}) \\
&= \sum_{\sigma \in \mathcal{S}_{p+q}} \operatorname{sign} \sigma \cdot \varphi(v_{\sigma(1)}, \ldots, v_{\sigma(p)}) \cdot \psi(v_{\sigma(p+1)}, \ldots, v_{\sigma(p+q)}).
\end{aligned}$$

Nun betrachten wir die Untergruppe $G \subset \mathcal{S}_{p+q}$, die aus allen Permutationen σ besteht, die $p+1, \ldots, p+q$ fest lassen. Dann gilt

$$\begin{aligned}
&\sum_{\sigma \in G} \operatorname{sign} \sigma \cdot \varphi(v_{\sigma(1)}, \ldots, v_{\sigma(p)}) \cdot \psi(v_{\sigma(p+1)}, \ldots, v_{\sigma(p+q)}) \\
&= \left(\sum_{\sigma' \in \mathcal{S}_p} \operatorname{sign} \sigma' \cdot \varphi(v_{\sigma'(1)}, \ldots, v_{\sigma'(p)}) \right) \cdot \psi(v_{p+1}, \ldots, v_{p+q}) \\
&= p!\mathrm{Alt}(\varphi) \cdot \psi(v_{p+1}, \ldots, v_{p+q}) = 0
\end{aligned}$$

Nun sei $\sigma_0 \notin G$. Es sei

$$G\sigma_0 := \{\sigma \circ \sigma_0 \,|\, \sigma \in G\}.$$

Wir setzen außerdem

$$(v_{\sigma_0(1)}, \ldots, v_{\sigma_0(p+q)}) = (w_1, \ldots, w_{p+q}).$$

Dann gilt

$$\begin{aligned}
&\sum_{\sigma \in G\sigma_0} \operatorname{sign} \sigma \cdot \varphi(v_{\sigma(1)}, \ldots, v_{\sigma(p)}) \cdot \psi(v_{\sigma(p+1)}, \ldots, v_{\sigma(p+q)}) \\
&= \left(\operatorname{sign} \sigma_0 \cdot \sum_{\sigma' \in G} \operatorname{sign} \sigma' \cdot \varphi(w_{\sigma'(1)}, \ldots, w_{\sigma'(p)}) \right) \cdot \psi(w_{p+1}, \ldots, w_{p+q}) \\
&= 0.
\end{aligned}$$

Man beachte, dass $G \cap G\sigma_0 = \emptyset$ gilt. Denn angenommen $\sigma \in G \cap G\sigma_0$. Dann gilt $\sigma = \sigma' \circ \sigma_0$ für ein $\sigma' \in G$, also $\sigma_0 = (\sigma')^{-1} \circ \sigma \in G$, ein Widerspruch. Wenn wir auf diese Weise fortfahren, können wir $\mathcal{S}_{p+q}$ so in disjunkte Teilmengen zerlegen, dass die Summe über jede dieser Teilmengen jeweils 0 ergibt. Also ergibt die Summe über ganz $\mathcal{S}_{p+q}$ Null.

Die andere Gleichung $\mathrm{Alt}(\psi \otimes \varphi) = 0$ wird analog bewiesen. □

Lemma 14.3 *Für $\omega \in \bigwedge^p(V)$, $\eta \in \bigwedge^q(V)$ und $\theta \in \bigwedge^r(V)$ gilt*

$$\mathrm{Alt}(\mathrm{Alt}(\omega \otimes \eta) \otimes \theta) = \mathrm{Alt}(\omega \otimes \eta \otimes \theta) = \mathrm{Alt}(\omega \otimes \mathrm{Alt}(\eta \otimes \theta)).$$

Beweis. Es gilt

$$\mathrm{Alt}(\mathrm{Alt}(\eta \otimes \theta) - \eta \otimes \theta) = \mathrm{Alt}(\eta \otimes \theta) - \mathrm{Alt}(\eta \otimes \theta) = 0.$$

Aus Lemma 14.2 folgt damit

$$\begin{aligned} 0 &= \mathrm{Alt}(\omega \otimes [\mathrm{Alt}(\eta \otimes \theta) - \eta \otimes \theta]) \\ &= \mathrm{Alt}(\omega \otimes \mathrm{Alt}(\eta \otimes \theta)) - \mathrm{Alt}(\omega \otimes \eta \otimes \theta). \end{aligned}$$

Die andere Gleichung wird analog bewiesen. □

Satz 14.3 *Für $\omega \in \bigwedge^p(V)$, $\eta \in \bigwedge^q(V)$ und $\theta \in \bigwedge^r(V)$ gilt*

$$(\omega \wedge \eta) \wedge \theta = \omega \wedge (\eta \wedge \theta) = \frac{(p+q+r)!}{p!q!r!} \mathrm{Alt}(\omega \otimes \eta \otimes \theta).$$

Beweis.

$$\begin{aligned} (\omega \wedge \eta) \wedge \theta &= \frac{((p+q)+r)!}{(p+q)!r!} \mathrm{Alt}((\omega \wedge \eta) \otimes \theta) \\ &= \frac{(p+q+r)!}{(p+q)!r!} \mathrm{Alt}\left(\frac{(p+q)!}{p!q!} \mathrm{Alt}(\omega \otimes \eta) \otimes \theta \right) \\ &= \frac{(p+q+r)!}{p!q!r!} \mathrm{Alt}(\omega \otimes \eta \otimes \theta) \quad \text{(nach Lemma 14.3).} \end{aligned}$$

□

Satz 14.4 *Es sei $(v_1, \ldots, v_n)$ eine Basis von V und $(v_1^*, \ldots, v_n^*)$ die duale Basis. Dann ist die Menge aller p-fachen Dachprodukte*

$$v_{i_1}^* \wedge \cdots \wedge v_{i_p}^*, \quad 1 \leq i_1 < \ldots < i_p \leq n,$$

eine Basis von $\bigwedge^p(V)$. Insbesondere hat $\bigwedge^p(V)$ die Dimension

$$\binom{n}{p} = \frac{n!}{p!(n-p)!}.$$

Beweis.

(a) Es sei $\omega \in \bigwedge^p(V) \subset T^p(V)$. Nach Satz 13.2 können wir schreiben:

$$\omega = \sum_{i_1,\ldots,i_p=1}^{n} a_{i_1,\ldots,i_p} v_{i_1}^* \otimes \cdots \otimes v_{i_p}^*.$$

Es folgt

$$\omega = \mathrm{Alt}(\omega) = \sum_{i_1,\ldots,i_p=1}^{n} a_{i_1,\ldots,i_p} \mathrm{Alt}(v_{i_1}^* \otimes \cdots \otimes v_{i_p}^*).$$

Nach Satz 14.3 gilt aber

$$\mathrm{Alt}(v_{i_1}^* \otimes \cdots \otimes v_{i_p}^*) = \text{Konstante} \cdot v_{i_1}^* \wedge \cdots \wedge v_{i_p}^*.$$

Also spannen die Elemente $v_{i_1}^* \wedge \cdots \wedge v_{i_p}^*$ den Raum $\bigwedge^p(V)$ auf.

(b) Die lineare Unabhängigkeit wird wie im Beweis von Satz 13.2 bewiesen. □

Hat V die Dimension n, so folgt aus Satz 14.4, dass $\bigwedge^n(V)$ die Dimension 1 hat. Das bedeutet, dass alle alternierenden n-Tensoren auf V Vielfache eines von Null verschiedenen n-Tensors sind. Für $V = K^n$ ist $\det \in \bigwedge^n(K^n)$ ein solcher Tensor. Er ist dadurch ausgezeichnet, dass $\det(e_1, \ldots, e_n) = 1$ gilt. Deswegen erhalten wir die folgende Charakterisierung der Determinante:

Korollar 14.1 *Die Determinante* det *ist der eindeutig bestimmte alternierende n-Tensor mit*

$$\det(e_1, \ldots, e_n) = 1.$$

Satz 14.5 *Es sei V ein K-Vektorraum, $(v_1, \ldots, v_n)$ eine Basis von V und $\omega \in \bigwedge^n(V)$. Ist $w_j = \sum_{i=1}^n a_{ij} v_i$, $j = 1, \ldots, n$, so gilt*

$$\omega(w_1, \ldots, w_n) = \det(a_{ij}) \cdot \omega(v_1, \ldots, v_n).$$

Beweis. Wir definieren $\eta \in T^n(K^n)$ durch

$$\eta((a_{11}, \ldots, a_{n1})^T, \ldots, (a_{1n}, \ldots, a_{nn})^T) = \omega(\sum a_{i1} v_i, \ldots, \sum a_{in} v_i).$$

Dann ist $\eta \in \bigwedge^n(K^n)$, also $\eta = \lambda \cdot \det$ für ein $\lambda \in K$. Es gilt

$$\lambda = \lambda \det(e_1, \ldots, e_n) = \eta(e_1, \ldots, e_n) = \omega(v_1, \ldots, v_n).$$

□

Es sei nun $K = \mathbb{R}$. Satz 14.5 zeigt, dass ein von Null verschiedener alternierender n-Tensor $\omega \in \bigwedge^n(V)$ die (geordneten) Basen von V in zwei disjunkte Klassen einteilt: eine mit $\omega(v_1, \ldots, v_n) > 0$ und eine mit $\omega(v_1, \ldots, v_n) < 0$. Wenn $(v_1, \ldots, v_n)$ und $(w_1, \ldots, w_n)$ zwei Basen von V sind und $A = (a_{ij})$ die durch $w_j = \sum a_{ij} v_i$ definierte Matrix des Basiswechsels, dann sind $(v_1, \ldots, v_n)$ und $(w_1, \ldots, w_n)$ genau dann in der gleichen Klasse, wenn $\det A > 0$ gilt. Also ist die Klasseneinteilung unabhängig von ω. Damit können wir definieren:

Definition Eine *Orientierung* eines reellen Vektorraums V ist eine Klasse von Basen, für die gilt: Für einen von Null verschiedenen alternierenden n-Tensor $\omega \in \bigwedge^n(V)$ ist $\omega(v_1, \ldots, v_n) > 0$ für alle Basen aus dieser Klasse oder $\omega(v_1, \ldots, v_n) < 0$ für alle Basen aus dieser Klasse.

