Kapitel 5 Euklidische und unitäre Vektorräume

Die im letzten Kapitel entwickelten Kriterien für die Diagonalisierbarkeit eines Endomorphismus Φ oder für die Herleitung der Jordanschen Normalform der Abbildungsmatrizen von Φ erlauben noch keine einfache geometrische Formulierung. Hierzu fehlen uns bisher die aus dem Anschauungsraum vertrauten Begriffe wie Länge, Abstand und Orthogonalität.

Wir wollen deshalb jetzt Vektorräume betrachten, in denen solche Begriffe definiert werden können. Eine sinnvolle Einführung des Längenbegriffs setzt die Vergleichbarkeit von Längen voraus. Dies ist in allgemeinen $\mathbb{K}$-Vektorräumen nicht möglich, daher wollen wir uns im folgenden zunächst auf reelle, später auf komplexe Vektorräume beschränken.

§ 1 Skalarprodukte

Es sei V ein reeller Vektorraum beliebiger Dimension. Wir betrachten eine Bilinearform β auf V, also eine 2 – fach multilineare Abbildung $\beta : V \times V \longrightarrow \mathbb{R}$.

β heißt *symmetrisch*, wenn $\beta(x,y) = \beta(y,x)$ für alle $x,y \in V$ gilt.

β heißt *positiv definit*, wenn $\beta(x,x) > 0$ für $x \neq o$ gilt ($\beta(o,o) = 0$ folgt aus der Bilinearität).

Definition. Eine positiv definite, symmetrische Bilinearform $\beta : V \times V \longrightarrow \mathbb{R}$ heißt *Skalarprodukt* oder *inneres Produkt* auf V. Für Skalarprodukte schreiben wir $<x,y>$ statt $\beta(x,y)$.

Das Paar $(V, <\cdot,\cdot>)$ oder kurz V heißt *euklidischer Vektorraum* oder auch *(reeller) Vektorraum mit Skalarprodukt.*

Beispiele. **(a)** $V = \mathbb{R}^n$, $<x,y> := \sum_{i=1}^{n} x_i\, y_i = x^T y$ ist Skalarprodukt. Es heißt das

Standardskalarprodukt . Im $\mathbb{R}^n$ werden wir mit $<\cdot,\cdot>$ immer dieses Standardskalarprodukt bezeichnen.

(b) $V = \mathbb{R}^2$, $\beta(x,y) := x_1 y_1 - (x_1 y_2 + x_2 y_1) + 2 x_2 y_2$ ist Skalarprodukt und offensichtlich verschieden vom Standardskalarprodukt $<\cdot,\cdot>$.

(c) Es sei $V = C([a,b])$ der Vektorraum der stetigen, reellen Funktionen auf dem Intervall $[a,b]$. Weiter sei s eine feste, positive, stetige Funktion auf $[a,b]$. Dann wird durch

$$<g,h> := \int_a^b g(t)\, h(t)\, s(t)\, dt$$

ein Skalarprodukt erklärt.

In einem euklidischen Vektorraum V können wir nun die gewünschten Begriffe Länge, Abstand und Orthogonalität einführen.

Definitionen und Bemerkungen. Es sei $(V, <\cdot,\cdot>)$ ein euklidischer Vektorraum.

(a) $\| x \| := \sqrt{<x,x>}$ heißt die *Länge* oder *Norm* von $x \in V$. Die Abbildung

$$\begin{aligned} \|\cdot\| : V &\longrightarrow \mathbb{R} \\ x &\longmapsto \| x \| \end{aligned}$$

heißt *Norm* . Die Norm $\|\cdot\|$ hat, wie man unmittelbar aus der Definition erkennt, folgende Eigenschaften:

(i) $\| x \| \geq 0$ für alle $x \in V$ und $\| x \| = 0 \iff x = o$,

(ii) $\| c\, x \| = | c | \, \| x \|$ für alle $x \in V$ und alle $c \in \mathbb{R}$.

Eine dritte Eigenschaft, die Minkowski-Ungleichung, werden wir in dem folgenden Satz 1 beweisen.

(b) $d(x,y) := \| x - y \|$ heißt *Abstand* oder *Distanz* von $x,y \in V$. Die Abbildung

$$\begin{aligned} d : V \times V &\longrightarrow \mathbb{R} \\ (x,y) &\longmapsto d(x,y) \end{aligned}$$

heißt *Metrik*. Für die Metrik d gilt:

(i) $d(x,y) \geq 0$ für alle $x,y \in V$ und $d(x,y) = 0 \iff x = y$,

(ii) $d(x,y) = d(y,x)$ für alle $x,y \in V$.

Weiterhin erfüllt die Metrik d die sogenannte Dreiecksungleichung, die wir ebenfalls in Satz 1 beweisen werden.

(c) x und y heißen *orthogonal*, wenn $<x,y> = 0$ gilt; Schreibweise : $x \perp y$.

Sind $A,B \subset V$, so nennen wir A und B *orthogonal*, bzw. A *orthogonal* zu B bzw. B *orthogonal* zu A, wenn $x \perp y$ für alle $x \in A$, $y \in B$ gilt; Schreibweise : $A \perp B$. Statt $\{x\} \perp B$ schreiben wir auch kürzer $x \perp B$.

Für $A \subset V$ sei

$$A^{\perp} := \{x \in V \mid x \perp A\}.$$

$A^{\perp}$ heißt *orthogonales Komplement* von A.

Für alle Teilmengen $A \subset V$ gilt:

(i) $A^{\perp}$ ist ein Untervektorraum von V,

(ii) $A^{\perp} = [A]^{\perp}$,

(iii) $A^{\perp} \cap [A] = \{o\}$,

(iv) $[A] \subset (A^{\perp})^{\perp}$.

Satz 1. *Es seien V ein euklidischer Vektorraum und x, y, $z \in V$. Dann gilt:*

(a) $|<x,y>| \leq \| x \| \cdot \| y \|$. (*"Cauchy–Schwarzsche Ungleichung"*)

Gleichheit tritt genau dann ein, wenn x und y linear abhängig sind.

(b) $\| x + y \| \leq \| x \| + \| y \|$. (*"Minkowskische Ungleichung"*)

(c) $d(x,y) \leq d(x,z) + d(z,y)$. (*"Dreiecksungleichung"*)

(d) $\| x + y \|^2 + \| x - y \|^2 = 2 \| x \|^2 + 2 \| y \|^2$. (*"Parallelogrammidentität"*)

(e) $4 <x,y> = \| x + y \|^2 - \| x - y \|^2$.

(f) $x \perp y \iff \| x+y \|^2 = \| x \|^2 + \| y \|^2$. (*"Satz von Pythagoras"*)

Beweis. (a) Für $y = o$ gilt Gleichheit und x und y sind linear abhängig. Sei nun $y \neq o$. Für alle $t \in \mathbb{R}$ und alle $x, y \in V$ gilt

$$(*) \qquad 0 \leq <x + t\,y, x + t\,y> \;=\; <x,x> + 2\,t <x,y> + t^2 <y,y> .$$

Setzen wir speziell $t = -\frac{<x,y>}{<y,y>}$, so erhalten wir

$$0 \leq <x,x> - \frac{<x,y>^2}{<y,y>} ,$$

woraus die behauptete Ungleichung folgt.

Gilt in der Cauchy–Schwarzschen Ungleichung das Gleichheitszeichen, so gilt in (*) für $t = -\frac{<x,y>}{<y,y>}$ das Gleichheitszeichen, woraus $x = -t\,y$ folgt. Die Umkehrung ist trivial.

(b) Mit (a) folgt

$$\| x+y \|^2 = \| x \|^2 + 2 <x,y> + \| y \|^2$$
$$\leq \| x \|^2 + 2 \| x \| \| y \| + \| y \|^2 = (\| x \| + \| y \|)^2.$$

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

(d) $$\| x+y \|^2 + \| x-y \|^2 = <x+y, x+y> + <x-y, x-y>$$
$$= 2 <x,x> + 2 <y,y> = 2 (\| x \|^2 + \| y \|^2).$$

(e) $$\| x+y \|^2 - \| x-y \|^2 = 4 <x,y> \qquad \text{(wie in (d)).}$$

(f) $$x \perp y \iff <x,y> = 0 \iff \| x+y \|^2 = \| x \|^2 + \| y \|^2$$ ∎

Bemerkungen. (a) Man kann auf einem reellen Vektorraum V eine *Norm* $\|\cdot\|$ einführen als Abbildung $\|\cdot\| : V \longrightarrow \mathbb{R}$, die folgende Bedingungen erfüllt:

(i) Für alle $x \in V$ gilt: $\| x \| \geq 0$ und $\| x \| = 0 \iff x = o$.

(ii) Für alle $x \in V$ und alle $c \in \mathbb{R}$ gilt: $\| c\,x \| = | c | \, \| x \|$.

(iii) Für alle $x,y \in V$ gilt: $\| x+y \| \leq \| x \| + \| y \|$.

$(V, \|\cdot\|)$ heißt dann *normierter Raum* .

Es gibt normierte Räume, bei denen die Norm nicht von einem Skalarprodukt $<\cdot,\cdot>$ induziert wird. Eine Norm wird genau dann von einem Skalarprodukt induziert, d.h. es gilt für alle $x \in V$, daß $\| x \| = \sqrt{<x,x>}$, wenn die Parallelogrammidentität (d) erfüllt ist. (Übungsaufgabe). Die Aussage (e) zeigt, daß in diesem Fall das Skalarprodukt eindeutig durch die Norm $\|\cdot\|$ bestimmt ist.

(b) Man kann auf einem Vektorraum V oder allgemeiner auf einer beliebigen Menge $V \neq \emptyset$ eine *Metrik* d einführen als Abbildung $d : V \times V \longrightarrow \mathbb{R}$, die den folgenden drei Bedingungen genügt:

(i) Für alle $x,y \in V$ gilt: $d(x,y) \geq 0$ und $d(x,y) = 0 \iff x = y$.

(ii) Für alle $x,y \in V$ gilt: $d(x,y) = d(y,x)$.

(iii) Für alle $x,y,z \in V$ gilt: $d(x,y) \leq d(x,z) + d(z,y)$.

(V,d) heißt dann *metrischer Raum.*

Es gibt metrische Räume, die keine Vektorräume sind, und selbst wenn V ein reeller Vektorraum ist, muß die Metrik d nicht von einer Norm und damit erst recht nicht von einem Skalarprodukt erzeugt werden.

(c) Die Cauchy–Schwarzsche Ungleichung gibt uns jetzt die Möglichkeit, in euklidischen Vektorräumen den Winkel ω zwischen zwei Vektoren x und y zu erklären. Dabei setzen wir sinnvollerweise $x \neq o$ und $y \neq o$ voraus. Dann ist

$$-1 \leq \frac{<x,y>}{\|x\| \, \|y\|} \leq 1 .$$

Damit existiert genau ein $\omega \in [0,\pi]$ mit

$$\cos \omega = \frac{<x,y>}{\|x\| \, \|y\|} .$$

Diese Zahl ω heißt *Winkel* zwischen x und y.

In der linearen Algebra ist der Winkelbegriff nicht so wichtig wie der Begriff der Orthogonalität. Wir werden deshalb im folgenden hauptsächlich letzteren untersuchen.

Mit Hilfe von Satz 1 wollen wir noch ein weiteres, wichtiges Beispiel für einen euklidischen Vektorraum angeben.

Beispiel. Es sei $V \subset \mathbb{R}^{\mathbb{N}_0}$ die Menge der reellen Folgen $x = (x_0, x_1, \ldots)$, die quadratisch summierbar sind, für die also $\sum_{i=0}^{\infty} x_i^2 < \infty$ gilt.

V ist ein Untervektorraum: Seien $x \in V$, $a \in \mathbb{R}$. Dann ist $a\,x = (a\,x_0, a\,x_1, \ldots)$, und es gilt

$$\sum_{i=0}^{\infty} a^2 x_i^2 < \infty ,$$

also ist $a\,x \in V$.

Aus $x = (x_0, x_1, \ldots)$, $y = (y_0, y_1, \ldots) \in V$ folgt mit der Minkowskischen Ungleichung für den Standardraum $\mathbb{R}^{m+1}$

$$\Big(\sum_{i=0}^{m} (x_i + y_i)^2\Big)^{\frac{1}{2}} \leq \Big(\sum_{i=0}^{m} x_i^2\Big)^{\frac{1}{2}} + \Big(\sum_{i=0}^{m} y_i^2\Big)^{\frac{1}{2}} \leq \Big(\sum_{i=0}^{\infty} x_i^2\Big)^{\frac{1}{2}} + \Big(\sum_{i=0}^{\infty} y_i^2\Big)^{\frac{1}{2}} =: c$$

mit $c \in \mathbb{R}$. Damit gilt für alle $m \in \mathbb{N}_0$

$$\sum_{i=0}^{m} (x_i + y_i)^2 \leq c^2$$

und somit

$$\sum_{i=0}^{\infty} (x_i + y_i)^2 < \infty ,$$

also $x + y \in V$.

Für $x, y \in V$ setzen wir nun

$$<x,y> := \sum_{i=0}^{\infty} x_i\, y_i\,.$$

Diese Reihe ist absolut konvergent, denn mit der Cauchy-Schwarzschen Ungleichung, angewendet im $\mathbb{R}^{m+1}$ erhalten wir für alle $m \in \mathbb{N}_0$

$$\sum_{i=0}^{m} |x_i\, y_i| = \sum_{i=0}^{m} |x_i| \cdot |y_i| \leq \Big(\sum_{i=0}^{m} x_i^2\Big)^{\frac{1}{2}} \cdot \Big(\sum_{i=0}^{m} y_i^2\Big)^{\frac{1}{2}}$$

$$\leq \Big(\sum_{i=0}^{\infty} x_i^2\Big)^{\frac{1}{2}} \cdot \Big(\sum_{i=0}^{\infty} y_i^2\Big)^{\frac{1}{2}} =: c$$

mit $c \in \mathbb{R}$ und somit

$$\sum_{i=0}^{\infty} |x_i\, y_i| \leq c\,.$$

Also ist die obige Reihe absolut konvergent und daher auch konvergent.

Damit ist $<\cdot,\cdot>$ offensichtlich ein Skalarprodukt auf V. Der Vektorraum V ist ein wichtiger unendlich dimensionaler euklidischer Vektorraum in der Mathematik. Er heißt *Standard-Hilbertraum* und wird üblicherweise mit l^2 bezeichnet. Der Vektorraum $\mathbb{R}[X]$ der reellen Polynome ist ein Untervektorraum von l^2.

Zum Abschluß dieses Paragraphen wollen wir überlegen, wie man Skalarprodukte auf endlich dimensionalen reellen Vektorräumen V bestimmen kann.

Es sei jetzt also $\dim V = n$ und $B = (v_1, \ldots, v_n)$ sei eine festgewählte Basis von V. Ist $\beta : V \times V \longrightarrow \mathbb{R}$ ein Skalarprodukt, so gilt wegen der Bilinearität von β für alle $x = \sum_{i=1}^{n} x_i\, v_i$, $x_i \in \mathbb{R}$, und alle $y = \sum_{j=1}^{n} y_j\, v_j$, $y_j \in \mathbb{R}$

$$\beta(x,y) = \sum_{i,\,j=1}^{n} x_i\, \beta(v_i, v_j)\, y_j = \hat{x}^{\mathsf{T}} A\, \hat{y}\,,$$

wo $\hat{x}$ und $\hat{y}$ die Koordinatenvektoren von x und y sind und $A = ((\beta(v_i,v_j)))$ gesetzt ist.

β ist also eindeutig festgelegt durch die Matrix A. Weil β symmetrisch ist, ist auch A symmetrisch. Weil β positiv definit ist, gilt für alle $\hat{x} \in \mathbb{R}^n$, $\hat{x} \neq o$,

$$(*) \qquad \hat{x}^{\mathsf{T}} A \hat{x} > 0 .$$

Eine symmetrische Matrix A, die (*) erfüllt, wollen wir ebenfalls *positiv definit* nennen.

Es sei nun umgekehrt $A \in \mathbb{R}^{n\times n}$ eine symmetrische, positiv definite Matrix. Dann wird durch

$$\beta(x,y) = \hat{x}^{\mathsf{T}} A \hat{y} ,$$

wo $\hat{x}$ und $\hat{y}$ die Koordinatendarstellungen von x und y bezüglich der festen Basis B sind, offensichtlich ein Skalarprodukt auf V erklärt. Wir haben also den folgenden Zusammenhang gezeigt:

Satz 2. *Es seien V ein reeller Vektorraum und B eine geordnete Basis von V. Dann ist $\beta: V\times V \longrightarrow \mathbb{R}$ genau dann ein Skalarprodukt, wenn eine symmetrische, positiv definite Matrix $A \in \mathbb{R}^{n\times n}$ existiert mit*

$$\beta(x,y) = \hat{x}^{\mathsf{T}} A \hat{y} , \qquad x,y \in V .$$

Dabei sind $\hat{x}$ und $\hat{y}$ die Koordinatenvektoren von x und y bezüglich B.

Es bleibt noch die Frage zu klären, wie man die positive Definitheit einer symmetrischen Matrix A nachprüfen kann. Wir zeigen dazu zwei Kriterien, benötigen aber zuvor eine Diagonalisierbarkeitsaussage für symmetrische Matrizen.

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

(a) *A ist diagonalisierbar.*

(b) *Eigenvektoren von A zu verschiedenen Eigenwerten sind orthogonal (bezüglich des Standardskalarproduktes im $\mathbb{R}^n$).*

Beweis. (a) Wir fassen A zunächst als komplexe Matrix auf. Sei $c = a + i\,b$, $a,b \in \mathbb{R}$, komplexe Nullstelle des charakteristischen Polynoms p von A und $z = u + i\,v$ ein zugehöriger Eigenvektor aus $\mathbb{C}^n$, $A\,z = c\,z$. Dann ist

$$A\,u = a\,u - b\,v\;,\;A\,v = b\,u + a\,v\,.$$

Damit folgt

$$(A\,u)^{\mathsf{T}}\,v = a\,u^{\mathsf{T}}v - b\,v^{\mathsf{T}}v$$

und

$$(A\,u)^{\mathsf{T}}\,v = u^{\mathsf{T}}A\,v = b\,u^{\mathsf{T}}u + a\,u^{\mathsf{T}}v\,.$$

Also ist $b\,\|\,u\,\|^2 = -\,b\,\|\,v\,\|^2$, woraus wegen $z \neq o$ unmittelbar $b = 0$ folgt. Das charakteristische Polynom p zerfällt also in Linearfaktoren mit lauter reellen Nullstellen $c_1,\dots,c_k$, woraus $\mathbb{R}^n = H_{c_1} \oplus\cdots\oplus H_{c_k}$ und $m = (X - c_1)^{s_1} \cdots (X - c_k)^{s_k}$, s_i Index von H_{c_i}, folgen.

Wir zeigen jetzt, daß der Index s_i von H_{c_i} eins ist: Sei $(A - c_i\,E_n)^2\,y = o$. Dann gilt

$$0 = y^{\mathsf{T}}(A - c_i\,E_n)^2\,y = [(A - c_i\,E_n)\,y]^{\mathsf{T}}[(A - c_i\,E_n)\,y]\;,$$

also $(A - c_i\,E_n)\,y = o$ und somit $s_i = 1$. Nach Satz 4.19 ist A daher diagonalisierbar.

(b) Aus $A\,u = c\,u$ und $A\,v = d\,v$ mit $u \neq o$, $v \neq o$ und $c,d \in \mathbb{R}$, $c \neq d$, folgt

$$c\,u^{\mathsf{T}}v = (A\,u)^{\mathsf{T}}\,v = u^{\mathsf{T}}A\,v = d\,u^{\mathsf{T}}v$$

und daraus $\langle u,v\rangle = u^{\mathsf{T}}v = 0$. ∎

Satz 4. *Es sei $A \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix. Dann ist A genau dann positiv definit, wenn alle Eigenwerte von A positiv sind.*

Beweis. Es seien A positiv definit, $c \in \mathbb{R}$ Eigenwert und x Eigenvektor von A. Dann gilt wegen $x \neq o$

$$0 < x^{\mathsf{T}}A\,x = x^{\mathsf{T}}(c\,x) = c\,x^{\mathsf{T}}x = c\,(x_1^2 + \cdots + x_n^2)\,,$$

also folgt $c > 0$.

Seien umgekehrt $c_1,\ldots,c_k$ die Eigenwerte von A und es gelte $c_i > 0$ für alle $i = 1,\ldots,k$. Nach Satz 3 (a) ist A diagonalisierbar, d.h. es gilt $\mathbb{R}^n = E_{c_1} \oplus \cdots \oplus E_{c_k}$. Zu jedem $x \in \mathbb{R}^n$ gibt es also eine eindeutige Zerlegung $x = x_1 + \cdots + x_k$ mit $x_i \in E_{c_i}$. Damit folgt

$$x^\mathsf{T} A\, x \;=\; \sum_{i=1}^{k} x_i^\mathsf{T} A\, x_i \;+\; \sum_{\substack{i,\, j=1 \\ i \neq j}}^{k} x_i^\mathsf{T} A\, x_j \;=\; \sum_{i=1}^{k} c_i \,\| x_i \|^2 \;+\; \sum_{\substack{i,\, j=1 \\ i \neq j}}^{k} c_j <x_i, x_j> ,$$

also wegen Satz 3 (b)

$$x^\mathsf{T} A\, x \;=\; \sum_{i=1}^{k} c_i \,\| x_i \|^2 .$$

Nun ist für $x \neq o$ mindestens ein $x_i \neq o$. Wegen $c_i > 0$ für $i = 1,\ldots,k$ folgt also insgesamt, daß $x^\mathsf{T} A\, x > 0$. ∎

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

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

mit dem charakteristischen Polynom

$$p = -\,(X-3)\,\left(X - \tfrac{1}{2}\,(3+\sqrt{5})\right)\left(X - \tfrac{1}{2}\,(3-\sqrt{5})\right).$$

Alle Eigenwerte sind positiv, also ist A positiv definit.

(b) Die Matrix

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

mit dem charakteristischen Polynom $p = X^2\,(X-2)^2$ besitzt den Eigenwert 0, ist also nicht positiv definit.

Oft ist es schwierig, die Eigenwerte einer Matrix A auszurechnen. Dann ist eventuell das folgende Kriterium nützlicher.

Satz 5. *Für eine symmetrische Matrix $A = (\!(a_{ij})\!) \in \mathbb{R}^{n\times n}$ sind folgende Aussagen äquivalent:*

(a) *A ist positiv definit.*

(b) *Es existiert eine reguläre Matrix $B \in \mathbb{R}^{n\times n}$ mit $A = B^{\mathsf{T}} B$.*

(c) *Alle Hauptunterdeterminanten von A sind positiv, d.h. für $k = 1,\ldots,n$ gilt*

$$\begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix} > 0 .$$

Bemerkung. Man nennt die Determinanten in (c) auch die *Hauptminoren* der Matrix A.

Beweis. (a) $\Longrightarrow$ (c): Ist A positiv definit, so ist auch jede der Matrizen

$$A_k = \begin{bmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{bmatrix} , \qquad k = 1,\ldots,n,$$

positiv definit ($A_n = A$). Ist nämlich $(x_1,\ldots,x_k) \neq (0,\ldots,0)$, so ist $x = (x_1,\ldots,x_k,0,\ldots,0) \neq o$, und es gilt

$$(x_1 \cdots x_k)\, A_k \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix} = x^{\mathsf{T}} A\, x > 0 .$$

Nach Satz 3 (a) ist jede der Matrizen A_k ähnlich zu einer Diagonalmatrix D_k:

$$A_k = S_k^{-1} D_k S_k .$$

Nach Satz 4 sind die Eigenwerte von A_k alle positiv, also gilt für $k = 1,\ldots,n$

$$\det A_k = \det D_k > 0 .$$

(c) $\Longrightarrow$ (b): Für $k = 2,\ldots,n$ schreiben wir die Matrizen A_k in der Form

$$A_k = \begin{bmatrix} A_{k-1} & y_k \\ y_k^{\mathsf T} & a_{kk} \end{bmatrix}$$

mit $y_k^{\mathsf T} = (a_{k1} \ldots a_{k,k-1})$. Weiterhin sei für $k = 2,\ldots,n$

$$T_k = \begin{bmatrix} E_{k-1} & -A_{k-1}^{-1}\, y_k \\ o^{\mathsf T} & 1 \end{bmatrix}$$

Hierbei haben wir benutzt, daß A_{k-1} wegen $\det A_{k-1} > 0$ invertierbar ist. Es folgt

$$T_k^{\mathsf T}\, A_k\, T_k = \begin{bmatrix} A_{k-1} & o \\ o^{\mathsf T} & a_{kk} - y_k^{\mathsf T}\, A_{k-1}^{-1}\, y_k \end{bmatrix}$$

Sei $b_k := a_{kk} - y_k^{\mathsf T}\, A_{k-1}^{-1}\, y_k$. Wegen $\det T_k = \det T_k^{\mathsf T} = 1$ folgt

$$\det A_k = \det (T_k^{\mathsf T}\, A_k\, T_k) = \det A_{k-1} \cdot b_k ,$$

also

$$b_k = \frac{\det A_k}{\det A_{k-1}} > 0 , \qquad k = 2,\ldots,n .$$

Für die (n,n)–Matrix

$$T := T_n \begin{bmatrix} T_{n-1} & O \\ O & 1 \end{bmatrix} \begin{bmatrix} T_{n-2} & O \\ O & E_2 \end{bmatrix} \cdots \begin{bmatrix} T_2 & O \\ O & E_{n-2} \end{bmatrix}$$

gilt nun nach Konstruktion, daß sie regulär ist und daß

$$T^{\mathsf{T}} A\, T = \begin{bmatrix} a_{11} & & & O \\ & b_2 & & \\ & & \ddots & \\ O & & & b_n \end{bmatrix}$$

Setzen wir noch

$$D = \begin{bmatrix} \sqrt{a_{11}} & & & O \\ & \sqrt{b_2} & & \\ & & \ddots & \\ O & & & \sqrt{b_n} \end{bmatrix}$$

und $B = D\, T^{-1}$, so folgt $B^{\mathsf{T}} B = (T^{-1})^{\mathsf{T}} D^2\, T^{-1} = A$ und $B \in \mathbb{R}^{n\times n}$ ist regulär.

(b) $\Longrightarrow$ (a): Sei $x \neq o$. Dann ist $B\,x \neq o$ und es gilt $x^{\mathsf{T}} A\, x = x^{\mathsf{T}} B^{\mathsf{T}} B\, x = (B\,x)^{\mathsf{T}} (B\,x) > 0$. Also ist A positiv definit. ∎

Bemerkung. Der Beweis zeigt, daß die reguläre Matrix B sogar eine obere Dreiecksmatrix ist. Die Zerlegung der positiv definiten Matrix A in der Form $A = B^{\mathsf{T}} B$ mit einer regulären Dreicksmatrix B heißt *Cholesky- Zerlegung*, sie spielt eine wichtige Rolle in der numerischen Mathematik.

Beispiel. Für welche $a \in \mathbb{R}$ ist

$$A = \begin{bmatrix} 1 & a & -1 \\ a & 9 & 0 \\ -1 & 0 & 4 \end{bmatrix}$$

positiv definit? Es gilt $p = -X^3 + 14\,X^2 - (48 - a^2)\,X + 27 - 4\,a^2$. Hier die Nullstellen in Abhängigkeit von a bestimmen zu wollen, ist sicher nicht einfach. Wir wenden deshalb Satz 5 an und erhalten

$$\det A_1 = 1 > 0\ , \ \det A_2 = 9 - a^2 > 0 \iff |\,a\,| < 3\,,$$

$$\det A_3 = \det A = 27 - 4\,a^2 > 0 \iff |\,a\,| < \tfrac{3}{2}\sqrt{3}\,.$$

Somit ist A genau dann positiv definit, wenn $-\frac{3}{2}\sqrt{3} < a < \frac{3}{2}\sqrt{3}$ gilt.