Beispiel 14.1 Die *Standardorientierung* des $\mathbb{R}^n$ ist die Orientierung, zu der die Standardbasis $(e_1, \ldots, e_n)$ gehört. Hier kommt es auf die Reihenfolge an: Vertauschen wir zwei dieser Basiselemente, so gehört die neue Basis zu der anderen Orientierung.

Ist nun V ein euklidischer Vektorraum, so gilt für die Transformationsmatrix A, die eine Orthonormalbasis in eine andere transformiert, $\det A = \pm 1$. Damit können wir definieren:

Definition Es sei V ein euklidischer Vektorraum und eine Orientierung von V gewählt. Das *Volumenelement* von V ist das eindeutig bestimmte Element $\omega \in \bigwedge^n(V)$ mit $\omega(v_1, \ldots, v_n) = 1$ für jede Orthonormalbasis $(v_1, \ldots, v_n)$ aus der Orientierung von V.

Beispiel 14.2 Für $V = \mathbb{R}^n$ mit dem gewöhnlichen euklidischen Skalarprodukt und der Standardorientierung ist det das Volumenelement und

$$|\det(v_1, \ldots, v_n)|$$

ist das Volumen des von den Vektoren $v_1, \ldots, v_n$ aufgespannten Parallelotops.

15 Symmetrische Multilinearformen

Analog zu alternierenden Multilinearformen kann man auch symmetrische Multilinearformen betrachten. Für den Grundkörper K setzen wir weiterhin $n_K \neq 0$ für alle $n \geq 1$ voraus.

Definition Eine Multilinearform $\varphi \in T^p(V)$ heißt *symmetrisch*, wenn für jede Permutation $\sigma \in \mathcal{S}_p$ gilt

$$\varphi(v_1, \ldots, v_p) = \varphi(v_{\sigma(1)}, \ldots, v_{\sigma(p)}) \text{ für alle } v_1, \ldots, v_p \in V.$$

Die Menge aller symmetrischen p-Tensoren bezeichnen wir mit $S^p(V)$.

Definition Es sei $\varphi \in T^p(V)$. Dann definieren wir

$$\mathrm{Sym}(\varphi)(v_1, \ldots, v_p) := \frac{1}{p!} \sum_{\sigma \in \mathcal{S}_p} \varphi(v_{\sigma(1)}, \ldots, v_{\sigma(p)}).$$

Satz 15.1 (i) *Für* $\varphi \in T^p(V)$ *ist* $\mathrm{Sym}(\varphi) \in S^p(V)$.

(ii) *Für* $\varphi \in S^p(V)$ *gilt* $\mathrm{Sym}(\varphi) = \varphi$.

(iii) *Für* $\varphi \in T^p(V)$ *gilt* $\mathrm{Sym}(\mathrm{Sym}(\varphi)) = \mathrm{Sym}(\varphi)$.

Beweis. Der Beweis ist analog zum Beweis von Satz 14.1. □

Definition Ist $\varphi \in S^p(V)$ und $\psi \in S^q(V)$, so definieren wir das *symmetrische Produkt* $\varphi \vee \psi \in S^{p+q}(V)$ durch

$$\varphi \vee \psi := \frac{(p+q)!}{p!q!} \mathrm{Sym}(\varphi \otimes \psi).$$

Wie im Fall der alternierenden Multilinearformen beweist man:

Satz 15.2 *Das symmetrische Produkt hat folgende Eigenschaften:*

(i) $(\varphi_1 + \varphi_2) \vee \psi = \varphi_1 \vee \psi + \varphi_2 \vee \psi$.

(ii) $\varphi \vee (\psi_1 + \psi_2) = \varphi \vee \psi_1 + \varphi \vee \psi_2$.

(iii) $(\lambda\varphi) \vee \psi = \varphi \vee (\lambda\psi) = \lambda(\varphi \vee \psi)$.

(iv) $\varphi \vee \psi = \psi \vee \varphi$.

(v) $(\varphi \vee \psi) \vee \theta = \varphi \vee (\psi \vee \theta)$.

(vi) $f^*(\varphi \vee \psi) = f^*(\varphi) \vee f^*(\psi)$.

Satz 15.3 *Es sei $(v_1, \dots, v_n)$ eine Basis von V und $(v_1^*, \dots, v_n^*)$ die duale Basis. Dann ist die Menge aller p-fachen symmetrischen Produkte*

$$v_{i_1}^* \vee \cdots \vee v_{i_p}^*, \quad 1 \leq i_1 \leq \ldots \leq i_p \leq n,$$

eine Basis von $S^p(V)$. Insbesondere hat $S^p(V)$ die Dimension

$$\binom{n+p-1}{p}.$$

Definition Ein Polynom

$$P(x_1, \dots, x_n) = \sum_{\text{endlich}} a_{\nu_1, \dots, \nu_n} x_1^{\nu_1} \cdots x_n^{\nu_n}$$

heißt *homogen vom Grad p*, wenn die Summe über alle n-Tupel $(\nu_1, \dots, \nu_n)$ mit $\sum_{i=1}^n \nu_i = p$ läuft. Es sei $K^p[x_1, \dots, x_n]$ der Vektorraum der homogenen Polynome in den Variablen $x_1, \dots, x_n$ vom Grad p.

Satz 15.4 *Es sei V ein K-Vektorraum der Dimension n. Dann gibt es einen Isomorphismus*

$$\Psi : S^p(V) \to K^p[x_1, \dots, x_n].$$

Beweis. Es sei $(v_1, \dots, v_n)$ eine Basis von V und $(v_1^*, \dots, v_n^*)$ die duale Basis. Dann definieren wir

$$\Psi : S^p(V) \to K^p[x_1, \dots, x_n]$$

durch

$$v_{i_1}^* \vee \cdots \vee v_{i_p}^* \mapsto x_{i_1} \cdots x_{i_p}, \quad 1 \leq i_1 \leq \ldots \leq i_p \leq n.$$

Da die $v_{i_1}^* \vee \cdots \vee v_{i_p}^*$ nach Satz 15.3 eine Basis von $S^p(V)$ und die Monome $x_{i_1} \cdots x_{i_p}$ eine Basis von $K^p[x_1, \dots, x_n]$ bilden, folgt, dass sich Ψ zu einem Isomorphismus zwischen $S^p(V)$ und $K^p[x_1, \dots, x_n]$ erweitern lässt. □

16 Der Quotientenraum

Wir wollen nun den Begriff des Quotientenraums einführen. Dazu betrachten wir Äquivalenzrelationen.

Es sei X eine Menge.

Definition Eine *Äquivalenzrelation* auf X ist eine Teilmenge $R \subset X \times X$ mit folgenden Eigenschaften:

(R) $(x,x) \in R$ für alle $x \in X$ (*reflexiv*).

(S) $(x,y) \in R \Rightarrow (y,x) \in R$ für alle $x,y \in X$ (*symmetrisch*).

(T) $(x,y) \in R, (y,z) \in R \Rightarrow (x,z) \in R$ für alle $x,y,z \in X$ (*transitiv*).

Notation $x \sim y :\Leftrightarrow (x,y) \in R$.

Beispiel 16.1 Es sei $X = \mathrm{Mat}(n,n;K)$ und

$$R := \{(A;B) \in X \times X \mid \exists T \in \mathrm{GL}(n;K) \text{ mit } A = T^{-1}BT\} \subset X \times X.$$

Dann gilt

$$A \sim B \Leftrightarrow A \text{ ist ähnlich zu } B.$$

Definition Es sei V ein K-Vektorraum und $U \subset V$ ein Unterraum. Wir definieren

$$R := \{(u,v) \in V \times V \mid u - v \in U\}.$$

Mit anderen Worten

$$u \sim v :\Leftrightarrow u - v \in U.$$

Lemma 16.1 *Die obige Relation R ist eine Äquivalenzrelation.*

Beweis.

(R) $(u,u) \in R$, da $u - u = 0 \in U$.

(S) $(u,v) \in R \Rightarrow u - v \in U \Rightarrow v - u \in U \Rightarrow (v,u) \in R$.

(T) $(u,v) \in R, (v,w) \in R \Rightarrow u - v \in U, v - w \in U \Rightarrow u - w \in U \Rightarrow (u,w) \in R$.

□

Definition Es sei $x \in X$. Die *Äquivalenzklasse von x*, in Zeichen $[x]$, ist definiert als

$$[x] := \{y \in X \mid x \sim y\} \subset X.$$

Jedes Element $y \in [x]$ heißt *Repräsentant* der Äquivalenzklasse $[x]$.

Lemma 16.2 (i) $x \in [x]$.

(ii) $[x] \cap [y] \neq \emptyset \Leftrightarrow x \sim y$.

(iii) $x \sim y \Leftrightarrow [x] = [y]$.

Beweis.

(i) folgt aus der Reflexivität.

(ii): Es sei $[x] \cap [y] \neq \emptyset$. Dann existiert ein $z \in [x] \cap [y]$. Dann gilt $z \sim x$ und $z \sim y$. Aus der Transitivität folgt $x \sim y$.

Ist umgekehrt $x \sim y$, so folgt $x \in [x] \cap [y]$, also $[x] \cap [y] \neq \emptyset$.

(iii): Es sei $x \sim y$. Ist $z \in [x]$, so gilt $z \sim x$. Aus der Transitivität folgt $z \sim y$, also $z \in [y]$. Also gilt $[x] \subset [y]$. Analog zeigt man $[y] \subset [x]$.

Die umgekehrte Richtung ist klar. □

Korollar 16.1 *Die Menge X ist die disjunkte Vereinigung der Äquivalenzklassen.*

Definition Die Menge der Äquivalenzklassen heißt die *Quotientenmenge* und wird mit $X/\sim$ bezeichnet. Man hat eine *kanonische Projektion*

$$\begin{array}{rccc} \pi: & X & \longrightarrow & X/\sim \\ & x & \longmapsto & [x] \end{array}.$$

Wir kehren nun zu dem Beispiel eines Unterraums $U \subset V$ in einem K-Vektorraum V zurück.

Definition Die Menge $a + U := \{a + u \,|\, u \in U\}$ heißt die *Nebenklasse von* $a \in V$ in V.

Lemma 16.3 *Es gilt $[a] = a + U$ bezüglich der Äquivalenzrelation $u \sim v \Leftrightarrow u - v \in U$.*

Beweis. $x \in [a] \Leftrightarrow x \sim a \Leftrightarrow x - a \in U \Leftrightarrow x \in a + U$. □

Definition Der *Quotientenraum* V/U ist die Menge

$$V/U := \{a + U \,|\, a \in V\}.$$

Definition Wir definieren eine Addition auf V/U durch

$$(a + U) + (b + U) := (a + b) + U \quad (a, b \in V)$$

und eine skalare Multiplikation durch

$$\lambda(a + U) := (\lambda a) + U \quad (\lambda \in K, a \in V).$$

Lemma 16.4 *Diese Verknüpfungen sind wohldefiniert (d.h. unabhängig von der Wahl der Repräsentanten a und b) und machen V/U zu einem K-Vektorraum.*

Beweis. Wir zeigen zunächst die *Unabhängigkeit von der Wahl der Repräsentanten.*

(a) Es sei $a \sim a', b \sim b'$. Es ist zu zeigen: $(a+b)+U = (a'+b')+U$, d.h.

$$a + b \sim a' + b'.$$

Es gilt

$$\begin{aligned} a \sim a', b \sim b' \quad &\Rightarrow \quad a - a' \in U, b - b' \in U \Rightarrow (a-a') + (b-b') \in U \\ &\Rightarrow \quad (a+b) - (a'+b') \in U \Rightarrow a + b \sim a' + b'. \end{aligned}$$

(b) Die entsprechende Aussage für die skalare Multiplikation zeigt man analog.

Das Nachrechnen der Vektorraumaxiome für V/U ist einfach, siehe Vorlesung. Was ist das neutrale Element von V/U? Was ist das additive Inverse von $a + U$? □

Der Quotientenraum V/U lässt sich folgendermaßen geometrisch deuten: Es sei W ein Komplement von U in V, d.h.

$$V = U \oplus W.$$

Es sei $a + U$ eine Nebenklasse. Dann gibt es eine eindeutig bestimmte Darstellung

$$a = a_W + a_U, \quad a_W \in W,\ a_U \in U.$$

Es gilt

$$[a] = a + U = a_W + U.$$

Das Element a_W hängt nur von der Äquivalenzklasse ab und ist durch diese eindeutig bestimmt. Wir können also die Elemente von V/U mit den Elementen von W identifizieren. Allerdings ist es günstiger mit V/U zu arbeiten, da das Komplement W nicht eindeutig bestimmt ist.

Satz 16.1 *Die Abbildung*

$$\begin{aligned} \pi: \ V &\longrightarrow V/U \\ a &\longmapsto a + U \end{aligned}$$

ist ein Epimorphismus mit $\operatorname{Ker} \pi = U$.

Beweis.

(a) Es gilt

$$\pi(a+b) = (a+b) + U = (a+U) + (b+U) = \pi(a) + \pi(b).$$

Analog zeigt man $\pi(\lambda a) = \lambda\pi(a)$.

(b) Nach Konstruktion ist π surjektiv.

(c) Es gilt

$$\pi(a) = 0 \Leftrightarrow a + U = 0 + U \Leftrightarrow a \sim 0 \Leftrightarrow a - 0 = a \in U,$$

d.h. $\operatorname{Ker}\pi = U$. □

Korollar 16.2 *Zu jedem Unterraum $U \subset V$ eines K-Vektorraums V gibt es eine kanonische kurze exakte Sequenz*

$$0 \longrightarrow U \hookrightarrow V \xrightarrow{\pi} V/U \longrightarrow 0.$$

Korollar 16.3 *Es sei U ein Unterraum eines endlich dimensionalen K-Vektorraums V. Dann gilt*

$$\dim U + \dim V/U = \dim V.$$

Beweis.

$$\dim V = \dim \operatorname{Ker}\pi + \dim \operatorname{Im}\pi = \dim U + \dim V/U.$$

□

Wir betrachten nun eine lineare Abbildung $f : V \to W$ zwischen K-Vektorräumen V und W.

Satz 16.2 (Kern-Bild-Satz) *Es gibt genau einen Isomorphismus*

$$\overline{f} : V/\operatorname{Ker} f \to \operatorname{Im} f,$$

so dass gilt:

$$\overline{f}(a + \operatorname{Ker} f) = f(a) \text{ für alle } a \in V.$$

Beweis. Falls $\overline{f}$ existiert, muss gelten

$$\overline{f}(a + \operatorname{Ker} f) = f(a) \text{ für alle } a \in V.$$

Wir benutzen daher diese Gleichung zur Definition von $\overline{f}$.

(a) Die Abbildung $\overline{f}$ ist wohldefiniert:

$$a \sim b \Rightarrow a - b \in \operatorname{Ker} f \Rightarrow f(a-b) = 0 \Rightarrow f(a) = f(b).$$

(b) Die Abbildung $\overline{f}$ ist linear:

$$\begin{aligned}\overline{f}((a + \operatorname{Ker} f) + (b + \operatorname{Ker} f)) &= \overline{f}((a+b) + \operatorname{Ker} f) = f(a+b) \\ &= f(a) + f(b) = \overline{f}(a + \operatorname{Ker} f) + \overline{f}(b + \operatorname{Ker} f).\end{aligned}$$

Der Beweis für die skalare Multiplikation geht analog.

(c) Die Abbildung $\overline{f}$ ist nach Konstruktion surjektiv.

(d) Die Abbildung $\overline{f}$ ist injektiv:

$$\overline{f}(a + \operatorname{Ker} f) = 0 \Leftrightarrow f(a) = 0 \Leftrightarrow a \in \operatorname{Ker} f \Leftrightarrow a + \operatorname{Ker} f = 0 + \operatorname{Ker} f.$$

□

Beispiel 16.2 Es sei $V = U \oplus W$ und $f : V \to W$ die Projektion auf W, d.h. $f(u + w) = w$ für alle $v = u + w \in V$ mit $u \in U$ und $w \in W$. Dann ist $U = \operatorname{Ker} f$ und $W = \operatorname{Im} f$. Dann ist die Abbildung

$$\overline{f} : V/U \to W, \quad \overline{f}(u + w + U) = w,$$

ein Isomorphismus. Dies ist die obige Deutung des Quotientenraums.

17 Projektive Räume

Es sei V ein endlich dimensionaler K-Vektorraum. Wir setzen

$$V' := V \setminus \{0\}.$$

Definition Für $u, v \in V'$ definieren wir

$$u \sim v :\Leftrightarrow \text{Es gibt } \lambda \in K^* = K \setminus \{0\} \text{ mit } u = \lambda v.$$

Bemerkung 17.1 (i) $\sim$ ist eine Äquivalenzrelation.

(ii) Es gilt

$$u \sim v \Leftrightarrow Ku = Kv \Leftrightarrow u, v \text{ spannen dieselbe Gerade auf.}$$

Definition Der zu V gehörige *projektive Raum* ist

$$\mathbb{P}(V) := V'/\sim .$$

Als Menge ist $\mathbb{P}(V)$ gerade die Menge der Ursprungsgeraden in V:

$$\mathbb{P}(V) = \{Kv \mid v \in V'\} = \{\text{Geraden in } V \text{ durch } 0\}.$$

Definition Die *Dimension* des projektiven Raums $\mathbb{P}(V)$ ist

$$\dim \mathbb{P}(V) := \dim V - 1.$$

Notation Der Raum $\mathbb{P}^n(K) := \mathbb{P}(K^{n+1})$ heißt der *n-dimensionale projektive Raum über K*.

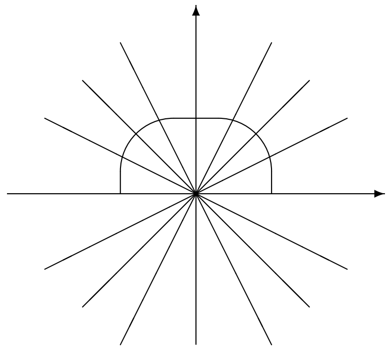

Abbildung 10: Die reelle projektive Gerade $\mathbb{P}^1(\mathbb{R})$

Bemerkung 17.2 Insbesondere ist $\mathbb{P}(\{0\}) = \emptyset$ und $\dim \emptyset = -1$.

Beispiel 17.1 Für die reelle projektive Gerade $\mathbb{P}^1(\mathbb{R})$ gilt:

$$\mathbb{P}^1(\mathbb{R}) = \mathbb{P}(\mathbb{R}^2) = S^1 = \mathbb{R} \cup \{\infty\} = \mathbb{R} \cup \mathbb{P}^0(\mathbb{R}).$$

Die Identifikation $S^1 = \mathbb{R} \cup \{\infty\}$ geschieht über die stereographische Projektion (Skizze siehe Vorlesung).

Beispiel 17.2 Wir betrachten die reelle projektive Ebene $\mathbb{P}^2(\mathbb{R}) = \mathbb{P}(\mathbb{R}^3)$. Man hat eine Zerlegung

$$\mathbb{P}^2(\mathbb{R}) = \mathbb{R}^2 \cup \mathbb{P}^1(\mathbb{R})$$

(Näheres siehe Vorlesung).