§ 2 Orthonormalbasen und Orthogonalprojektionen

Da wir in euklidischen Vektorräumen die Orthogonalität von Vektoren eingeführt haben, liegt es nahe, nach Basen zu fragen, deren Vektoren paarweise orthogonal sind. Solche "rechtwinkligen Koordinatensysteme" spielen ja auch im Anschauungsraum eine wichtige Rolle bei der Beschreibung geometrischer Phänomene.

Definition. Es sei V ein euklidischer Vektorraum. Eine nichtleere Teilmenge $A \subset V$ heißt *Orthogonalsystem*, wenn gilt:

(a) $o \notin A$,

(b) Je zwei verschiedene Vektoren $x,y \in A$ sind orthogonal.

Gilt außerdem noch

(c) $\| x \| = 1$ für alle $x \in A$,

so heißt A *Orthonormalsystem.* Ist zusätzlich A Basis von V, so heißt A *Orthogonal*– bzw. *Orthonormalbasis* von V. Für letztere schreiben wir kurz ONB.

Bemerkung. Eine endliche Basis $B = (x_1,...,x_n)$ bzw. eine abzählbare Basis $B = (x_1,x_2,...)$ eines euklidischen Vektorraums ist genau dann eine ONB, wenn für alle $i,j \in \{1,...,n\}$ bzw. für alle $i,j \in \mathbb{N}$ gilt: $<x_i,x_j> = \delta_{ij}$. Wie hier werden wir auch in Zukunft bei endlichen bzw. abzählbaren Basen meistens davon ausgehen, daß sie geordnet sind.

Satz 6. *Es seien V ein euklidischer Vektorraum und $A \subset V$ ein Orthogonalsystem. Dann ist A linear unabhängig.*

Beweis. Seien $x_1,...,x_k \in A$ paarweise verschieden und $a_1 x_1 + \cdots + a_k x_k = o$, $a_i \in \mathbb{R}$, $i = 1,...,k$. Dann folgt für $j = 1,...,k$:

$$0 = <a_1 x_1 + \cdots + a_k x_k , x_j> = a_1 <x_1,x_j> + \cdots + a_k <x_k,x_j> = a_j \| x_j \|^2 ,$$

also $a_j = 0$. Somit ist A linear unabhängig. ∎

Ein Orthogonalsystem ist also genau dann eine Orthogonalbasis, wenn es V erzeugt. In unendlich dimensionalen Vektorräumen kann es Orthogonalsysteme geben, die keine Basen sind, aber dennoch maximal sind, d.h. nicht durch Hinzufügen von weiteren Vektoren vergrößert werden können. In solchen euklidischen Vektorräumen gibt es keine Orthonormalbasis. Für endlich dimensionale euklidische Vektorräume kann dieser Fall jedoch nicht eintreten, wie der folgende Satz zeigt.

Satz 7. *Es sei V ein endlich dimensionaler euklidischer Vektorraum. Dann existiert eine Orthonormalbasis B von V.*

Bemerkung. Wir werden nicht nur die Existenz von B nachweisen, sondern auch ein Konstruktionsverfahren angeben, mit dem man aus einer beliebigen geordneten Basis B' von V eine ONB erhalten kann. Das Verfahren heißt *Gram–Schmidtsches Orthogonalisierungsverfahren.*

Beweis. Sei $\dim V = n$ und $B' = (x_1,\dots,x_n)$ eine Basis von V. Wir definieren zunächst induktiv eine Orthogonalbasis $\overline{B} = (y_1,\dots,y_n)$ von V und normieren die Basisvektoren anschließend. Sei

$$
\begin{aligned}
y_1 &:= x_1\,, \\
y_2 &:= x_2 - \frac{<x_2,y_1>}{\|\, y_1 \,\|^2}\, y_1\,, \\
&\;\vdots \\
y_{k+1} &:= x_{k+1} - \sum_{i=1}^{k} \frac{<x_{k+1},y_i>}{\|\, y_i \,\|^2}\, y_i\,, \qquad k = 1,\dots,n-1.
\end{aligned}
$$

Diese Definition ist sinnvoll, da alle y_i von o verschieden sind. Offensichtlich ist $y_1 \neq o$. Wäre $y_{k+1} = o$ für ein $k \in \{1,\dots,n-1\}$, so wäre $x_{k+1} \in [y_1,\dots,y_k] \subset [x_1,\dots,x_k]$, ein Widerspruch.

Durch vollständige Induktion nach k zeigen wir nun, daß $(y_1,\dots,y_{k+1})$ ein Orthogonalsystem ist.

$$k = 1\colon\quad <y_1,y_2> \;=\; <x_1,x_2> \;-\; \frac{<x_2,x_1>}{<x_1,x_1>}\,<x_1,x_1> \;=\; 0\,.$$

$k \longrightarrow k+1$: Nach Induktionsvoraussetzung sind die Vektoren $y_1,\dots,y_k$ paarweise orthogonal, also müssen wir nur noch $<y_{k+1},y_j> = 0$ für $j = 1,\dots,k$ nachweisen:

$$<y_{k+1},y_j> \;=\; <x_{k+1},y_j> \;-\; \sum_{i=1}^{k} \frac{<x_{k+1},y_i>}{\| y_i \|^2} <y_i,y_j>$$

$$= <x_{k+1},y_j> \;-\; \frac{<x_{k+1},y_j>}{\| y_j \|^2} <y_j,y_j> \;=\; 0\,.$$

Also ist $\overline{B} = (y_1,\dots,y_n)$ ein Orthogonalsystem und daher nach Satz 5 linear unabhängig. Wegen $x_j \in [y_1,\dots,y_j]$, $j = 1,\dots,n$, ist $[\overline{B}] = V$. Daher ist $\overline{B}$ Orthogonalbasis und somit ist

$$B \;=\; \left(\frac{y_1}{\|y_1\|},\dots,\frac{y_n}{\|y_n\|}\right)$$

eine ONB von V. ■

Bemerkungen. **(a)** Der Beweis zeigt, daß das Orthogonalisierungsverfahren auch dann funktioniert, wenn V unendlich dimensional ist und in V eine abzählbare Basis $B' = (x_1,x_2,\dots)$ existiert.

(b) Die Aussage von Satz 7 folgt auch aus Satz 5 :

Beweis. Sei o.B.d.A. $V = \mathbb{R}^n$, und sei β ein Skalarprodukt auf $\mathbb{R}^n$. Bezüglich der Standardbasis gilt dann für alle $x,y \in \mathbb{R}^n$

$$\beta\,(x,y) = x^{\mathsf{T}} A\, y$$

mit einer symmetrischen positiv definiten Matrix A. Nach Satz 5 existiert eine reguläre Matrix $B \in \mathbb{R}^{n\times n}$ mit $A = B^{\mathsf{T}}B$. Somit gilt $\beta\,(x,y) = (B\,x)^{\mathsf{T}}\,(B\,y)$, also ist $(B^{-1}e_1,\dots,B^{-1}e_n)$ eine Orthonormalbasis bezüglich β. ■

Allerdings ergibt sich so kein einfaches Rechenverfahren zur Bestimmung einer Orthonormalbasis.

Beispiele. **(a)** Im $\mathbb{R}^n$ ist die Standardbasis $B = (e_1,\dots,e_n)$ ONB bezüglich des Standardskalarprodukts $<\cdot,\cdot>$.

(b) In dem Vektorraum $\mathbb{R}[X]$, aufgefaßt als Untervektorraum des Hilbertraumes l^2, ist $A = \{e_i |\ i \in \mathbb{N}_0\}$ mit

$$e_i = (0,\dots,0,\underbrace{1}_{\text{i-te Stelle}},0,\dots)$$

eine ONB. In l^2 ist A ein Orthonormalsystem.

(c) Wir wollen im $\mathbb{R}^3$ von der Basis

$$B' = \left(\begin{bmatrix}1\\1\\0\end{bmatrix},\begin{bmatrix}1\\1\\1\end{bmatrix},\begin{bmatrix}1\\0\\1\end{bmatrix}\right)$$

ausgehend eine ONB bezüglich des Standardskalarproduktes bestimmen:

$$y_1 = x_1 = \begin{bmatrix}1\\1\\0\end{bmatrix}, \ \| y_1 \| = \sqrt{2},$$

$$y_2 = x_2 - \frac{<x_2,y_1>}{\| y_1 \|^2} y_1 = \begin{bmatrix}1\\1\\1\end{bmatrix} - \frac{2}{2}\begin{bmatrix}1\\1\\0\end{bmatrix} = \begin{bmatrix}0\\0\\1\end{bmatrix}, \ \| y_2 \| = 1,$$

$$y_3 = x_3 - \frac{<x_3,y_1>}{\| y_1 \|^2} y_1 - \frac{<x_3,y_2>}{\| y_2 \|^2} y_2 = \begin{bmatrix}1\\0\\1\end{bmatrix} - \frac{1}{2}\begin{bmatrix}1\\1\\0\end{bmatrix} - \frac{1}{1}\begin{bmatrix}0\\0\\1\end{bmatrix} = \frac{1}{2}\begin{bmatrix}1\\-1\\0\end{bmatrix},$$

$$\| y_3 \| = \frac{1}{2}\sqrt{2}.$$

Also ist

$$B = \left(\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\\0\end{bmatrix}\right)$$

eine Orthonormalbasis von $\mathbb{R}^3$.

(d) Wir rechnen Beispiel (c) nochmals, wobei jetzt aber der $\mathbb{R}^3$ mit dem durch

$$\beta(x,y) = x^{\mathrm{T}} A\, y,$$

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

definierten Skalarprodukt β versehen ist:

$$y_1 = x_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \| y_1 \|^2 = \beta(y_1, y_1) = 4 ,$$

$$y_2 = x_2 - \frac{\beta(x_2, y_1)}{\| y_1 \|^2} y_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \tfrac{3}{4} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \tfrac{1}{4} \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}, \quad \| y_2 \|^2 = \tfrac{7}{4},$$

$$y_3 = x_3 - \frac{\beta(x_3, y_1)}{\| y_1 \|^2} y_1 - \frac{\beta(x_3, y_2)}{\| y_2 \|^2} y_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \tfrac{0}{4} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \tfrac{1}{7} \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix} = \tfrac{1}{7} \begin{bmatrix} 6 \\ -1 \\ 3 \end{bmatrix},$$

$$\| y_3 \|^2 = \tfrac{3}{7}$$

Also ist

$$B = \left(\tfrac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \frac{1}{2\sqrt{7}} \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}, \frac{1}{\sqrt{21}} \begin{bmatrix} 6 \\ -1 \\ 3 \end{bmatrix} \right)$$

bezüglich β eine ONB von $\mathbb{R}^3$.

Bemerkungen. **(a)** Ist $B = (x_1, \ldots, x_n)$ eine ONB von V, so gilt für alle $x \in V$

$$x = \sum_{i=1}^{n} \langle x, x_i \rangle \, x_i ,$$

denn aus $x = a_1 x_1 + \cdots + a_n x_n$ folgt $\langle x, x_j \rangle = \sum_{i=1}^{n} a_i \langle x_i, x_j \rangle = a_j$ für alle $j = 1, \ldots, n$.

Damit folgt weiter für alle x, $y \in V$

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

$<x,y>$ stimmt also mit dem Standardskalarprodukt $\hat{x}^{\mathrm{T}}\,\hat{y}$ der Koordinatenvektoren $\hat{x}$ und $\hat{y}$ überein.

Daraus ergibt sich für alle $x \in V$

$$\| x \|^2 = \sum_{i=1}^{n} <x,x_i>^2 = \| \hat{x} \|^2 .$$

(b) Ist V ein euklidischer Vektorraum beliebiger Dimension und ist $A \subset V$ Orthonormalsystem, so gilt für alle $x \in V$

$$(*) \qquad \| x \|^2 \geq \sum_{y \in A} <x,y>^2 , \qquad (\text{"}\textit{Besselsche Ungleichung}\text{"})$$

wobei in der Summe jeweils höchstens abzählbar viele Summanden positiv sind (Beweis als Übungsaufgabe).

Statt nach einer ONB, die es im allgemeinen nicht gibt, sucht man dann nach Orthonormalsystemen A, die *vollständig* sind, d.h. für die in (*) Gleichheit gilt für alle $x \in V$ (*"Parsevalsche Gleichung"*).

Jedes vollständige Orthonormalsystem A ist maximal, denn aus $x \perp A$ folgt wegen der Parsevalschen Gleichung $\| x \|^2 = 0$, also $x = o$.

Beispiel. Sei $V = l^2$. Hier ist $A = \{e_i |\ i \in \mathbb{N}_0\}$ mit

$$e_i = (0,\ldots,0,\underbrace{1}_{\text{i-te Stelle}},0,\ldots)$$

ein Orthonormalsystem. Es ist wegen $[A] = \mathbb{R}[X]$ keine Basis von l^2, aber es ist vollständig, denn für alle $x \in l^2$ gilt:

$$\| x \|^2 = \sum_{i=0}^{\infty} x_i^2 = \sum_{i=0}^{\infty} <x,e_i>^2$$

Es seien nun V ein euklidischer Vektorraum, $U \subset V$ ein endlich dimensionaler Untervektorraum von V und $B = (x_1,\dots,x_k)$ eine ONB von U. Dann gilt für alle $u \in U$

$$u = \sum_{i=1}^{k} <u,x_i> x_i .$$

Setzen wir also für $x \in V$

$$\pi(x) := \sum_{i=1}^{k} <x,x_i> x_i ,$$

so erhalten wir eine lineare Abbildung $\pi : V \longrightarrow V$ mit Bild $\pi = U$ und $\pi|_U = \mathrm{id}_U$, d.h. $\pi^2 = \pi$. Somit ist π eine Projektion auf den Untervektorraum $U =$ Bild π. Weiterhin gilt

$$<\pi(x) - x\, ,\, \pi(x)> = <\pi(x),\pi(x)> - <\pi(x),x>$$

$$= \sum_{i=1}^{k} <x,x_i>^2 - \sum_{i=1}^{k} <x,x_i> <x,x_i> = 0,$$

also ist

$$(\pi(x) - x) \perp \pi(x) .$$

Eine solche Abbildung π wollen wir Orthogonalprojektion nennen.