Definition Es sei

$$\begin{array}{rccl} \pi: & V' & \longrightarrow & \mathbb{P}(V) \\ & v & \longmapsto & [v] = Kv \end{array}$$

die kanonische Projektion. Ist $U \subset \mathbb{P}(V)$ eine Teilmenge, so definieren wir

$$\widetilde{U} := \pi^{-1}(U) \cup \{0\}.$$

Eine Teilmenge $U \subset \mathbb{P}(V)$ heißt ein *projektiver Unterraum* von $\mathbb{P}(V)$, falls $\widetilde{U} \subset V$ ein Untervektorraum von V ist.

Lemma 17.1 *Für einen projektiven Unterraum U gilt $U = \mathbb{P}(\widetilde{U})$. Insbesondere ist U selbst wieder ein projektiver Raum und hat die Dimension*

$$\dim U = \dim \widetilde{U} - 1.$$

Sprechweise

(i) $\dim U = -1$: $\emptyset$

(ii) $\dim U = 0$: Punkt

(iii) $\dim U = 1$: projektive Gerade

(iv) $\dim U = 2$: projektive Ebene

(v) $\dim U = \dim \mathbb{P}(V) - 1$: (projektive) Hyperebene.

Es sei nun

$$V = K^{n+1} = \{(x_0, x_1, \ldots, x_n) \,|\, x_i \in K, i = 0, 1, \ldots, n\}.$$

Definition Es sei

$$v = (x_0, x_1, \ldots, x_n) \neq 0 \quad \text{(d.h. } x_i \neq 0 \text{ für ein } i).$$

Den Punkt $Kv \in \mathbb{P}^n(K)$ bezeichnen wir mit $(x_0 : x_1 : \ldots : x_n)$. Wir nennen $(x_0 : x_1 : \ldots : x_n)$ die *homogenen Koordinaten* des Punktes Kv.

Bemerkung 17.3 Es gilt $(x_0 : x_1 : \ldots : x_n) = (x'_0 : x'_1 : \ldots : x'_n)$ genau dann, wenn es ein $\lambda \in K$, $\lambda \neq 0$, gibt mit $x_i = \lambda x'_i$ für alle $i = 0, 1, \ldots, n$.

Es sei nun $\widetilde{U} \subset K^{n+1}$ ein Unterraum. Dann ist $\widetilde{U}$ Lösungsmenge eines homogenen Gleichungssystems

$$\begin{array}{ccccc} a_{10}x_0 + & \cdots & +a_{1n}x_n & = & 0 \\ \vdots & & \vdots & \vdots & \vdots \\ a_{m0}x_0 + & \cdots & +a_{mn}x_n & = & 0 \end{array}$$

Mit $(x_0, \ldots, x_n)$ ist auch $(\lambda x_0, \ldots, \lambda x_n)$ für $\lambda \in K^*$ eine Lösung dieses Gleichungssystems. Deswegen kann man schreiben:

$$\mathbb{P}(\widetilde{U}) = \left\{ (x_0 : \ldots : x_n) \in \mathbb{P}^n(K) \,\middle|\, \sum_{j=0}^{n} a_{ij}x_j = 0, i = 1, \ldots, m \right\}.$$

Wir betrachten nun speziell die Gleichung $x_0 = 0$.

Definition Die projektive Hyperebene

$$H_\infty := \{(x_0 : \ldots : x_n) \in \mathbb{P}^n(K) \,|\, x_0 = 0\}$$

heißt die *Hyperebene im Unendlichen*.

Definition Die Abbildung

$$\begin{array}{rccl} \iota: & K^n & \longrightarrow & \mathbb{P}^n(K) \\ & (x_1,\dots,x_n) & \longmapsto & (1:x_1:\dots:x_n) \end{array}$$

heißt die *kanonische Einbettung* von K^n in den projektiven Raum $\mathbb{P}^n(K)$.

Diese Abbildung ist injektiv und hat als Bild gerade die Menge

$$A := \mathbb{P}^n(K) \setminus H_\infty.$$

Denn sei $y = (y_0 : \dots : y_n) \in A$. Dann ist $y_0 \neq 0$ und damit

$$(y_0 : y_1 : \dots : y_n) = (1 : \tfrac{y_1}{y_0} : \dots : \tfrac{y_n}{y_0}) = \iota(\tfrac{y_1}{y_0} : \dots : \tfrac{y_n}{y_0}).$$

Also ist

$$\iota : K^n \to A$$

eine Bijektion. Wenn wir K^n mit A mittels ι identifizieren, erhalten wir

$$\mathbb{P}^n(K) = A \cup H_\infty = K^n \cup \mathbb{P}^{n-1}(K).$$

Definition Man nennt $A = \mathbb{P}^n(K) \setminus H_\infty$ den *affinen Teil* von $\mathbb{P}^n(K)$. Man sagt, dass sich der projektive Raum $\mathbb{P}^n(K) = A \cup H_\infty$ aus dem affinen Teil A und der Hyperebene im Unendlichen zusammensetzt.

Bemerkung 17.4 Statt der Hyperebene $H_\infty = \{x_0 = 0\}$ hätte man auch jede andere Hyperebene als Hyperebene im Unendlichen auswählen können.

Definition Eine Teilmenge $W \subset K^n$ heißt ein *affiner Unterraum* des Vektorraums K^n, falls es einen Unterraum W_0 von K^n und ein Element $a \in K^n$ gibt mit

$$W = a + W_0.$$

Zur Vermeidung von Fallunterscheidungen soll auch die leere Menge ein affiner Unterraum sein. Die *Dimension* eines affinen Unterraums W wird als $\dim W_0$ definiert. Wir definieren $\dim \emptyset := -1$.

Beispiel 17.3 Die Lösungsmenge eines linearen Gleichungssystems

$$\begin{array}{cccc} a_{11}x_1 + & \cdots & + a_{1n}x_n & = & b_1 \\ \vdots & & \vdots & \vdots & \vdots \\ a_{m1}x_1 + & \cdots & + a_{mn}x_n & = & b_m \end{array}$$

ist ein affiner Unterraum. Umgekehrt ist jeder affine Unterraum $W \subset K^n$ Lösungsmenge eines solchen linearen Gleichungssystems.

Lemma 17.2

(i) *Es sei* $\overline{W} \subset \mathbb{P}^n(K)$ *ein projektiver Unterraum. Dann ist* $W := \overline{W} \cap A \subset A = K^n$ *ein affiner Unterraum.*

(ii) *Ist* $\emptyset \neq W \subset A = K^n$ *ein affiner Unterraum, so gibt es genau einen projektiven Unterraum* $\overline{W} \subset \mathbb{P}^n(K)$ *mit* $\overline{W} \cap A = W$.

Beweis. (i) Die Menge $\widetilde{\overline{W}} := \pi^{-1}(\overline{W}) \cup \{0\}$ ist ein linearer Unterraum von K^{n+1}, also Lösungsmenge eines homogenen Gleichungssystems

$$\begin{array}{cccccc} a_{10}y_0+ & \cdots & +a_{1n}y_n & = & 0 \\ \vdots & & \vdots & \vdots & \vdots \\ a_{m0}y_0+ & \cdots & +a_{mn}y_n & = & 0. \end{array}$$

Für einen Punkt

$$y = (y_0 : \ldots : y_n) \notin H_\infty \quad \text{(d.h. } y_0 \neq 0\text{)}$$

gilt

$$a_{i0}y_0 + \cdots + a_{in}y_n = 0 \Leftrightarrow a_{i0} + a_{i1}\frac{y_1}{y_0} + \cdots + a_{in}\frac{y_n}{y_0} = 0.$$

Also ist $\overline{W} \cap A$ die Lösungsmenge des Gleichungssystems

$$\begin{array}{cccccc} a_{11}x_1+ & \cdots & +a_{1n}x_n & = & -a_{10} \\ \vdots & & \vdots & \vdots & \vdots \\ a_{m1}x_1+ & \cdots & +a_{mn}x_n & = & -a_{m0}. \end{array}$$

Daher ist $\overline{W} \cap A$ ein affiner Unterraum.

(ii) Es sei umgekehrt $W \subset A$ ein affiner Unterraum. Dann ist W die Lösungsmenge eines Gleichungssystems

$$\begin{array}{cccccc} a_{11}x_1+ & \cdots & +a_{1n}x_n & = & b_1 \\ \vdots & & \vdots & \vdots & \vdots \\ a_{m1}x_1+ & \cdots & +a_{mn}x_n & = & b_m. \end{array}$$

Sind $(y_0 : \ldots : y_n)$ wieder die homogenen Koordinaten des $\mathbb{P}^n(K)$, so betrachten wir das homogene lineare Gleichungssystem

$$\begin{array}{cccccc} -b_1y_0 + a_{11}y_1+ & \cdots & +a_{1n}y_n & = & 0 \\ \vdots & & \vdots & \vdots & \vdots \\ -b_my_0 + a_{m1}y_1+ & \cdots & +a_{mn}y_n & = & 0. \end{array}$$

Dann ist die Lösungsmenge $\overline{W}$ dieses Gleichungssystems ein projektiver Unterraum des $\mathbb{P}^n(K)$ und nach Konstruktion gilt $\overline{W} \cap A = W$.

Wir müssen noch die Eindeutigkeit zeigen. Es sei $\overline{U}$ ein weiterer projektiver Unterraum mit $\overline{U} \cap A = W$. Wir setzen

$$\overline{W}_\infty := \overline{W} \cap H_\infty, \quad \overline{U}_\infty := \overline{U} \cap H_\infty.$$

Dann gilt

$$\pi^{-1}(W) = \pi^{-1}(\overline{W} - \overline{W}_\infty) = \pi^{-1}(\overline{U} - \overline{U}_\infty) \neq \emptyset.$$

Da das mengentheoretische Komplement eines echten Unterraums den ganzen Vektorraum erzeugt, gilt

$$\widetilde{\overline{W}} = \mathrm{Span}(\pi^{-1}(\overline{W} - \overline{W}_\infty)) = \mathrm{Span}(\pi^{-1}(\overline{U} - \overline{U}_\infty)) = \widetilde{\overline{U}}.$$

Daraus folgt $\overline{W} = \overline{U}$. □

Lemma 17.3 *Der Durchschnitt von zwei projektiven Unterräumen U_1 und U_2 ist wieder ein projektiver Unterraum.*

Beweis. Es sei

$$U_1 = \mathbb{P}(\widetilde{U}_1), \quad U_2 = \mathbb{P}(\widetilde{U}_2).$$

Dann gilt

$$U_1 \cap U_2 = \mathbb{P}(\widetilde{U}_1 \cap \widetilde{U}_2).$$

□

Definition Es seien $U_1, \ldots, U_r \subset \mathbb{P}(V)$ projektive Unterräume. Dann ist der *Spann* von $U_1, \ldots, U_r$, in Zeichen $U_1 \vee \ldots \vee U_r$, der kleinste projektive Unterraum von $\mathbb{P}(V)$, der $U_1, \ldots, U_r$ enthält.

Bemerkung 17.5 Es gilt

$$U_1 \vee \ldots \vee U_r = \mathbb{P}(\widetilde{U}_1 + \cdots + \widetilde{U}_r).$$

Lemma 17.4 (Dimensionsformel) *Es seien $U_1, U_2 \subset \mathbb{P}(V)$ projektive Unterräume. Dann gilt*

$$\dim U_1 + \dim U_2 = \dim(U_1 \vee U_2) + \dim(U_1 \cap U_2).$$

Beweis. Es sei $U_i = \mathbb{P}(\widetilde{U}_i)$, $i = 1, 2$. Dann gilt

$$\begin{aligned} \dim U_1 + \dim U_2 &= \dim \widetilde{U}_1 - 1 + \dim \widetilde{U}_2 - 1 \\ &= \dim(\widetilde{U}_1 + \widetilde{U}_2) - 1 + \dim(\widetilde{U}_1 \cap \widetilde{U}_2) - 1 \\ &= \dim(U_1 \vee U_2) + \dim(U_1 \cap U_2). \end{aligned}$$

□

Beispiel 17.4 Wir betrachten die projektive Ebene $\mathbb{P}^2(K)$. Es seien $\overline{L}_1, \overline{L}_2 \subset \mathbb{P}^2(K)$ zwei projektive Geraden. Dann gilt

$$\dim(\overline{L}_1 \cap \overline{L}_2) = \dim \overline{L}_1 + \dim \overline{L}_2 - \dim(\overline{L}_1 \vee \overline{L}_2) \geq 2 - 2 = 0.$$

Daraus folgt $\overline{L}_1 \cap \overline{L}_2 \neq \emptyset$, d.h. in einer projektiven Ebene schneiden sich je zwei Geraden stets. Sind $\overline{L}_1$ und $\overline{L}_2$ verschieden, so ist der Durchschnitt genau ein Punkt. Liegt dieser Punkt in H_∞, so sind $L_1 = \overline{L}_1 \cap A$ und $L_2 := \overline{L}_2 \cap A$ zwei parallele affine Geraden.

Nun wollen wir auch Abbildungen von projektiven Räumen betrachten. Es seien V, W endlich dimensionale K-Vektorräume und $\mathbb{P}(V), \mathbb{P}(W)$ die zugehörigen projektiven Räume. Es sei $F : V \to W$ eine injektive lineare Abbildung. Dann gilt für $v \in V$, $v \neq 0$,

$$F(Kv) = KF(v) \neq \{0\}.$$

Daher induziert F eine injektive Abbildung

$$\begin{array}{rccc} \overline{F}: & \mathbb{P}(V) & \longrightarrow & \mathbb{P}(W) \\ & Kv & \longmapsto & KFv \end{array}.$$

Definition Eine Abbildung

$$f : \mathbb{P}(V) \to \mathbb{P}(W)$$

heißt *projektiv*, falls es eine injektive lineare Abbildung $F : V \to W$ gibt mit $\overline{F} = f$. Eine bijektive projektive Abbildung heißt *Projektivität*.

Beispiel 17.5 Für $m \geq n$ haben wir eine kanonische Einbettung

$$\begin{array}{rccc} \overline{J}: & \mathbb{P}^n(K) & \longrightarrow & \mathbb{P}^m(K) \\ & (x_0 : \ldots : x_n) & \longmapsto & (x_0 : \ldots : x_n : 0 : \ldots : 0) \end{array}.$$

Sie entsteht aus der linearen Abbildung

$$\begin{array}{rccc} J: & K^{n+1} & \longrightarrow & K^{m+1} \\ & (x_0, \ldots, x_n) & \longmapsto & (x_0, \ldots, x_n, 0, \ldots, 0) \end{array}.$$

Lemma 17.5 *Für zwei injektive lineare Abbildungen $F, F' : V \to W$ gilt $\overline{F} = \overline{F}'$ genau dann, wenn es ein $\lambda \in K^*$ gibt mit $F' = \lambda F$.*

Beweis. Ist $F' = \lambda F$, so gilt offensichtlich $\overline{F}' = \overline{F}$. Es bleibt die Umkehrung zu zeigen: Ist $\overline{F} = \overline{F}'$, so gibt es zu jedem $v \in V$ ein $\lambda_v \in K^*$ mit $F'(v) = \lambda_v F(v)$. Es ist zu zeigen, dass man zu jedem v das gleiche λ_v wählen kann.

Für $\dim V \leq 1$ ist das klar. Andernfalls gibt es linear unabhängige $v, w \in V$. Dann gibt es $\lambda_v, \lambda_w, \lambda_{v+w} \in K^*$ mit

$$F'(v) = \lambda_v F(v), \quad F'(w) = \lambda_w F(w), \quad F'(v+w) = \lambda_{v+w} F(v+w).$$

Aus der Linearität von F und F' folgt

$$(\lambda_v - \lambda_{v+w})F(v) + (\lambda_w - \lambda_{v+w})F(w) = 0.$$

Da F injektiv ist, sind auch $F(v), F(w)$ linear unabhängig. Also folgt

$$\lambda_v = \lambda_{v+w} = \lambda_w.$$

□

Um projektive Abbildungen durch Matrizen zu beschreiben, führen wir Koordinatensysteme ein. Es sei $\dim V = n + 1$, also $\dim \mathbb{P}(V) = n$.

Definition Die Punkte $P_0, \ldots, P_k \in \mathbb{P}(V)$ heißen *projektiv unabhängig*, falls

$$\dim(P_0 \vee \ldots \vee P_k) = k$$

gilt.

Bemerkung 17.6 Es sei $P_i = Kv_i$, $i = 0, \ldots, k$. Dann gilt

$$P_0, \ldots, P_k \text{ projektiv unabhängig} \Leftrightarrow v_0, \ldots, v_k \text{ linear unabhängig.}$$

Definition Ein $(n + 2)$-Tupel $(P_0, \ldots, P_{n+1})$ von Punkten aus $\mathbb{P}(V)$ heißt *projektive Basis*, falls je $n + 1$ Punkte davon projektiv unabhängig sind.

Lemma 17.6 *Es sei $(P_0, \ldots, P_{n+1})$ eine projektive Basis. Dann gibt es eine Basis $(v_0, \ldots, v_n)$ von V mit*

(i) $P_i = Kv_i$ *($i = 0, \ldots, n$).*

(ii) $P_{n+1} = K(v_0 + \cdots + v_n)$.

Die Basis $(v_0, \ldots, v_n)$ ist bis auf einen Skalar eindeutig bestimmt.

Beweis. Da $P_0, \ldots, P_n$ projektiv unabhängig sind, gibt es eine Basis $(w_0, \ldots, w_n)$ von V mit

$$P_0 = Kw_0, \ldots, P_n = Kw_n.$$

Weiter gibt es $\lambda_0, \ldots, \lambda_n \in K$ mit

$$P_{n+1} = K(\lambda_0 w_0 + \cdots + \lambda_n w_n).$$

Wäre $\lambda_0 = 0$, so wären $P_1, \dots, P_{n+1}$ nicht projektiv unabhängig. Also ist $\lambda_0 \neq 0$ und analog $\lambda_1 \neq 0, \dots, \lambda_n \neq 0$. Daher ist durch

$$v_0 := \lambda_0 w_0, \dots, v_n := \lambda_n w_n$$

die gesuchte Basis gegeben. □

Es sei nun $(P_0, \dots, P_{n+1})$ eine projektive Basis und $B = (v_0, \dots, v_n)$ eine zugehörige Basis, die bis auf einen Skalar eindeutig bestimmt ist. Es sei $P = Kv \in \mathbb{P}(V)$. Dann ist v auch bis auf einen Skalar eindeutig festgelegt. Wir ordnen dem Punkt P den Koordinatenvektor $(x_0, \dots, x_n)$ von v bezüglich der Basis B zu. Er ist damit bis auf einen Skalar festgelegt.

Definition Das Element $(x_0 : \ldots : x_n) \in \mathbb{P}^n(K)$ heißt der *homogene Koordinatenvektor* des Punktes P bezüglich der projektiven Basis $(P_0, \dots, P_{n+1})$.

Notation Wir schreiben $P = (x_0 : \ldots : x_n)$. Damit gilt:

$$\begin{array}{rcl} P_0 & = & (1 : 0 : \ldots : 0 : 0) \\ \vdots & \vdots & \vdots \\ P_n & = & (0 : 0 : \ldots : 0 : 1) \\ P_{n+1} & = & (1 : 1 : \ldots : 1 : 1). \end{array}$$

Durch die Einführung von Koordinaten reduziert sich das Studium der Projektivitäten beliebiger projektiver Räume auf das Studium von Projektivitäten des $\mathbb{P}^n(K)$. Es sei

$$f : \mathbb{P}^n(K) \to \mathbb{P}^n(K)$$

eine Projektivität. Dann gibt es einen Isomorphismus

$$F : K^{n+1} \to K^{n+1} \quad \text{mit } \overline{F} = f.$$

Die lineare Abbildung F wird durch eine Matrix $A \in \mathrm{GL}(n+1; K)$ gegeben, wobei F und somit auch A bis auf einen Skalar $\lambda \neq 0$ festgelegt sind.