Definition. Es seien V ein euklidischer Vektorraum und $\pi \in \mathrm{Hom}\,(V,V)$. Die Abbildung π heißt *Orthogonalprojektion* (auf $U =$ Bild π), wenn π Projektion ist, also $\pi^2 = \pi$ erfüllt, und wenn

$$(\pi(x) - x) \perp \pi(x)$$

für alle $x \in V$ gilt. Der Vektor $\pi(x)$ heißt dann die *Orthogonalprojektion* von x auf U.

Bemerkung. Ist π Orthogonalprojektion auf U, so gilt $(\pi(x) - x) \perp U$ und somit auch Kern $\pi \perp$ Bild π. Umgekehrt ist jede Projektion π, die Kern $\pi \perp$ Bild π erfüllt, eine Orthogonalprojektion.

Beweis. Es sei $u \in U$. Dann gilt $\pi(u - \pi(x) + x) = \pi(u) = u$ und daher $\pi(u - \pi(x) + x) - (u - \pi(x) + x) \perp u$. Also folgt $(\pi(x) - x) \perp u$.

Aus $(\pi(x) - x) \perp U$ erhalten wir unmittelbar Kern $\pi \perp U$, und umgekehrt ist jede Projektion π auf U mit dieser Eigenschaft wegen $\pi(x) - x \in$ Kern π eine Orthogonalprojektion. ∎

Nicht jeder Untervektorraum U besitzt eine Orthogonalprojektion. Wann dies der Fall ist, besagt der folgende Satz.

Satz 8. *Es seien V ein euklidischer Vektorraum und U ein Untervektorraum von V. Genau dann existiert eine Orthogonalprojektion auf U, wenn $V = U \oplus U^{\perp}$ gilt. Ist dies der Fall, so gibt es genau eine Orthogonalprojektion auf U und es gilt $U = (U^{\perp})^{\perp}$.*

Beweis. Sei π eine Orthogonalprojektion auf U. Für $x \in V$ gilt dann

$$x = \pi(x) + (x - \pi(x)) \in U + U^{\perp}.$$

Somit ist $V = U + U^{\perp}$, und wegen $U \cap U^{\perp} = \{o\}$ ist diese Summe direkt.

Sei nun umgekehrt $V = U \oplus U^{\perp}$. Jeder Vektor $x \in V$ besitzt somit eine eindeutige Zerlegung $x = u + v$, $u \in U$, $v \in U^{\perp}$. Die durch $\pi(x) = u$ erklärte Abbildung π ist dann eine Projektion von V auf U und es gilt $<\pi(x) - x, \pi(x)> = -<v,u> = 0$. Also ist π Orthogonalprojektion auf U.

Eindeutigkeit von π: Ist $u = \pi(x)$, und ist $u' = \pi'(x)$ ebenfalls Orthogonalprojektion von x auf U, so folgt

$$\pi(x) - x \perp U \quad \text{und} \quad \pi'(x) - x \perp U$$

also

$$0 = <u - x, u - u'> \quad \text{und} \quad 0 = <u' - x, u - u'> .$$

Damit ergibt sich durch Subtraktion $0 = <u - u', u - u'>$, also $u = u'$.

Seien nun $V = U \oplus U^{\perp}$ und $x \in (U^{\perp})^{\perp}$. Dann gilt $x = u + v$ mit Vektoren $u \in U$ und $v \in U^{\perp}$. Daraus folgt $<v,v> \; = \; <x,v> - <u,v> \; = \; 0$. Also ist $v = o$ und somit $x \in U$. Da umgekehrt stets $U \subset (U^{\perp})^{\perp}$ gilt, folgt $U = (U^{\perp})^{\perp}$. ■

Daß die Bedingung $V = U \oplus U^{\perp}$ nicht immer erfüllt ist, wollen wir an einem Beispiel sehen.

Beispiel. Sei $V = l^2$ der Standardhilbertraum und $U = \mathbb{R}[X]$ sei der Untervektorraum der Polynome. Dann ist $U^{\perp} = \{o\}$, also $V \neq U \oplus U^{\perp}$.

Beweis. Sei $x = (x_0, x_1, \ldots) \in U^{\perp}$. Dann ist $<x,y> \; = 0$ für alle $y \in U$ und somit $<x,e_i> \; = 0$ für alle $i \in \mathbb{N}_0$ mit

$$e_i = (0,\ldots,0,\underbrace{1}_{i\text{-te Stelle}},0,\ldots) \in U.$$

Daraus folgt $x_i = 0$ für alle i, also $x = o$. Somit ist $U^{\perp} = \{o\}$. ■

In dem euklidischen Vektorraum l^2 gibt es daher keine Orthogonalprojektion auf den Untervektorraum U.

Wie wir zu Beginn unserer Überlegungen über Projektionen gesehen haben, existiert π immer, wenn $\dim U < \infty$ ist. Wir wollen dies als Satz festhalten.

Satz 9. *Es seinen V ein euklidischer Vektorraum beliebiger Dimension und $U \subset V$ ein Untervektorraum endlicher Dimension. Dann existiert die Orthogonalprojektion π auf U, und es gilt $V = U \oplus U^{\perp}$ und $(U^{\perp})^{\perp} = U$.*

Beispiel. Sei V der Vektorraum $\mathbb{R}^5$, versehen mit dem Standardskalarprodukt, und seien

$$U = [\begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}] \quad \text{sowie} \quad x = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}.$$

Gesucht ist die Orthogonalprojektion von x auf U.

1. Schritt: Wir bestimmen eine ONB in U:

$$y_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad ; \quad \| y_1 \| = \sqrt{3},$$

$$y_2 = \begin{bmatrix} 2 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} - \frac{2+1}{3} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \; ; \; \| y_2 \| = \sqrt{3},$$

$$y_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \frac{2}{3} \begin{bmatrix} -1 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \; ; \; \| y_3 \| = \frac{2}{\sqrt{3}}.$$

Damit erhalten wir als ONB in U:

$$x_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \; x_2 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \; x_3 = \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}.$$

2. Schritt: Es ist $\pi(x) = <x,x_1> x_1 + <x,x_2> x_2 + <x,x_3> x_3 = \frac{1}{3} \begin{bmatrix} 3 \\ 5 \\ 8 \\ 13 \\ 0 \end{bmatrix}.$

Orthogonalprojektionen lassen sich auch durch eine Minimalitätseigenschaft des Abstandes charakterisieren.

Satz 10 und Definition. *Es seien V ein euklidischer Vektorraum, $U \subset V$ ein Untervektorraum und $\pi \in$* Hom (V, V) *mit* Bild $\pi \subset U$. *Dann ist π genau dann Orthogonalprojektion auf U, wenn für alle $x \in V$ und alle $u \in U$ gilt:*

$$\| x - \pi(x) \| \leq \| x - u \| .$$

Ist π Orthogonalprojektion auf U, so heißt $d(x,U) := \| x - \pi(x) \|$ der Abstand von x zu U.

Beweis. Sei π Orthogonalprojektion auf U. Dann gilt für $x \in V$, $u \in U$:

$$x - u = x - \pi(x) + \pi(x) - u .$$

Wegen $\pi(x) - u \in U$ ist $x - \pi(x) \perp \pi(x) - u$. Damit folgt mit dem Satz von Pythagoras

$$\| x - u \|^2 = \| x - \pi(x) \|^2 + \| \pi(x) - u \|^2 \geq \| x - \pi(x) \|^2 .$$

Umgekehrt gelte nun $\| x - \pi(x) \| \leq \| x - u \|$ für alle $x \in V$, $u \in U$. Dann folgt speziell für $x \in U$

$$\| x - \pi(x) \| \leq \| x - x \| = 0,$$

also $x = \pi(x)$. Somit ist wegen Bild $\pi \subset U$ $\pi^2 = \pi$ und Bild $\pi = U$.

Nun müssen wir noch die Orthogonalitätseigenschaft $\pi(x) - x \perp \pi(x)$ nachweisen. Für $\pi(x) = o$ ist dies trivial. Für $\pi(x) \neq o$ betrachten wir $(1 + a)\, \pi(x)$, $a \in \mathbb{R}$. Es gilt

$$\| x - \pi(x) \|^2 \leq \| x - (1 + a)\, \pi(x) \|^2 = \| x - \pi(x) - a\, \pi(x) \|^2$$

$$= \| x - \pi(x) \|^2 - 2\, a < x - \pi(x), \pi(x) > + a^2 \| \pi(x) \|^2 ,$$

also

$$2\, a < x - \pi(x), \pi(x) > \leq a^2 \| \pi(x) \|^2.$$

Wählen wir speziell

$$a = \frac{< x - \pi(x), \pi(x) >}{\| \pi(x) \|^2},$$

so erhalten wir

$$2\, a^2 \| \pi(x) \|^2 \leq a^2 \| \pi(x) \|^2.$$

Also ist $a = 0$ und somit $< x - \pi(x), \pi(x) > = 0$. ∎

Beispiel. In dem Beispiel von S. 243/244 ist $d\,(x,U) = \| \pi(x) - x \| = \frac{2}{3}\sqrt{57}$.

Bemerkungen. **(a)** Orthogonalprojektionen sind kontrahierend, d.h. sie verkleinern die Norm und den Abstand:

(i) $\| \pi(x) \| \leq \| x \|$, für alle $x \in V$

(ii) $\| \pi(x) - \pi(y) \| \leq \| x - y \|$, für alle $x,y \in V$,

Umgekehrt ist jede Projektion π auf U, die (i) oder (ii) erfüllt, Orthogonalprojektion (Übungsaufgabe).

(b) Geometrische Interpretation des Orthogonalisierungsverfahrens von Gram–Schmidt:

Es sei $(x_1,x_2,\ldots)$ ein System linear unabhängiger Vektoren und $(y_1,y_2,\ldots)$ sei das daraus konstruierte Orthogonalsystem. Bezeichnen wir für $k = 1,2,\ldots$ mit π_k die Orthogonalprojektion auf $U_k = [y_1,\ldots,y_k] = [x_1,\ldots,x_k]$, so gilt

$$y_{k+1} = x_{k+1} - \pi_k(x_{k+1}) .$$

Wendet man das Verfahren allgemeiner auf ein beliebiges System $(x_1,x_2,\ldots)$ an mit $x_i \neq o$, und ist $x_{k+1} \in [x_1,\ldots,x_k]$, so folgt wegen $\pi_k(x_{k+1}) = x_{k+1}$, daß $y_{k+1} = o$ ist. Läßt man bei dem modifizierten Verfahren aus dem System $(y_1,y_2,\ldots)$ alle Vektoren $y_i = o$ weg, so erhält man wieder ein Orthogonalsystem.

§ 3 Adjungierte Abbildungen

Wir betrachten zwei euklidische Vektorräume V, W sowie eine lineare Abbildung $\Phi \in \mathrm{Hom}(V,W)$. Mit Hilfe der Skalarprodukte in V und W wollen wir eine weitere lineare Abbildung $\Psi \in \mathrm{Hom}(W,V)$ finden, die eng mit Φ zusammenhängt.

Definition. Es seien V und W euklidische Vektorräume und $\Phi \in \mathrm{Hom}\,(V,W)$. Eine Abbildung $\Psi \in \mathrm{Hom}(W,V)$ heißt *adjungierte Abbildung* von Φ, wenn

$$(*) \qquad <\Phi(v),w> \; = \; <v,\Psi(w)>$$

für alle $v \in V$ und alle $w \in W$ gilt.

Bemerkungen. **(a)** Man beachte, daß in obiger Definition links das Skalarprodukt in W, rechts das in V steht.

(b) Es gibt höchstens eine zu Φ adjungierte Abbildung, wir bezeichnen sie mit Φ^*.

Beweis. Aus $<\Phi(v),w> \; = \; <v,\Psi_1(w)> \; = \; <v,\Psi_2(w)>$ folgt $<v,(\Psi_1 - \Psi_2)(w)> \; = 0$. Da hier $v \in V$ und $w \in W$ beliebig sind, ist $(\Psi_1 - \Psi_2)(w) = o$ für alle $w \in W$, d.h. $\Psi_1 = \Psi_2$. ■

(c) Existiert die adjungierte Abbildung Φ^* von Φ, so existiert auch $(\Phi^*)^*$, nämlich $(\Phi^*)^* = \Phi$.

(d) Nicht zu jedem Homomorphismus $\Phi \in \mathrm{Hom}(V,W)$ existiert die adjungierte Abbildung Φ^*.

Beispiel. Sei $V = \mathbb{R}[X]$, $W = l^2$ und $\Phi : \mathbb{R}[X] \longrightarrow l^2$, $v \longmapsto v$, die identische Abbildung. Würde die adjungierte Abbildung $\Phi^* : l^2 \longrightarrow \mathbb{R}[X]$ existieren, so müsste für alle $v \in \mathbb{R}[X]$, $w \in l^2$

$$<v,w> \; = \; <\Phi(v),w> \; = \; <v,\Phi^*(w)>$$

gelten. Somit wäre $w - \Phi^*(w) \in \mathbb{R}[X]^\perp = \{o\}$ für alle $w \in l^2$, also $\Phi^*(w) = w$ und damit $\mathbb{R}[X] = l^2$. Dies ist ein Widerspruch.

Um festzustellen, ob zu gegebenem Φ die adjungierte Abbildung Φ^* existiert, wäre es naheliegend, Φ^* durch die Gleichung (*) zu erklären, d.h. zu festem $w \in W$ das Bild $v = \Phi^*(w)$ so zu wählen, daß

$$<\Phi(x),w> \;=\; <x,v>$$

für alle $x \in V$ gilt. Nun ist $x \longmapsto <\Phi(x),w>$ bei festem $w \in W$ eine Linearform auf V, also ein Element x^* von V^* ; ebenso ist $x \longmapsto <x,v>$ eine Linearform. Unsere obige Definition von Φ^* wäre also sinnvoll, wenn zu jedem $x^* \in V^*$ ein $v \in V$ existierte mit

$$x^*(x) \;=\; <x,v> .$$

Wir werden das jetzt im Fall $\dim V < \infty$ zeigen, im Fall $\dim V = \infty$ ist es im allgemeinen falsch, wie sich aus dem obigen Beispiel ergibt.

Satz 11 (Rieszscher Darstellungssatz). *Es seien V ein euklidischer Vektorraum mit* $\dim V < \infty$ *und $x^* \in V^*$ eine Linearform. Dann existiert genau ein $v \in V$ mit*

$$x^*(x) \;=\; <x,v>$$

für alle $x \in V$.

Beweis. Seien $\dim V = n$ und $B = (x_1,\dots,x_n)$ eine ONB von V. Wir setzen

$$v \;:=\; \sum_{i=1}^{n} x^*(x_i)\, x_i .$$

Dann gilt

$$<x,v> \;=\; \sum_{i=1}^{n} x^*(x_i) <x,x_i> \;=\; x^*\Big(\sum_{i=1}^{n} <x,x_i>\, x_i\Big) \;=\; x^*(x)$$

für alle $x \in V$.

Erfüllt $v' \in V$ ebenfalls $x^*(x) = <x,v'>$ für alle $x \in V$, so folgt $<x,v - v'> = 0$ für alle $x \in V$, also $v - v' = o$. ■

Mit Hilfe von Satz 11 beweisen wir nun die Existenz der zu Φ adjungierten Abbildung Φ^*, falls V und W endlich dimensional sind.

Satz 12. *Es seien V, W endlich dimensionale euklidische Vektorräume und Φ aus* Hom (V, W). *Dann existiert die adjungierte Abbildung* $\Phi^* \in \operatorname{Hom}(W, V)$.

Beweis. Wie schon angedeutet, können wir für $w \in W$ den Bildvektor $\Phi^*(w)$ durch

$$<\Phi(x),w> = <x,\Phi^*(w)> \qquad \text{für alle } x \in V$$

erklären, weil nach Satz 11 ein solches Element $\Phi^*(w) := v$ existiert. Damit ist eine Abbildung Φ^*: $W \longrightarrow V$ erklärt. Es bleibt noch die Linearität von Φ^* nachzuweisen: Für alle $x \in V$ ist

$$<x,\Phi^*(a_1 w_1 + a_2 w_2)> = <\Phi(x), a_1 w_1 + a_2 w_2> = a_1 <\Phi(x),w_1> + a_2 <\Phi(x),w_2>$$

$$= a_1 <x,\Phi^*(w_1)> + a_2 <x,\Phi^*(w_2)> = <x,a_1 \Phi^*(w_1) + a_2 \Phi^*(w_2)> .$$

Also folgt für alle $w_1, w_2 \in W$ und alle $a_1, a_2 \in \mathbb{R}$

$$\Phi^*(a_1 w_1 + a_2 w_2) = a_1 \Phi^*(w_1) + a_2 \Phi^*(w_2) . \qquad \blacksquare$$

Bemerkungen. **(a)** In Kapitel 3, S.144 haben wir gezeigt, daß es für einen endlich dimensionalen Vektorraum V einen natürlichen basisunabhängigen Isomorphismus $\Psi : V \longrightarrow V^{**}$ gibt, während wir für V und V^* keinen basisunabhängigen Isomorphismus angeben konnten. Für einen euklidischen Vektorraum V mit $\dim V < \infty$ ist dies anders, denn hier ist die Abbildung

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

ein solcher Isomorphismus. Wir können deshalb in diesem Fall V mit V^* identifizieren und werden dies ab jetzt auch tun.

(b) Wird V mit V^* identifiziert, so fällt eine Basis B von V genau dann mit der dualen Basis B^* in V^* zusammen, wenn B Orthonormalbasis ist.

(c) Setzt man $V = V^*$ und $W = W^*$ für $\dim V < \infty$ und $\dim W < \infty$ (wie in (a)), so fällt die adjungierte Abbildung $\Phi^* : W \longrightarrow V$ mit der transponierten Abbildung $\Phi^\mathsf{T} : W^* \longrightarrow V^*$ zusammen.

(d) Als Folgerung von (b),(c) ergibt sich nun, daß für endlich dimensionale euklidische Vektorräume V, W, für $\Phi \in \mathrm{Hom}\,(V, W)$ und für Orthonormalbasen B von V und C von W die Abbildungsmatrix von Φ^* gerade die Transponierte der Abbildungsmatrix von Φ ist (bzgl. B, C) :

$$A_{\Phi^*} = A_\Phi^{\mathsf{T}}.$$

Dies läßt sich auch direkt beweisen: Seien $B = (v_1,\dots,v_n)$, $C = (w_1,\dots,w_m)$ Orthonormalbasen von V, W, $A_\Phi = (\!(a_{ij})\!)$, $A_{\Phi^*} = (\!(b_{ij})\!)$. Dann gilt für $j = 1,\dots,n$ und $i = 1,\dots,m$

$$\Phi(v_j) = \sum_{i=1}^{m} a_{ij}\, w_i = \sum_{i=1}^{m} <\Phi(v_j), w_i>\, w_i \ ,$$

$$\Phi^*(w_i) = \sum_{j=1}^{n} b_{ji}\, v_j = \sum_{j=1}^{n} <\Phi^*(w_i), v_j>\, v_j \ .$$

Also gilt

$$a_{ij} = <\Phi(v_j), w_i> = <v_j, \Phi^*(w_i)> = b_{ji}\,,\ \text{d.h.}\ A_{\Phi^*} = A_\Phi^{\mathsf{T}}. \qquad \blacksquare$$

Für Endomorphismen Φ eines endlich dimensionalen euklidischen Vektorraumes V folgt daraus $\det \Phi = \det \Phi^*$.

(e) Für Φ^* gelten die folgenden Rechenregeln, die wir schon von der transponierten Abbildung her kennen. Für endlich dimensionale Vektorräume folgen die Aussagen daher direkt aus (c). Im allgemeinen Fall kann man sie auch leicht aus der Definition von Φ^* herleiten (Übungsaufgabe):

(i) $(\Phi^*)^* = \Phi$,

(ii) $(a\,\Phi)^* = a\,\Phi^*\ ,\ a \in \mathbb{R}$,

(iii) $(\Phi_1 + \Phi_2)^* = \Phi_1^* + \Phi_2^*\,$,

(iv) $(\Phi \circ \Psi)^* = \Psi^* \circ \Phi^*$.

Über die adjungierte Abbildung erhalten wir eine wichtige Klasse von Abbildungen.

Definition. Es seien V ein euklidischer Vektorraum und $\Phi \in \operatorname{Hom}(V,V)$. Die Abbildung Φ heißt *selbstadjungiert*, wenn

$$<\Phi(x),y> \;=\; <x,\Phi(y)>$$

für alle $x,y \in V$ gilt. Φ heißt *antiselbstadjungiert*, wenn

$$<\Phi(x),y> \;=\; -<x,\Phi(y)>$$

für alle $x,y \in V$ gilt.

Bemerkungen. **(a)** Φ ist genau dann selbstadjungiert bzw. antiselbstadjungiert, wenn Φ^* existiert und $\Phi = \Phi^*$ bzw. $\Phi = -\Phi^*$ erfüllt.

(b) Ist $\dim V < \infty$ und ist B eine ONB von V, so ist Φ genau dann selbstadjungiert bzw. antiselbstadjungiert, wenn für die Abbildungsmatrix von Φ bezüglich B

$$A_\Phi = A_\Phi{}^{\mathsf T} \text{ bzw. } A_\Phi = -A_\Phi{}^{\mathsf T}$$

gilt. Matrizen A mit $A^{\mathsf T} = -A$ heißen *schiefsymmetrisch*.

Im endlich dimensionalen Fall werden selbstadjungierte Endomorphismen Φ also durch symmetrische Abbildungsmatrizen beschrieben und sind daher nach Satz 3 diagonalisierbar. Der folgende Satz ist eine Verschärfung von Satz 3.

Satz 13. *Es seien V ein endlich dimensionaler euklidischer Vektorraum und Φ ein Endomorphismus von V. Genau dann ist Φ selbstadjungiert, wenn es in V eine Orthonormalbasis aus Eigenvektoren von Φ gibt.*

Beweis. Sei zunächst $B = (x_1,\dots,x_n)$ eine ONB aus Eigenvektoren von Φ und $c_1,\dots c_n$ die zugehörigen Eigenwerte. Dann folgt $<\Phi(x_i),x_j> \;=\; c_i <x_i,x_j> \;=\; c_i\,\delta_{ij} = c_j\,\delta_{ij} = <x_i,\Phi(x_j)>$. Folglich ist $<\Phi(x),y> \;=\; <x,\Phi(y)>$ für alle $x,y \in V$, also $\Phi = \Phi^*$.

Umgekehrt sei $\Phi = \Phi^*$ und B sei eine ONB von V. Dann gilt $A_\Phi = A_\Phi{}^{\mathsf T}$. Nach Satz 3 ist A_Φ diagonalisierbar, und die Eigenräume zu verschiedenen Eigenwerten sind zueinander orthogonal. Wählen wir in jedem Eigenraum eine ONB, so erhalten wir insgesamt eine ONB $(\hat{x}_1,\dots,\hat{x}_n)$ des $\mathbb{R}^n$ aus Eigenvektoren von A_Φ. Die zugehörigen Vektoren $x_1,\dots,x_n$ sind dann Eigenvektoren von Φ und bilden eine ONB in V. ■

Bemerkung und Definition. Die Matrix S, die die symmetrische Matrix A_Φ auf Diagonalgestalt transformiert, kann also so gewählt werden, daß ihre Spalten eine ONB von $\mathbb{R}^n$ bilden. Dies ist gleichwertig mit $S^\top S = E_n$.

Eine Matrix $A \in \mathbb{R}^{n\times n}$, die $A^\top A = E_n$ erfüllt, heißt *orthogonal.* Es gilt:

$$A \text{ orthogonal} \iff A^{-1} = A^\top \iff A\,A^\top = E_n\,.$$

Quadratische (n,n)–Matrizen A und A', für die $A' = S^\top A\,S$ mit einer orthogonalen Matrix S gilt, heißen *orthogonal äquivalent.*

Damit lautet die Aussage von Satz 13 für Matrizen:

Genau dann ist die Matrix $A \in \mathbb{R}^{n\times n}$ symmetrisch, wenn sie zu einer Diagonalmatrix orthogonal äquivalent ist.

Bemerkung. Für schiefsymmetrische Matrizen kann man ebenfalls eine Normalform angeben (siehe S.274).

Beispiele. **(a)** Zu der symmetrischen Matrix

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

wird eine orthogonale Matrix S gesucht, so daß $S^\top A\,S$ eine Diagonalmatrix ist.

Aus dem charakteristischen Polynom $p = -\,(2 + X)^2\,X$ von A ergeben sich die Eigenwerte $c_1 = 0$ und $c_2 = -2$. Für die zugehörigen Eigenräume E_{c_1} und E_{c_2} erhalten wir:

$$A\,v = o \iff v = \begin{bmatrix} a \\ 0 \\ a \end{bmatrix},\ a \in \mathbb{R}\ ;\quad (A + 2\,E_3)\,v = o \iff v = \begin{bmatrix} a \\ b \\ -a \end{bmatrix},\ a,b \in \mathbb{R}\,,$$

also

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

Nun sind

$$\left\{\begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}\right\} \quad \text{bzw.} \quad \left\{\begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right\}$$

Orthonormalbasen von E_{c_1} bzw. von E_{c_2}. Also ist

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

eine orthogonale Matrix mit

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

(b) Orthogonalprojektionen π_U sind selbstadjungiert: Ist V endlich dimensional, so kann man eine ONB von U wählen und diese zu einer ONB von V ergänzen. Die Abbildungsmatrix von π_U bezüglich dieser Basis hat dann die Gestalt

$$A_{\pi_U} = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & O & \\ & & 1 & & & \\ & & & 0 & & \\ & O & & & \ddots & \\ & & & & & 0 \end{bmatrix}$$

Für unendlich dimensionale Vektorräume überlassen wir den Beweis als Übungsaufgabe.

In unendlich dimensionalen euklidischen Vektorräumen haben selbstadjungierte Abbildungen nicht die gleichen guten Eigenschaften, sie müssen noch nicht einmal Eigenwerte besitzen.

Beispiel. In dem euklidischen Vektorraum $V = C([0,1])$ mit dem Skalarprodukt

$$<f,g> = \int_0^1 f(t)\, g(t)\, \mathrm{d}t$$

betrachten wir die lineare Abbildung

$$\Phi : V \longrightarrow V, f \longmapsto \Phi(f) \text{ mit } \Phi(f)(t) = t \cdot f(t), t \in [0,1].$$

Φ ist selbstadjungiert:

$$<\Phi(f),g> = \int_0^1 t\, f(t)\, g(t)\, \mathrm{d}t = <f,\Phi(g)>.$$

Ist $c \in \mathbb{R}$ Eigenwert von Φ, so existiert ein Eigenvektor f, d.h. eine von der Nullfunktion o verschiedene Funktion $f \in C([0,1])$ mit $\Phi(f) = c\, f$. Daraus folgt $t\, f(t) = c\, f(t)$ für alle $t \in [0,1]$, also $f(t) = 0$ für alle $t \neq c$. Für $c \notin [0,1]$ folgt daher $f = o$, für $c \in [0,1]$ gilt wegen der Stetigkeit von f ebenfalls $f = o$. Dies ist ein Widerspruch, daher besitzt Φ keine Eigenwerte.