Wir untersuchen nun den Zusammenhang zwischen Affinitäten und Projektivitäten. Es sei

$$f_a : K^n \to K^n, \quad x \mapsto Ax + b \quad (A \in \mathrm{GL}(n; K)),$$

eine Affinität. Dann betrachten wir die kanonische Einbettung

$$\begin{array}{rccc} \iota : & K^n & \longrightarrow & \mathbb{P}^n(K) \\ & (x_1, \dots, x_n) & \longmapsto & (1 : x_1 : \ldots : x_n) \end{array}$$

von K^n in den $\mathbb{P}^n(K)$. Setzen wir

$$\mathbf{A} = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline b_1 & & & \\ \vdots & & A & \\ b_n & & & \end{array}\right),$$

so ist durch

$$\begin{pmatrix} x_0 \\ \vdots \\ x_n \end{pmatrix} \mapsto \mathbf{A} \begin{pmatrix} x_0 \\ \vdots \\ x_n \end{pmatrix}$$

eine bijektive lineare Abbildung $F_a : K^{n+1} \to K^{n+1}$ bestimmt, die eine Projektivität $\overline{F}_a : \mathbb{P}^n(K) \to \mathbb{P}^n(K)$ induziert. Es gilt $\overline{F}_a(H_\infty) = H_\infty$. Damit haben wir eine Motivation für die Konstruktion in § 11 nachgeliefert.

Es sei nun umgekehrt $f : \mathbb{P}^n(K) \to \mathbb{P}^n(K)$ eine Projektivität mit $f(H_\infty) = H_\infty$. Eine zugehörige lineare Abbildung $F : K^{n+1} \to K^{n+1}$ sei gegeben durch die Matrix $\mathbf{A} \in \mathrm{GL}(n+1; K)$ mit

$$\mathbf{A} = \begin{pmatrix} a_{00} & a_{01} & \cdots & a_{0n} \\ a_{10} & a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n0} & a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

Wegen $f(H_\infty) = H_\infty$ muss gelten

$$\mathbf{A} = \begin{pmatrix} a_{00} & 0 & \cdots & 0 \\ a_{10} & a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n0} & a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

Wegen $\mathbf{A} \in \mathrm{GL}(n+1; K)$ ist $a_{00} \neq 0$. Da $\mathbf{A}$ nur bis auf einen Skalar eindeutig festgelegt ist, kann man annehmen dass

$$\mathbf{A} = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline a_{10} & & & \\ \vdots & & A & \\ a_{n0} & & & \end{array}\right)$$

für eine Matrix $A \in \mathrm{GL}(n; K)$. Es gilt daher

$$\left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline a_{10} & & & \\ \vdots & & A & \\ a_{n0} & & & \end{array}\right) \begin{pmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 1 \\ x'_1 \\ \vdots \\ x'_n \end{pmatrix}.$$

Das bedeutet, dass $f_a := f|_{K^n} : K^n \to K^n$ eine Affinität ist, die durch

$$f_a \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} a_{10} \\ \vdots \\ a_{n0} \end{pmatrix}$$

gegeben wird.

Wir haben damit bewiesen:

Satz 17.1 *Ist $f_a : K^n \to K^n$ eine Affinität, so gibt es eine Projektivität $f : \mathbb{P}^n(K) \to \mathbb{P}^n(K)$ mit $f|_{K^n} = f_a$ und $f(H_\infty) = H_\infty$.*

Ist umgekehrt $f : \mathbb{P}^n(K) \to \mathbb{P}^n(K)$ eine Projektivität mit $f(H_\infty) = H_\infty$, dann ist $f_a := f|_{K^n} : K^n \to K^n$ eine Affinität.

18 Projektive Quadriken

Wir wollen nun Quadriken in projektiven Räumen betrachten.

Es sei K ein Körper mit $n_K \neq 0$ für $n = 2$.

Definition Unter einem *homogenen Polynom zweiten Grades* in den Unbestimmten $x_0, x_1, \ldots, x_n$ versteht man einen Ausdruck der Form

$$q(x_0, \ldots, x_n) = \sum_{0 \leq i \leq j \leq n} \alpha_{ij} x_i x_j,$$

wobei $\alpha_{ij} \in K$ für $0 \leq i \leq j \leq n$.

Definition Eine Teilmenge $C \subset K^{n+1}$ heißt *Kegel*, wenn für jedes $(x_0, \ldots, x_n) \in C$ und $\lambda \in K$ auch $(\lambda x_0, \ldots, \lambda x_n) \in C$ ist.

Dies bedeutet, dass C die Vereinigung von Geraden durch den Ursprung ist. Eine Gerade durch den Ursprung von K^{n+1} ist ein Punkt von $\mathbb{P}^n(K)$. Für jedes homogene Polynom zweiten Grades q ist die Menge

$$C = \{(x_0, \ldots, x_n) \in K^{n+1} \,|\, q(x_0, \ldots, x_n) = 0\}$$

ein Kegel. Dies folgt aus

$$q(\lambda x_0, \ldots, \lambda x_n) = \lambda^2 q(x_0, \ldots, x_n) \quad \text{für } \lambda \in K.$$

Definition Eine Teilmenge $Q \subset \mathbb{P}^n(K)$ heißt (*projektive*) *Quadrik* (oder (*projektive*) *Hyperfläche zweiter Ordnung*), wenn es ein homogenes Polynom zweiten Grades q gibt, so dass

$$Q = \{(x_0 : \ldots : x_n) \in \mathbb{P}^n(K) \,|\, q(x_0, \ldots, x_n) = 0\}.$$

Bemerkung 18.1 Man beachte, dass im Gegensatz zum affinen Fall das Polynom q keine Funktion auf dem $\mathbb{P}^n(K)$ definiert, denn für $\lambda \in K^*$ gilt

$$\begin{aligned}(x_0 : \ldots : x_n) &= (\lambda x_0 : \ldots : \lambda x_n), \quad \text{aber} \\ q(\lambda x_0, \ldots, \lambda x_n) &= \lambda^2 q(x_0, \ldots, x_n).\end{aligned}$$

Aber die Nullstellenmenge von q in $\mathbb{P}^n(K)$ ist wohldefiniert, denn es gilt für $\lambda \in K^*$:

$$q(x_0, \ldots, x_n) = 0 \Leftrightarrow q(\lambda x_0, \ldots, \lambda x_n) = 0.$$

Beispiel 18.1 Es sei $K = \mathbb{R}$ und $n = 2$.

(1) Es sei $q(x_0, x_1, x_2) = x_0^2 + x_1^2 - x_2^2$. Dann ist

$$C = \{(x_0, x_1, x_2) \in \mathbb{R}^3 \,|\, x_0^2 + x_1^2 - x_2^2 = 0\}$$

ein Kegel und

$$Q = \{(x_0 : x_1 : x_2) \in \mathbb{P}^2(\mathbb{R}) \,|\, x_0^2 + x_1^2 - x_2^2 = 0\}$$

die Menge der in C liegenden Geraden durch den Ursprung. Um uns ein Bild von Q zu beschaffen, betrachten wir den affinen Teil, wobei wir eine geeignete Hyperebene H_∞ im Unendlichen entfernen.

(a) Es sei $H_\infty := \{(x_0 : x_1 : x_2) \in \mathbb{P}^2(\mathbb{R}) \,|\, x_0 = 0\}$. Auf $\mathbb{R}^2 = \mathbb{P}^2(\mathbb{R}) \setminus H_\infty$ erhalten wir die Gleichung

$$1 + x_1^2 - x_2^2 = 0.$$

Diese Gleichung definiert eine Hyperbel. Sie entsteht als Schnitt des Kegels C mit der Ebene $x_0 = 1$.

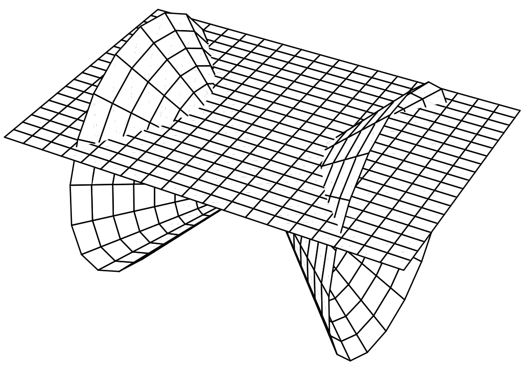

Abbildung 11: Schnitt von C mit der Ebene $x_0 = 1$

(b) Es sei $H_\infty := \{(x_0 : x_1 : x_2) \in \mathbb{P}^2(\mathbb{R}) \,|\, x_2 = 0\}$. Dann erhalten wir auf $\mathbb{R}^2 = \mathbb{P}^2(\mathbb{R}) \setminus H_\infty$ die Gleichung

$$x_0^2 + x_1^2 - 1 = 0.$$

Diese Gleichung definiert einen Kreis. Er entsteht als Schnitt des Kegels C mit der Ebene $x_2 = 1$. Wählt man als Hyperebene im Unendlichen $H_\infty := \{(x_0 : x_1 : x_2) \in \mathbb{P}^2(\mathbb{R}) \,|\, \frac{1}{2}x_0 + x_2 = 0\}$, so kann man leicht zeigen, dass man im affinen Teil eine Ellipse erhält.