Weitere Struktureigenschaften selbstadjungierter Endomorphismen werden bei der Untersuchung einer allgemeineren Klasse linearer Abbildungen in § 5.5 hergeleitet.

§ 4 Isometrien

Es seien V und W euklidische Vektorräume. Die naheliegendste und interessanteste Klasse von linearen Abbildungen $\Phi : V \longrightarrow W$ wird von denjenigen Abbildungen gebildet, die außer mit der Vektorraumstruktur auch mit dem Skalarprodukt in V bzw. W verträglich sind, also

$$<\Phi(x),\Phi(y)> \;=\; <x,y>$$

für alle $x,y \in V$ erfüllen.

Definition. Es seien V, W euklidische Vektorräume. $\Phi \in \operatorname{Hom}(V,W)$ heißt *Isometrie* oder *orthogonale Abbildung*, wenn

$$<\Phi(x),\Phi(y)> \;=\; <x,y>$$

für alle $x,y \in V$ gilt.

Bemerkung. Jede Isometrie Φ ist injektiv. Ist Φ außerdem surjektiv, so heißen die zwei euklidischen Vektorräume V und W *isometrisch isomorph*.

Der Name Isometrie wird durch folgenden Satz gerechtfertigt.

Satz 14. *Es seien V, W euklidische Vektorräume und $\Phi \in \operatorname{Hom}(V,W)$. Dann sind folgende Aussagen äquivalent:*

(a) *Φ ist Isometrie.*

(b) $\| \Phi(x) \| = \| x \|$ *für alle* $x \in V$.

(c) $d(\Phi(x),\Phi(y)) = d(x,y)$ *für alle* $x,y \in V$.

Beweis. (a) $\Longrightarrow$ (b): $\| \Phi(x) \| \;=\; \sqrt{<\Phi(x),\Phi(x)>} \;=\; \sqrt{<x,x>} \;=\; \| x \|$.

(b) $\Longrightarrow$ (c): $\| \Phi(x) - \Phi(y) \| \;=\; \| \Phi(x-y) \| \;=\; \| x - y \|$.

(c) $\Longrightarrow$ (a): $\quad <\Phi(x),\Phi(y)> \;=\; \frac{1}{4}(\| \Phi(x) + \Phi(y) \|^2 - \| \Phi(x) - \Phi(y) \|^2)$

$$= \frac{1}{4}(\| \Phi(x) - \Phi(-y) \|^2 - \| \Phi(x) - \Phi(y) \|^2) \;=\; \frac{1}{4}(\| x - (-y) \|^2 - \| x - y \|^2)$$

$$= \frac{1}{4}(\| x+y \|^2 - \| x-y \|^2) = <x,y> . \qquad \blacksquare$$

Im Fall $\dim V = \dim W = n$ lassen sich für Isometrien weitere Kennzeichnungen angeben.

Satz 15. *Es seien V, W euklidische Vektorräume mit* $\dim V = \dim W = n$ *sowie* $\Phi \in \mathrm{Hom}\,(V, W)$. *Dann sind äquivalent:*

(a) *Φ ist Isometrie.*

(b) *Für jede ONB $(x_1,\ldots,x_n)$ von V ist $(\Phi(x_1),\ldots,\Phi(x_n))$ ONB von W.*

(c) *Es gibt eine ONB $(x_1,\ldots,x_n)$ von V, für die $(\Phi(x_1),\ldots,\Phi(x_n))$ ONB von W ist.*

(d) $\Phi^* \circ \Phi = \mathrm{id}_V$, $\Phi \circ \Phi^* = \mathrm{id}_W$.

(e) *Bezüglich jeder ONB B von V und jeder ONB C von W ist A_Φ orthogonal.*

(f) *Es gibt eine ONB B von V und eine ONB C von W, für die A_Φ orthogonal ist.*

Beweis. (a) $\Longrightarrow$ (b): Φ ist injektiv, also wegen $\dim V = \dim W = n$ ein Isomorphismus. Daher ist $(\Phi(x_1),\ldots,\Phi(x_n))$ eine Basis von W. Wegen $<\Phi(x_i),\Phi(x_j)> = <x_i,x_j> = \delta_{ij}$ für $i,j = 1,\ldots,n$ ist $(\Phi(x_1),\ldots,\Phi(x_n))$ ONB von W.

(b) $\Longrightarrow$ (c): Nach Satz 7 existiert in V eine ONB $(x_1,\ldots,x_n)$. Dann ist $(\Phi(x_1),\ldots,\Phi(x_n))$ ONB von W.

(c) $\Longrightarrow$ (f): Bezüglich der Orthonormalbasen $(x_1,\ldots,x_n)$ von V und $(\Phi(x_1),\ldots,\Phi(x_n))$ von W ist $A_\Phi = E_n$, also ist A_Φ orthogonal.

(f) $\Longrightarrow$ (d): Bezüglich der Orthonormalbasen B und C gilt $A_{\Phi^* \circ \Phi} = A_{\Phi^*} A_\Phi = A_\Phi{}^\top A_\Phi = E_n$, also $\Phi^* \circ \Phi = \mathrm{id}_V$. Analog folgt $\Phi \circ \Phi^* = \mathrm{id}_W$.

(d) $\Longrightarrow$ (a): Es gilt $<\Phi(x),\Phi(y)> = <x,\Phi^* \circ \Phi(y)> = <x,y>$ für alle $x,y \in V$.

(d) $\Longrightarrow$ (e): Bezüglich jeder ONB B von V und jeder ONB C von W gilt $A_{\Phi^*} = A_\Phi^\top$, also $A_\Phi^\top A_\Phi = A_{\Phi^*} A_\Phi = A_{\Phi^* \circ \Phi} = E_n$. Somit ist A_Φ orthogonal.

(e) $\Longrightarrow$ (f): Nach Satz 7 gibt es in V eine ONB B und in W eine ONB C. $\blacksquare$

Wir betrachten nun speziell eine Isometrie $\Phi : V \longrightarrow V$, $\dim V = n$. Wegen $\Phi^* \circ \Phi = \mathrm{id}_V$ und $\det \Phi = \det \Phi^*$ folgt $\det \Phi = \pm 1$.

Ist c Eigenwert von Φ, so erhalten wir aus $\Phi(x) = c\,x$, $x \neq o$, unmittelbar $\| \Phi(x)\| = |c|\, \| x \|$, also $|c| = 1$.

Isometrien $\Phi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ heißen auch *Drehungen* des $\mathbb{R}^n$; ist $\det \Phi = 1$, so spricht man von *eigentlichen Drehungen*, bei $\det \Phi = -1$ von *uneigentlichen Drehungen* oder *Drehspiegelungen*.

Für $\dim V < \infty$ bilden die Isometrien $\Phi : V \longrightarrow V$ eine Untergruppe $O(V)$ von $\mathrm{Aut}(V)$. Die zugehörige Matrizengruppe heißt *orthogonale Gruppe* und wird mit $O(n)$ bezeichnet. Die Matrizen A aus dieser Gruppe mit $\det A = 1$ bilden wieder eine Untergruppe, die *spezielle orthogonale Gruppe* $SO(n)$.

Diese Bezeichnungen beinhalten schon eine geometrische Interpretation von Isometrien, die wir nun nachvollziehen wollen. Dazu leiten wir für Isometrien eine Normalform her.

Satz 16. *Es seien V ein n-dimensionaler euklidischer Vektorraum und $\Phi \in \mathrm{Hom}(V,V)$. Dann ist Φ genau dann Isometrie, wenn eine geordnete Orthonormalbasis von V existiert, bezüglich der A_Φ die Form*

$$A_\Phi = \begin{bmatrix} 1 & & & & & & & & & & \\ & \ddots & & & & & & & & O & \\ & & 1 & & & & & & & & \\ & & & -1 & & & & & & & \\ & & & & \ddots & & & & & & \\ & & & & & -1 & & & & & \\ & & & & & & \cos\omega_1 & -\sin\omega_1 & & & \\ & & & & & & \sin\omega_1 & \cos\omega_1 & & & \\ & & O & & & & & & \ddots & & \\ & & & & & & & & & \cos\omega_k & -\sin\omega_k \\ & & & & & & & & & \sin\omega_k & \cos\omega_k \end{bmatrix} \qquad (*)$$

besitzt mit $\omega_i \in (0,\pi)$, $i = 1,\dots,k$.

Beweis. Jede Matrix A_Φ dieser Form erfüllt $A_\Phi^{\mathsf{T}} A_\Phi = E_n$, also ist A_Φ orthogonal und Φ somit eine Isometrie.

Die Umkehrung beweisen wir so, daß damit zugleich ein praktisches Verfahren zur Aufstellung der Normalform gewonnen wird.

Sei Φ eine Isometrie. Wir betrachten die Hilfsabbildung

$$\Psi := \Phi + \Phi^* = \Phi + \Phi^{-1}.$$

Ψ ist selbstadjungiert, also existiert nach Satz 13 in V eine ONB aus Eigenvektoren von Ψ. Somit gilt

$$V = E_2 \oplus E_{-2} \oplus E_{c_1} \oplus \cdots \oplus E_{c_k}$$

mit $c_i \neq \pm 2$, und alle Eigenräume von Ψ sind paarweise orthogonal. Dabei tritt der Eigenraum E_2 von Ψ zum Eigenwert 2 genau dann auf, wenn Φ den Eigenwert 1 besitzt, und E_{-2} tritt genau dann auf, wenn Φ den Eigenwert -1 hat. Wir zeigen dies für den ersten Fall, der andere folgt dann analog:

Aus $\Phi(x) = x$ folgt $\Phi^{-1}(x) = x$ und somit $\Psi(x) = 2\,x$. Sei umgekehrt $\Psi(x) = 2\,x$, also $(\Phi + \Phi^*)(x) = 2\,x$. Dann ist $2 <x,x> = <\Phi(x),x> + <\Phi^*(x),x> = 2 <\Phi(x),x>$ und somit $<\Phi(x) - x, \Phi(x) - x> = 0$. Also gilt $\Phi(x) = x$.

Alle Eigenräume E_c von Ψ sind Φ-invariant, denn für $x \in E_c$ folgt

$$\Psi(\Phi(x)) = (\Phi + \Phi^{-1})(\Phi(x)) = \Phi \circ (\Phi + \Phi^{-1})(x) = \Phi(\Psi(x)) = \Phi(c\,x) = c\,\Phi(x),$$

also $\Phi(x) \in E_c$. Somit folgt $\Phi(E_c) \subset E_c$. Da Φ Isometrie ist, gilt sogar $\Phi(E_c) = E_c$.

Wir können damit alle Eigenräume von Ψ getrennt behandeln und dort geeignete Orthonormalbasen suchen.

1. Fall: $c = \pm 2$. Wie wir oben schon gesehen haben, ist $E_2(\Psi)$ gerade der Eigenraum von Φ zum Eigenwert 1 und $E_{-2}(\Psi)$ ist der Eigenraum von Φ zum Eigenwert -1. In jedem dieser Eigenräume können wir beliebige Orthonormalbasen wählen.

2. Fall: $c \neq \pm 2$. Dann ist $\dim E_c$ gerade und es gibt $l = \frac{1}{2} \dim E_c$ paarweise orthogonale, zweidimensionale und Φ-invariante Untervektorräume $U_1, \ldots, U_l$ mit $E_c = U_1 \oplus \cdots \oplus U_l$.

Beweis. Wäre dim E_c ungerade, so hätte das charakteristische Polynom von $\Phi|_{E_c}$ eine Nullstelle. Diese wäre Eigenwert von Φ, also ± 1, und somit gäbe es in E_c einen Eigenvektor von Ψ zum Eigenwert ± 2. Dies ist ein Widerspruch.

Sei nun $x_1 \in E_c$, $\| x_1 \| = 1$, beliebig gewählt. Dann sind $x_1, \Phi(x_1)$ linear unabhängig, da sonst $\Phi(x_1) = \pm x_1$ und somit $\Psi(x_1) = \pm 2 x_1$ folgen würde. Der Untervektorraum $U_1 = [x_1, \Phi(x_1)] \subset E_c$ ist also zweidimensional.

U_1 ist Φ-invariant: Aus $\Phi \circ (\Phi + \Phi^{-1})(x_1) = \Phi(\Psi(x_1)) = \Phi(c\, x_1) = c\, \Phi(x_1)$ folgt nämlich $\Phi^2(x_1) + x_1 = c\, \Phi(x_1)$, also $\Phi^2(x_1) \in U_1$. Somit gilt $\Phi(U_1) \subset U_1$ und, weil Φ Isometrie ist, sogar $\Phi(U_1) = U_1$.

$U_1^{\perp}$ ist ebenfalls Φ-invariant: Seien $x \in U_1^{\perp}$ und $y \in U_1$ beliebig gewählt. Dann gilt $\Phi^{-1}(y) \in U_1$ und somit $<\Phi(x), y> = <x, \Phi^*(y)> = <x, \Phi^{-1}(y)> = 0$, also $\Phi(x) \in U_1^{\perp}$. Wegen $\Phi(E_c) = E_c$ und $\Phi(U_1^{\perp}) = U_1^{\perp}$ gilt daher auch $\Phi(U_1^{\perp} \cap E_c) = U_1^{\perp} \cap E_c$.

Ist $U_1^{\perp} \cap E_c \neq \{o\}$, so wählen wir einen beliebigen normierten Vektor x_2 aus dieser Menge und bilden den Untervektorraum $U_2 = [x_2, \Phi(x_2)]$. Dieser ist wieder zweidimensional, Φ-invariant und wegen $U_2 \subset U_1^{\perp} \cap E_c$ orthogonal zu U_1. Danach wählen wir x_3 aus $(U_1 \oplus U_2)^{\perp} \cap E_c$, $\| x_3 \| = 1$, usw. Das Verfahren bricht nach $l = \frac{1}{2} \dim E_c$ Schritten ab, und wir erhalten die Darstellung $E_c = U_1 \oplus \cdots \oplus U_l$.

In jedem dieser l Unterräume $U = [x, \Phi(x)] \subset E_c$ konstruieren wir von der Basis $(x, \Phi(x))$ ausgehend eine Orthonormalbasis (x, y) mit

$$y := \frac{\Phi(x) - <\Phi(x), x>\, x}{\|\Phi(x) - <\Phi(x), x>\, x\|}.$$

Wir stellen den Vektor y in anderer Form dar: Zunächst folgt aus $\Psi(x) = c\, x$ die Gleichung $<\Phi(x), x> = \frac{c}{2} <x, x> = \frac{c}{2}$. Ist $\omega \in (0, \pi)$ der Winkel (vgl. S. 226) zwischen x und $\Phi(x)$, so gilt

$$\cos \omega = \frac{c}{2}, \ \sin \omega = \sqrt{1 - c^2/4}.$$

Daraus folgt

$$y = (1/\sin \omega)\, (\Phi(x) - \cos \omega\, x)$$

und somit

$$\Phi(x) = \cos\omega\, x + \sin\omega\, y$$

und wegen $\Phi^2(x) = c\,\Phi(x) - x = 2\cos\omega\,\Phi(x) - x$

$$\begin{aligned}\Phi(y) &= (1/\sin\omega)\,(\Phi^2(x) - \cos\omega\,\Phi(x)) \\ &= (1/\sin\omega)\,(2\cos\omega\,\Phi(x) - x - \cos\omega\,\Phi(x)) \\ &= (1/\sin\omega)\,(\cos^2\omega\, x + \cos\omega\sin\omega\, y - x) = -\sin\omega\, x + \cos\omega\, y\,.\end{aligned}$$

Bezüglich dieser Basis gilt also

$$A_{\Phi|_U} = \begin{bmatrix} \cos\omega & -\sin\omega \\ \sin\omega & \cos\omega \end{bmatrix}$$

und damit

$$A_{\Phi|_{E_c}} = \begin{bmatrix} \cos\omega & -\sin\omega & & & \\ \sin\omega & \cos\omega & & O & \\ & & \ddots & & \\ & O & & \cos\omega & -\sin\omega \\ & & & \sin\omega & \cos\omega \end{bmatrix}$$

mit $l(c) = \frac{1}{2}\dim E_c$ Kästchen.

Insgesamt erhalten wir somit die gesuchte Normalform A_Φ. ■

Bemerkungen. **(a)** In der Literatur wird manchmal auch eine andere Matrix (**) als Normalform bezeichnet. Sie unterscheidet sich von der obigen Form (*) dadurch, daß die Zweierkästchen in (**) die Gestalt

$$\boxed{\begin{matrix} \cos\omega & \sin\omega \\ -\sin\omega & \cos\omega \end{matrix}}$$

besitzen. Jede der Normalformen geht aus der anderen dadurch hervor, daß bei jedem Zweierkästchen jeweils die Reihenfolge der zugehörigen orthonormalen Basisvektoren x,y vertauscht wird oder daß y durch $-y$ ersetzt wird.

(b) Ist $A \in \mathbb{R}^{n\times n}$ orthogonal, so liefert das obige Beweisverfahren, angewendet auf die symmetrische Hilfsmatrix $B = A + A^{\mathsf{T}}$, eine orthogonale Matrix $S \in \mathbb{R}^{n\times n}$, so daß $S^{\mathsf{T}} A\, S$ die Normalform (*) besitzt.

(c) Im nächsten Paragraphen werden wir noch einen anderen Beweis kennenlernen, bei dem A als komplexe Matrix angesehen wird und der die reelle Normalform aus Satz 4.22 benutzt.

Geometrische Interpretation im $\mathbb{R}^n$

Die Abbildungsmatrix

$$A_1 = \begin{bmatrix} \cos\omega & -\sin\omega \\ \sin\omega & \cos\omega \end{bmatrix}$$

beschreibt eine Abbildung $\Phi_1 : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, die folgende Eigenschaften hat. Schreiben wir $x \in \mathbb{R}^2$ in Polarkoordinaten

$$x = \begin{bmatrix} r\cos\alpha \\ r\sin\alpha \end{bmatrix},$$

so ist

$$\Phi_1(x) = \begin{bmatrix} \cos\omega & -\sin\omega \\ \sin\omega & \cos\omega \end{bmatrix} \begin{bmatrix} r\cos\alpha \\ r\sin\alpha \end{bmatrix} = \begin{bmatrix} r\cos(\alpha+\omega) \\ r\sin(\alpha+\omega) \end{bmatrix}.$$

Also ist Φ_1 eine Drehung im $\mathbb{R}^2$ um den Ursprung mit dem Drehwinkel ω. Damit ist auch klar, daß Φ_1 bezüglich jeder ONB in $\mathbb{R}^2$ eine der Matrizen A_1 oder A_1^{T} als Abbildungsmatrix besitzt.

Setzen wir nun diese Information über die einzelnen Kästchen von A_Φ zusammen, so erhalten wir die folgende geometrische Beschreibung einer Isometrie $\Phi : V \longrightarrow V$. Es gibt eine Orthogonalzerlegung

$$V = U_1 \oplus U_2 \oplus W_1 \oplus \cdots \oplus W_k,$$

so daß $\Phi|_{U_1} = \mathrm{id}_{U_1}$ ist, also U_1 Fixunterraum ist, so daß $\Phi|_{U_2} : x \longmapsto -x$ gilt, also Φ

die Vektoren aus U_2 an $U_2^\perp = U_1 \oplus W_1 \oplus \cdots \oplus W_k$ spiegelt, und so daß $\Phi|_{W_i}$, $i = 1,\ldots k$, Drehungen in den zweidimensionalen Untervektorräumen W_i darstellen.

Nun kann man die speziellen Matrizen

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ und } \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

ebenfalls als Drehmatrizen mit Winkel 0 bzw. π ansehen. Folglich ergeben sich genaue Beschreibungen von Φ in den Dimensionen 2 und 3:

n = 2 : Für $\det \Phi = 1$ ist Φ eine Drehung um einen Winkel $\omega \in [0,\pi]$, für $\det \Phi = -1$ ist Φ Spiegelung an einer Geraden durch den Ursprung.

n = 3 : $\det \Phi = 1$: Es gibt einen Vektor $v \neq o$, der Fixvektor ist, im Orthogonalraum $v^\perp$ ist Φ eine Drehung. Φ heißt dann *Drehung* des Vektorraums V mit der *Drehachse* $[v]$ und der *Drehebene* $v^\perp$.

$\det \Phi = -1$: Es gibt ein $v \neq o$ mit $\Phi(v) = -v$, im Orthogonalraum $v^\perp$ ist Φ eine Drehung. Φ ist dann Produkt einer Drehung um die Achse $[v]$ mit einer Spiegelung an der Drehebene $v^\perp$ und heißt deswegen eine *Drehspiegelung* von V .

Beispiel. Die Matrix

$$A = \frac{1}{9\sqrt{2}} \begin{bmatrix} 9 & 4 & 8 & 1 \\ -4 & 9 & -1 & 8 \\ -8 & 1 & 9 & -4 \\ -1 & -8 & 4 & 9 \end{bmatrix} \in \mathbb{R}^{4\times 4}$$

ist orthogonal, denn es gilt $A^\top A = E_4$. Wir wollen ihre Normalform $\tilde{A}$ bestimmen und betrachten hierzu zunächst die symmmetrische Matrix

$$A + A^\top = \frac{1}{9\sqrt{2}} \begin{bmatrix} 18 & 0 & 0 & 0 \\ 0 & 18 & 0 & 0 \\ 0 & 0 & 18 & 0 \\ 0 & 0 & 0 & 18 \end{bmatrix} = \sqrt{2}\, E_4 \,.$$

$A + A^{\mathsf{T}}$ hat den 4-fachen Eigenwert $c = \sqrt{2}$. Die Normalform $\tilde{A}$ von A besteht daher aus 2 Kästchen der Form

$$\boxed{\begin{matrix} \cos\,\omega & -\sin\,\omega \\ \sin\,\omega & \cos\,\omega \end{matrix}}$$

mit $\cos\,\omega = \frac{1}{2}\sqrt{2}$. Daraus folgt $\omega = \frac{\pi}{4}$ und $\sin\,\omega = \frac{1}{2}\sqrt{2}$ sowie

$$\tilde{A} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 & 0 \\ 0 & 0 & 1/\sqrt{2} & -1/\sqrt{2} \\ 0 & 0 & 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

Berechnung der orthogonalen Transformationsmatrix S:

Sei x Eigenvektor von $A + A^{\mathsf{T}}$ zum Eigenwert c. Wegen $E_c = \mathbb{R}^4$ können wir x beliebig wählen, etwa $x = e_1$. Dann ist

$$A\,x = \frac{1}{9\sqrt{2}} \begin{bmatrix} 9 \\ -4 \\ -8 \\ -1 \end{bmatrix}.$$

Orthogonalisieren ergibt

$$x_1 = e_1\ ,\ \tilde{x}_2 = A\,x_1 - \langle Ax_1, x_1 \rangle\, x_1 = \frac{1}{9\sqrt{2}} \begin{bmatrix} 0 \\ -4 \\ -8 \\ -1 \end{bmatrix},$$

also

$$x_2 = \frac{\tilde{x}_2}{\|\tilde{x}_2\|} = \frac{1}{9} \begin{bmatrix} 0 \\ -4 \\ -8 \\ -1 \end{bmatrix}.$$

Nun muß $[x_1,x_2]^\perp$ bestimmt werden:

$$[x_1,x_2]^\perp = \left\{ \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} \middle| \ a_1 = 0,\ 4\ a_2 + 8\ a_3 + a_4 = 0 \right\} = \left[\begin{bmatrix} 0 \\ 1 \\ -1 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} \right].$$

In $[x_1,x_2]^\perp$ bestimmen wir ebenfalls eine ONB : Wir wählen $x_3 \in [x_1,x_2]^\perp$, $\| x_3 \| = 1$, beliebig, bilden $A\ x_3$ und orthogonalisieren:

$$x_3 = \frac{1}{3\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ -1 \\ 4 \end{bmatrix}, \quad \tilde{x}_4 = A\ x_3 - \langle Ax_3,x_3\rangle\ x_3 = \frac{1}{9} \begin{bmatrix} 0 \\ 7 \\ -4 \\ 4 \end{bmatrix} - \frac{1}{6} \begin{bmatrix} 0 \\ 1 \\ -1 \\ 4 \end{bmatrix} = \frac{1}{18} \begin{bmatrix} 0 \\ 11 \\ -5 \\ -4 \end{bmatrix},$$

also

$$x_4 = \frac{\tilde{x}_4}{\|\tilde{x}_4\|} = \frac{1}{9\sqrt{2}} \begin{bmatrix} 0 \\ 11 \\ -5 \\ -4 \end{bmatrix}.$$

Damit erhalten wir

$$S = (x_1 \mid x_2 \mid x_3 \mid x_4) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -4/9 & 1/3\sqrt{2} & 11/9\sqrt{2} \\ 0 & -8/9 & -1/3\sqrt{2} & -5/9\sqrt{2} \\ 0 & -1/9 & 4/3\sqrt{2} & -4/9\sqrt{2} \end{bmatrix}$$

§ 5 Unitäre Vektorräume

Wir wollen in diesem Paragraphen zunächst die Ergebnisse aus den vorigen Abschnitten 5.1 bis 5.4 soweit wie möglich auf komplexe Vektorräume übertragen. Ein Skalarprodukt β auf V muß, wenn $\sqrt{\beta(x,x)}$ wieder als Länge von x interpretiert werden soll, die gleichen Definitheitseigenschaften haben wie im reellen Fall. Andererseits bewirken Linearitätsforderungen, daß β eine Abbildung nach $\mathbb{C}$ sein muß. Damit kann β aber, im Gegensatz zum reellen Fall, keine symmetrische Bilinearform sein.

Definition. $\beta : V \times V \longrightarrow \mathbb{C}$ heißt *hermitesche Form*, wenn gilt:

(a) $\beta(a\,x + b\,x', y) = a\,\beta(x,y) + b\,\beta(x',y)$ für alle $a,b \in \mathbb{C}$, $x, x', y \in V$,

(b) $\beta(x,y) = \overline{\beta(y,x)}$ für alle $x,y \in V$.

Bemerkungen. **(a)** β ist im ersten Argument linear, im zweiten gilt für alle $a,b \in \mathbb{C}$ und alle $x, y, y' \in V$

$$\beta(x, a\,y + b\,y') = \overline{a}\,\beta(x,y) + \overline{b}\,\beta(x,y').$$

(b) Für alle $x \in V$ ist $\beta(x,x) \in \mathbb{R}$. Damit kann die positive Definitheit wie in 5.1 erklärt werden: β heißt *positiv definit*, wenn $\beta(x,x) > 0$ für alle $x \neq o$ gilt.

Definition. Eine positiv definite hermitesche Form β auf V heißt *Skalarprodukt* ; Schreibweise : $<\cdot,\cdot>$. Das Paar $(V, <\cdot,\cdot>)$ oder kurz V heißt *unitärer Vektorraum* oder auch (*komplexer*) *Vektorraum mit Skalarprodukt.*

Beispiel. $V = \mathbb{C}^n$, $<x,y> := x_1\,\overline{y}_1 + \cdots + x_n\,\overline{y}_n = x^T\,\overline{y}$ ist Skalarprodukt, das *Standardskalarprodukt* auf $\mathbb{C}^n$. Auf der Teilmenge $\mathbb{R}^n$ ($\mathbb{R}^n$ ist kein Untervektorraum von $\mathbb{C}^n$!) geht $<\cdot,\cdot>$ in das reelle Standardskalarprodukt auf $\mathbb{R}^n$ über.

Analog wie im Reellen lassen sich in unitären Vektorräumen nun die *Länge*, der *Abstand* und alle mit der *Orthogonalität* zusammenhängenden Begriffe einführen.

Beispiel. In $\mathbb{C}^n$ ist $\| x \|^2 = x_1\,\overline{x}_1 + \cdots + x_n\overline{x}_n = |x_1|^2 + \cdots + |x_n|^2 = \| \overline{x} \|^2$.