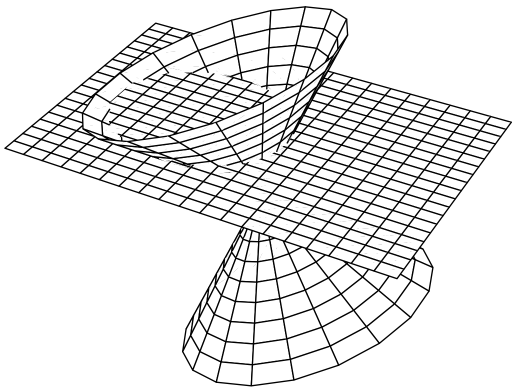

Abbildung 12: Schnitt von C mit der Ebene $\frac{1}{2}x_0 + x_2 = 1$

(c) Es sei $H_\infty := \{(x_0 : x_1 : x_2) \in \mathbb{P}^2(\mathbb{R}) \,|\, x_0 + x_2 = 0\}$. Dann erhalten wir auf $\mathbb{R}^2 = \mathbb{P}^2(\mathbb{R}) \setminus H_\infty$ die Gleichung

$$(1 - x_2)^2 + x_1^2 - x_2^2 = 0 \Leftrightarrow x_1^2 - 2x_2 = -1.$$

Das ist eine Parabel. Sie entsteht als Schnitt des Kegels C mit der Ebene $x_0 + x_2 = 1$.

(2) Es sei $q(x_0, x_1, x_2) = x_0^2 - x_1^2$. Dann gilt

$$q(x_0, x_1, x_2) = x_0^2 - x_1^2 = (x_0 - x_1)(x_0 + x_1).$$

Damit besteht Q aus zwei Geraden.

Wir wollen die Gleichung für eine Quadrik wieder durch Matrizen ausdrücken. Gegeben sei ein quadratisches Polynom

$$q(x_0, \ldots, x_n) = \sum_{0 \leq i \leq j \leq n} \alpha_{ij} x_i x_j.$$

Dann setzen wir

$$a_{ij} := \begin{cases} \alpha_{ij} & \text{für } i = j, \\ \frac{1}{2}\alpha_{ij} & \text{für } i < j, \\ \frac{1}{2}\alpha_{ji} & \text{für } i > j, \end{cases}$$

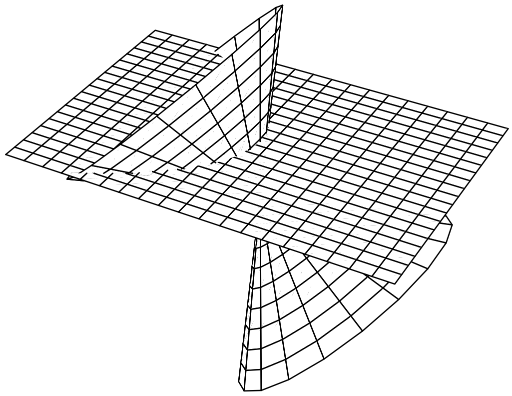

Abbildung 13: Schnitt von C mit der Ebene $x_0 + x_2 = 1$

wobei jeweils $0 \leq i, j \leq n$, und $A := (a_{ij}) \in \mathrm{Mat}(n+1, n+1; K)$. Dann ist A eine symmetrische Matrix und es gilt für $x \in K^{n+1}$

$$q(x) = x^T A x.$$

Wir haben gesehen, dass affine Quadriken unter Affinitäten invariant bleiben. Das Gleiche gilt auch für projektive Quadriken unter Projektivitäten.

Satz 18.1 *Ist $Q \subset \mathbb{P}^n(K)$ eine Quadrik und $f : \mathbb{P}^n(K) \to \mathbb{P}^n(K)$ eine Projektivität, so ist auch $f(Q) \subset \mathbb{P}^n(K)$ eine Quadrik.*

Beweis. Es sei Q gegeben durch

$$x^T A x = 0, \quad A \in \mathrm{Mat}(n+1, n+1; K),$$

und die zu $f : \mathbb{P}^n(K) \to \mathbb{P}^n(K)$ gehörige lineare Abbildung $F : K^{n+1} \to K^{n+1}$ sei gegeben durch

$$y = Sx, \quad S \in \mathrm{GL}(n+1; K).$$

Es sei C der Kegel

$$C := \{x \in K^{n+1} \,|\, x^T A x = 0\}.$$

Dann gilt für $y = (y_0, \ldots, y_n)^T \in K^{n+1}$ und $T := S^{-1}$

$$\begin{aligned} y = Sx \in F(C) \quad &\Leftrightarrow \quad x = Ty \in C \\ &\Leftrightarrow \quad 0 = x^T A x = (Ty)^T A (Ty) = y^T (T^T A T) y. \end{aligned}$$

Die Matrix $B := T^T A T$ ist wieder symmetrisch und es gilt

$$f(Q) = \{(y_0 : \ldots : y_n) \in \mathbb{P}^n(K) \,|\, y^t B y = 0\}.$$

$\square$

Definition Zwei Quadriken $Q, Q' \subset \mathbb{P}^n(K)$ heißen (projektiv) *äquivalent*, falls es eine Projektivität $f : \mathbb{P}^n(K) \to \mathbb{P}^n(K)$ gibt mit $f(Q) = Q'$.

Beispiel 18.2 Die beiden Quadriken

$$\begin{aligned} Q &= \{(x_0 : x_1 : x_2) \in \mathbb{P}^2(\mathbb{R}) \,|\, x_0^2 + x_1^2 - x_2^2 = 0\} \quad \text{und} \\ Q' &= \{(x_0 : x_1 : x_2) \in \mathbb{P}^2(\mathbb{R}) \,|\, x_0^2 - x_1^2 = 0\} \end{aligned}$$

sind nicht äquivalent, da Q' aus zwei Geraden besteht, aber Q keine Gerade enthält.

Wir kommen nun zur Klassifikation der projektiven Quadriken über dem Körper $K = \mathbb{R}$. Dazu erinnern wir an die Definition der Signatur einer symmetrischen Matrix A mit Einträgen in $\mathbb{R}$:

$$\operatorname{Sign} A = \sharp\{\text{positive Eigenwerte}\} - \sharp\{\text{negative Eigenwerte}\}.$$

Theorem 18.1 (Projektive Klassifikation von Quadriken) *Jede Quadrik $Q \subset \mathbb{P}^n(\mathbb{R})$ ist äquivalent zu genau einer der folgenden Quadriken:*

$$Q_{k,m} = \{(x_0 : \ldots : x_n) \in \mathbb{P}^n(\mathbb{R}) \,|\, x_0^2 + \ldots + x_k^2 - x_{k+1}^2 - \ldots - x_m^2 = 0\}$$

mit $-1 \leq k \leq m \leq n$ und $k+1 \geq m-k$. Insbesondere gilt für zwei Quadriken $Q = \{x^T A x = 0\}$ und $Q' = \{x^T A' x = 0\}$:

$$Q \text{ und } Q' \text{ sind äquivalent } \Leftrightarrow \operatorname{Rang} A = \operatorname{Rang} A', |\operatorname{Sign} A| = |\operatorname{Sign} A'|.$$

Um diesen Satz zu beweisen, brauchen wir noch einen Hilfssatz.

Lemma 18.1 *Es seien $q(x) = x^T A x$ und $q'(x) = x^T A' x$ quadratische Formen auf $V = \mathbb{R}^{n+1}$. Es gelte*

$$C := \{x \in V \,|\, q(x) = 0\} = \{x \in V \,|\, q'(x) = 0\}.$$

Gibt es ein $v_0 \in C$, so dass

$$v_0^T A w \neq 0 \quad \text{für mindestens ein } w \in V, \tag{1}$$

so gilt $A' = \rho \cdot A$ für ein $\rho \in \mathbb{R}^*$.

Beweis. Es seien $b(x, y) := x^T A y$ bzw. $b'(x, y) = x^T A' y$ die zu q bzw. q' gehörigen symmetrischen Bilinearformen. Wir wählen ein festes $v_0 \in C$ mit der Eigenschaft (1) und betrachten für jedes $w \in V$ die Gerade

$$g_w := \{w + \lambda v_0 \,|\, \lambda \in \mathbb{R}\}.$$

Ihre Schnittpunkte mit C sind bestimmt durch die Gleichungen

$$q(w + \lambda v_0) = 0 \quad \text{bzw.} \quad q'(w + \lambda v_0) = 0,$$

oder (wegen $q(v_0) = q'(v_0) = 0$)

$$2\lambda b(v_0, w) + q(w) = 0 \quad \text{bzw.} \quad 2\lambda b'(v_0, w) + q'(w) = 0. \tag{2}$$

Dass diese Gerade den Kegel C nicht in genau einem Punkt schneidet, ist gleichwertig mit

$$b(v_0, w) = 0 \quad \text{bzw.} \quad b'(v_0, w) = 0.$$

Also gilt für alle $w \in V$

$$b(v_0, w) = 0 \Leftrightarrow b'(v_0, w) = 0.$$

Wegen der Bedingung (1) ist die Menge

$$H := \{w \in V \mid b(v_0, w) = 0\}$$

eine Hyperebene in V. Da sie gleich der Menge

$$H' = \{w \in V \mid b'(v_0, w) = 0\}$$

ist, muss es ein $\rho \in \mathbb{R}^*$ geben mit

$$b'(v_0, w) = \rho \cdot b(v_0, w) \quad \text{für alle } w \in V.$$

Aus Gleichung (2) folgt

$$b(v_0, w) \cdot q'(w) = b'(v_0, w) \cdot q(w) \quad \text{für alle } w \in V.$$

Daraus ergibt sich

$$q'(w) = \rho \cdot q(w) \quad \text{für alle } w \in V \setminus H.$$

Die Funktion

$$\begin{array}{rccl} f: & V & \longrightarrow & \mathbb{R} \\ & w & \longmapsto & q'(w) - \rho \cdot q(w) \end{array}$$

verschwindet auf $V \setminus H$. Da sie außerdem stetig ist, verschwindet sie auch auf H. Damit gilt

$$q'(w) = \rho \cdot q(w) \quad \text{für alle } w \in V.$$

Daraus folgt durch Polarisierung

$$b'(v, w) = \rho \cdot b(v, w) \quad \text{für alle } v, w \in V$$

und damit

$$A' = \rho A.$$

$\square$

Beweis von Theorem 18.1. Es sei

$$Q = \{(x_0 : \ldots : x_n) \in \mathbb{P}^n(\mathbb{R}) \,|\, x^T A x = 0\}, A \in \mathrm{Mat}(n+1, n+1; \mathbb{R}), A = A^T.$$

Nach Satz 6.1 gibt es ein $S \in \mathrm{GL}(n+1; \mathbb{R})$ mit

$$S^T A S = \begin{pmatrix} E_{k+1} & & 0 \\ & -E_l & \\ 0 & & 0 \end{pmatrix} =: A_{k,m}, \quad k + l = m.$$

Da

$$x^T A x = 0 \Leftrightarrow x^T(-A)x = 0,$$

können wir annehmen, dass $k + 1 \geq m - k$ ist. Also ist Q äquivalent zu

$$Q_{k,m} = \{(x_0 : \ldots : x_n) \in \mathbb{P}^n(\mathbb{R}) \,|\, x^T A_{k,m} x = 0\}.$$

Aus dem Trägheitssatz von Sylvester folgt auch die Implikation "$\Leftarrow$" der zweiten Aussage des Theorems.

Nun beweisen wir die Richtung "$\Rightarrow$" der zweiten Aussage. Dazu seien $Q = \{x^T A x = 0\}$ und $Q' = \{x^T A' x = 0\}$ äquivalent. O. B. d. A. können wir annehmen, dass $Q = Q_{k,m}$ und $A = A_{k,m}$ gilt. Nach Voraussetzung gibt es eine Projektivität $f : \mathbb{P}^n(\mathbb{R}) \to \mathbb{P}^n(\mathbb{R})$ mit $f(Q) = Q'$, d.h. es gibt eine Matrix $T \in \mathrm{GL}(n+1; \mathbb{R})$, so dass Q auch durch die Bilinearform mit der Matrix

$$B := T^T A' T$$

beschrieben wird. Nun wollen wir Lemma 18.1 anwenden. Dazu müssen wir nachprüfen, wann die Bedingung (1) für $A_{k,m}$ erfüllt ist. Diese Bedingung ist aber erfüllt, wenn $k < m$ ist. Denn dann gilt: Bezeichnen wir mit $(e_0, \ldots, e_n)$ die Standardbasis von $\mathbb{R}^{n+1}$, so sind $v_0 := e_0 + e_m$ und $w := e_0$ Vektoren mit

$$v_0^T A_{k,m} v_0 = 0, \quad \text{aber} \quad v_0^T A_{k,m} w = 1 \neq 0.$$

Aus Lemma 18.1 folgt damit $B := \rho \cdot A_{k,m}$ für ein $\rho \in \mathbb{R}^*$. Aus dem Trägheitssatz von Sylvester (Satz 6.2) folgt dann die Behauptung.

Es bleibt noch der Fall $k = m$ zu behandeln. Dann ist Q ein linearer projektiver Unterraum der Dimension $n-(m+1)$, also auch Q'. Daraus folgt aber

$$\mathrm{Rang}\, A' = \mathrm{Rang}\, A_{m,m} = |\mathrm{Sign}\, A'| = |\mathrm{Sign}\, A_{m,m}| = m + 1.$$

Rang	|Sign|	Gleichung	Beschreibung
0	0	$0 = 0$	$\mathbb{P}^2(\mathbb{R})$
1	1	$x_0^2 = 0$	(Doppel-)Gerade
2	2	$x_0^2 + x_1^2 = 0$	Punkt
2	0	$x_0^2 - x_1^2 = 0$	Geradenpaar
3	3	$x_0^2 + x_1^2 + x_2^2 = 0$	leere Quadrik
3	1	$x_0^2 + x_1^2 - x_2^2 = 0$	nicht ausgeartete Quadrik

Tabelle 4: Normalformen von Quadriken im $\mathbb{P}^2(\mathbb{R})$

Rang	|Sign|	Gleichung	Beschreibung
0	0	$0 = 0$	$\mathbb{P}^3(\mathbb{R})$
1	1	$x_0^2 = 0$	(Doppel-)Ebene
2	2	$x_0^2 + x_1^2 = 0$	Gerade
2	0	$x_0^2 - x_1^2 = 0$	Ebenenpaar
3	3	$x_0^2 + x_1^2 + x_2^2 = 0$	Punkt
3	1	$x_0^2 + x_1^2 - x_2^2 = 0$	Kegel
4	4	$x_0^2 + x_1^2 + x_2^2 + x_3^2 = 0$	leere Quadrik
4	2	$x_0^2 + x_1^2 + x_2^2 - x_3^2 = 0$	Ovalfläche
4	0	$x_0^2 + x_1^2 - x_2^2 - x_3^2 = 0$	Regelfläche

Tabelle 5: Normalformen von Quadriken im $\mathbb{P}^3(\mathbb{R})$

Aus diesem Beweis folgt auch, dass zwei Quadriken $Q_{k,m}$ für unterschiedliche k und m mit $k + 1 \geq m - k$ nicht äquivalent sind. $\square$

Speziell für $n = 2, 3$ erhalten wir die Tabellen 4 und 5.

Beispiel 18.3 Es sei $K = \mathbb{R}$ und $n = 3$.

(1) Wir betrachten die nicht ausgeartete Quadrik

$$Q := \{(x_0 : x_1 : x_2 : x_3) \in \mathbb{P}^3(\mathbb{R}) \,|\, x_0^2 + x_1^2 + x_2^2 - x_3^2 = 0\}.$$

(a) Es sei $H_\infty := \{(x_0 : x_1 : x_2 : x_3) \in \mathbb{P}^3(\mathbb{R}) \,|\, x_3 = 0\}$. Dann ist

$$Q \cap A = \{(x_0, x_1, x_2) \in \mathbb{R}^3 \,|\, x_0^2 + x_1^2 + x_2^2 = 1\}$$

eine Kugel und es gilt $Q \cap H_\infty = \emptyset$.

(b) Es sei $H_\infty := \{(x_0 : x_1 : x_2 : x_3) \in \mathbb{P}^3(\mathbb{R}) \,|\, x_0 = 0\}$. Dann ist

$$Q \cap A = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \,|\, x_3^2 - x_1^2 - x_2^2 = 1\}$$

ein zweischaliges Hyperboloid (Bild 7) und es gilt

$$Q \cap H_\infty = \{(x_1 : x_2 : x_3) \in \mathbb{P}^2(\mathbb{R}) \mid x_1^2 + x_2^2 - x_3^2 = 0\},$$

eine nicht ausgeartete Quadrik in der projektiven Ebene.

(c) Es sei $H_\infty := \{(x_0 : x_1 : x_2 : x_3) \in \mathbb{P}^3(\mathbb{R}) \mid x_2 + x_3 = 0\}$. Dann ist

$$\begin{aligned} Q \cap A &= \{(x_0, x_1, x_2) \in \mathbb{R}^3 \mid x_0^2 + x_1^2 + x_2^2 - (1 - x_2)^2 = 0\} \\ &= \{(x_0, x_1, x_2) \in \mathbb{R}^3 \mid x_0^2 + x_1^2 + 2x_2 = 1\} \end{aligned}$$

ein elliptisches Paraboloid (Bild 8) und es gilt

$$\begin{aligned} Q \cap H_\infty &= \{(x_0 : x_1 : x_2 : x_3) \in \mathbb{P}^3(\mathbb{R}) \mid x_3 = -x_2,\ x_0^2 + x_1^2 = 0\} \\ &= \{(0 : 0 : 1 : -1)\}. \end{aligned}$$

(2) Nun betrachten wir die nicht ausgeartete Quadrik

$$Q := \{(x_0 : x_1 : x_2 : x_3) \in \mathbb{P}^3(\mathbb{R}) \mid x_0^2 + x_1^2 - x_2^2 - x_3^2 = 0\}.$$

(a) Es sei $H_\infty := \{(x_0 : x_1 : x_2 : x_3) \in \mathbb{P}^3(\mathbb{R}) \mid x_3 = 0\}$. Dann ist

$$Q \cap A = \{(x_0, x_1, x_2) \in \mathbb{R}^3 \mid x_0^2 + x_1^2 - x_2^2 = 1\}.$$

Dies ist ein einschaliges Hyperboloid (Bild 6). Es enthält Geraden und daher nennt man eine solche Quadrik eine *Regelfläche*. Tatsächlich ist Q die Vereinigung von Geraden (Fadenmodell des einschaligen Hyperboloids, siehe Vorlesung). Der Schnitt mit der Hyperebene H_∞ ist eine nicht ausgeartete ebene Quadrik ("Kreis")

$$Q \cap H_\infty = \{(x_0 : x_1 : x_2) \in \mathbb{P}^2(\mathbb{R}) \mid x_0^2 + x_1^2 - x_2^2 = 0\}.$$

(b) Es sei $H_\infty := \{(x_0 : x_1 : x_2 : x_3) \in \mathbb{P}^3(\mathbb{R}) \mid x_1 + x_2 = 0\}$. Dann ist

$$Q \cap A = \{(x_0, x_1, x_3) \in \mathbb{R}^3 \mid x_0^2 - x_3^2 + 2x_1 = 1\}.$$

Dies ist ein hyperbolisches Paraboloid (Bild 9). Es enthält ebenfalls Geraden. Sein Durchschnitt mit H_∞ ist ein Paar von Geraden

$$Q \cap H_\infty = \{(x_0 : x_1 : x_2 : x_3) \in \mathbb{P}^3(\mathbb{R}) \mid x_2 = -x_1,\ x_0^2 - x_3^2 = 0\}.$$

Damit haben wir alle interessanten Quadriken im $\mathbb{R}^3$ aus §11 als geeignete affine Teile von projektiven Quadriken im $\mathbb{P}^3(\mathbb{R})$ zurückerhalten.

Inhaltsverzeichnis