Der Satz 5.1 überträgt sich dann mit kleinen Modifikationen in der Formulierung und im Beweis. So lautet (e) jetzt

$$4 <x,y> \; = \; \| x+y \|^2 - \| x-y \|^2 + i \, \| x+i\,y \|^2 - i \, \| x - i\,y \|^2 ,$$

denn es gilt

$$\| x+y \|^2 - \| x-y \|^2 \; = \; 2 <x,y> + 2 \, \overline{<x,y>}$$

und

$$\| x+i\,y \|^2 - \| x-i\,y \|^2 \; = \; 2 <x,i\,y> + 2 \, \overline{<x,i\;y>} \; = \; -2\,i <x,y> + 2\,i \, \overline{<x,y>} \; .$$

Im Satz von Pythagoras gilt nur die eine Richtung

$$x \perp y \Longrightarrow \| x+y \|^2 \; = \; \| x \|^2 + \| y \|^2 .$$

Die Umkehrung ist falsch, wie das Beispiel $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $y = \begin{bmatrix} i \\ 0 \end{bmatrix}$ zeigt.

Wie in Satz 5.2 lassen sich in endlich dimensionalen unitären Vektorräumen V Skalarprodukte $\beta : V \times V \longrightarrow \mathbb{C}$ nach Auszeichnung einer Basis $B = (v_1,\ldots,v_n)$ durch Matrizen $A \in \mathbb{C}^{n\times n}$ mit

$$\beta\,(x,y) \; = \; \hat{x}^{\mathsf{T}} A \, \overline{\hat{y}}$$

beschreiben. Dabei sind $\hat{x}$ und $\hat{y}$ wieder die Koordinatendarstellungen von x und y bezüglich B und A ist die Matrix $(\!(\beta\,(v_i,v_j))\!)$ mit

$$A \; = \; \overline{A}^{\,\mathsf{T}} .$$

Solche Matrizen heißen *hermitesch.* Weiterhin ist A *positiv definit,* d.h. es gilt $x^{\mathsf{T}} A \, \overline{x} > 0$ für alle $x \in \mathbb{C}^n \setminus \{o\}$.

Die Sätze 6 und 7 (Gram–Schmidtsches Othogonalisierungsverfahren) aus § 5.2 übertragen sich wörtlich auf den komplexen Fall.

Für die Darstellung eines Vektors x bezüglich einer ONB $(x_1,\ldots,x_n)$ von V gilt wieder

$$x = \sum_{i=1}^{n} <x,x_i> x_i ,$$

für das Quadrat der Norm aber

$$\| x \|^2 = \sum_{i=1}^{n} |<x,x_i>|^2 .$$

Orthogonalprojektionen lassen sich wie im euklidischen Fall definieren und die Sätze 8, 9 und 10 gelten entsprechend für unitäre Vektorräume. Im Beweis von Satz 10 muß dabei $a \in \mathbb{C}$ gewählt und die Rechnung entsprechend modifiziert werden. Insbesondere besitzt also jeder endlich dimensionale unitäre Vektorraum eine Orthonormalbasis und Orthogonalprojektionen auf alle Untervektorräume.

Im Unterschied zum reellen Fall kann man einen endlich dimensionalen unitären Vektorraum V nicht wie in § 5.3 mit seinem Dualraum in natürlicher Weise identifizieren. Zwar gilt Satz 5.11 auch im Komplexen, im Beweis muß man nur

$$v = \sum_{i=1}^{n} \overline{x^*(x_i)}\, x_i$$

setzen, aber die Abbildung $\Psi : V \longrightarrow V^*$, $x \longmapsto <\cdot,x>$, ist jetzt nicht mehr linear.

Sind V und W unitäre Vektorräume, so wird die *adjungierte Abbildung* Φ^* einer Abbildung $\Phi \in \operatorname{Hom}(V,W)$ wieder über die Bedingung

$$<\Phi(v),w> = <v,\Phi^*(w)>$$

für alle $v \in V$, $w \in W$ erklärt. Der Existenzsatz 5.12 überträgt sich. Bei den Rechenregeln (S. 250) muß man (ii) durch

(ii') $$(a\,\Phi)^* = \overline{a}\,\Phi^* \ , \ a \in \mathbb{C} ,$$

ersetzen. Ist $\dim V = \dim W = n$, und ist A_Φ die Abbildungsmatrix von Φ bezüglich zweier Orthonormalbasen in V und W, so gilt

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

Weil V und V^* bzw. W und W^* nicht mehr identifiziert werden können, muß man im komplexen Fall die adjungierte Abbildung Φ^* von der transponierten Abbildung $\Phi^\top$ unterscheiden.

Abbildungen $\Phi \in \mathrm{Hom}\,(V,V)$ mit $\Phi^* = \Phi$ bzw. $\Phi^* = -\Phi$ heißen wieder *selbstadjungiert* bzw. *antiselbstadjungiert*. Im Fall $\dim V = n$ gilt dann für die Abbildungsmatrix einer selbstadjungierten Abbildung bezüglich einer Orthonormalbasis

$$A_\Phi = \overline{A_\Phi}^\top,$$

d.h. sie ist hermitesch und für die Abbildungsmatrix einer antiselbstadjungierten Abbildung bezüglich einer Orthonormalbasis gilt

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

Matrizen mit dieser Eigenschaft heißen *schiefhermitesch*.

Isometrien werden wie im reellen Fall durch

$$<\Phi(x),\Phi(y)> = <x,y>$$

für alle $x,y \in V$ definiert; statt Isometrie ist auch der Name *unitäre Abbildung* üblich. Für Isometrien gilt wieder $\Phi^* = \Phi^{-1}$.

Für $\dim V = n$ erfüllt die Abbildungsmatrix A_Φ einer unitären Abbildung $\Phi : V \longrightarrow V$ bezüglich einer ONB von V

$$A_\Phi \cdot \overline{A_\Phi}^\top = \overline{A_\Phi}^\top \cdot A_\Phi = E_n.$$

Solche Matrizen heißen *unitär*. Die unitären Matrizen bilden eine Untergruppe $U(n)$ der allgemeinen linearen Gruppe $GL(n,\mathbb{C})$ und die unitären Matrizen A mit $\det A = 1$ bilden wiederum eine Untergruppe $SU(n)$ von $U(n)$. Die Gruppen $U(n)$ bzw. $SU(n)$ heißen die *unitäre* bzw. die *spezielle unitäre Gruppe*.

Satz 5.14 überträgt sich wörtlich, beim Beweis von (c) $\Longrightarrow$ (a) muß man nur die entsprechende komplexe Darstellung von $<x,y>$ durch die Norm nehmen. In Satz 5.15 müssen in den Aussagen (e) und (f) die Abbildungsmatrizen A_Φ unitär sein.

In euklidischen und unitären Vektorräumen besitzt jeder Endomorphismus Φ, der selbstadjungiert oder antiselbstadjungiert oder eine surjektive Isometrie ist, eine adjungierte Abbildung Φ^*, die $\Phi \circ \Phi^* = \Phi^* \circ \Phi$ erfüllt. Wir nennen solche Endomorphismen normal.

Definition. Es seien V ein euklidischer oder unitärer Vektorraum und $\Phi \in \operatorname{End}(V)$. Φ heißt *normaler Endomorphismus* oder kurz *normal*, wenn Φ^* existiert und

$$\Phi \circ \Phi^* = \Phi^* \circ \Phi$$

erfüllt. Analog heißt die Matrix $A \in \mathbb{C}^{n \times n}$ *normal*, wenn

$$A \cdot \overline{A}^{\top} = \overline{A}^{\top} \cdot A$$

gilt.

Beispiele. **(a)** Wie schon bemerkt, sind selbstadjungierte und antiselbstadjungierte Endomorphismen sowie Isometrien normal.

(b) Symmetrische, schiefsymmetrische und orthogonale Matrizen $A \in \mathbb{R}^{n \times n}$ sind normal, ebenso hermitesche, schiefhermitesche und unitäre Matrizen $A \in \mathbb{C}^{n \times n}$.

Normale Abbildungen lassen sich auch noch durch andere Eigenschaften kennzeichnen.

Satz 17. *Es seien V ein euklidischer oder unitärer Vektorraum und $\Phi \in \operatorname{Hom}(V,V)$. Dann sind folgende Aussagen äquivalent:*

(a) *Φ ist normal.*

(b) *Φ^* existiert und für alle $x,y \in V$ gilt $<\Phi(x),\Phi(y)> = <\Phi^*(x),\Phi^*(y)>$.*

(c) *Φ^* existiert und für alle $x \in V$ gilt $\| \Phi(x) \| = \| \Phi^*(x) \|$.*

Für dim $V = n$ *sind die obigen Aussagen außerdem äquivalent zu:*

(d) *Für jede ONB von V ist die Abbildungsmatrix A_Φ normal.*

(e) *Es gibt eine ONB von V, so daß die Abbildungsmatrix A_Φ normal ist.*

Beweis. (a) $\Longleftrightarrow$ (b): $\Phi^* \circ \Phi = \Phi \circ \Phi^* \iff <x,\Phi^* \circ \Phi(y)> = <x,\Phi \circ \Phi^*(y)>$ für alle $x,y \in V \iff <\Phi(x),\Phi(y)> = <\Phi^*(x),\Phi^*(y)>$ für alle $x,y \in V$.

(b) $\Longleftrightarrow$ (c): Die eine Richtung ist offensichtlich, die andere folgt aus der Darstellung des Skalarproduktes mit Hilfe der Norm.

(a) $\Longrightarrow$ (d): Dies gilt, weil die Abbildungsmatrix von Φ^* bezüglich einer beliebigen ONB von V $A_{\Phi^*} = \overline{A_\Phi}^{\mathsf{T}}$ erfüllt und somit

$$A_\Phi \cdot \overline{A_\Phi}^{\mathsf{T}} = A_\Phi \cdot A_{\Phi^*} = A_{\Phi \circ \Phi^*} = A_{\Phi^* \circ \Phi} = A_{\Phi^*} \cdot A_\Phi = \overline{A_\Phi}^{\mathsf{T}} \cdot A_\Phi$$

gilt.

(d) $\Longrightarrow$ (e): offensichtlich.

(e) $\Longrightarrow$ (a): Für die Abbildungsmatrix von Φ^* bezüglich der gegebenen ONB von V gilt $A_{\Phi^*} = \overline{A_\Phi}^{\mathsf{T}}$ und aus der Normalität von A_Φ folgt dann analog zu oben $\Phi \circ \Phi^* = \Phi^* \circ \Phi$. ■

Bemerkungen. **(a)** Sei Φ normal. Dann gilt Kern Φ = Kern Φ^*.

Dies folgt unmittelbar aus Satz 17(c).

(b) Sei Φ normal. Dann gilt $\Phi(x) = c\,x$ genau dann, wenn $\Phi^*(x) = \overline{c}\,x$ gilt, d.h. die Eigenräume E_c von Φ und $E_{\overline{c}}$ von Φ^* sind gleich.

Beweis. $\Phi - c\,\mathrm{id}_V$ ist normal, denn es ist

$$(\Phi - c\,\mathrm{id}_V) \circ (\Phi - c\,\mathrm{id}_V)^* = (\Phi - c\,\mathrm{id}_V) \circ (\Phi^* - \overline{c}\,\mathrm{id}_V)$$
$$= \Phi \circ \Phi^* - \overline{c}\,\Phi - c\,\Phi^* + c\,\overline{c}\,\mathrm{id}_V = \Phi^* \circ \Phi - c\,\Phi^* - \overline{c}\,\Phi + \overline{c}\,c\,\mathrm{id}_V$$
$$= (\Phi - c\,\mathrm{id}_V)^* \circ (\Phi - c\,\mathrm{id}_V)\,.$$

Damit folgt mit Bemerkung (a) und den Rechenregeln für adjungierte Abbildungen

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

und daraus die Behauptung. ■

(c) Bei einem normalen Endomorphismus Φ sind die Eigenvektoren zu verschiedenen Eigenwerten stets orthogonal.

Beweis. Seien $\Phi(x) = a\,x$, $\Phi(y) = b\,y$ und $a \neq b$. Dann folgt $a < x,y > \, = \, < \Phi(x),y > \, = \, < x,\Phi^*(y) > \, = \, < x,\overline{b}\,y > \, = b < x,y >$, also $< x,y > \, = 0$. ■

Als Hauptergebnis leiten wir nun auch für normale Abbildungen bzw. Matrizen eine Normalform her. Hierzu müssen wir aber zwischen dem komplexen und dem reellen Fall unterscheiden.

Satz 18. *Es seien V ein n–dimensionaler unitärer Vektorraum und $\Phi \in \mathrm{End}\,(V)$. Genau dann ist Φ normal, wenn es in V eine Orthonormalbasis aus Eigenvektoren von Φ gibt.*

Beweis. Sei $(x_1,...x_n)$ eine ONB aus Eigenvektoren von Φ und $\Phi(x_i) = c_i\, x_i$. Dann gilt $\Phi^* \circ \Phi(x_i) = \Phi^*(c_i\, x_i) = c_i\, \overline{c}_i\, x_i = \Phi(\overline{c}_i\, x_i) = \Phi \circ \Phi^*(x_i)$, $i = 1,...,n$, also $\Phi \circ \Phi^* = \Phi^* \circ \Phi$, d.h. Φ ist normal.

Sei umgekehrt Φ normal. Nach dem Fundamentalsatz der Algebra besitzt Φ wenigstens einen Eigenwert c. Sei x_1 ein zugehöriger Eigenvektor mit $\| x_1 \| = 1$. Wir führen den weiteren Beweis nun mit vollständiger Induktion nach n :

$n = 1$: Hier ist $B = \{x_1\}$ ONB von V.

$n-1 \longrightarrow n$: Sei $U = \{x_1\}^\perp$. Dann ist $\dim U = n-1$, und es gilt $\Phi(x_1) \perp U$. Wegen $\Phi^*(x_1) = \overline{c}\; x_1$ folgt auch $\Phi^*(x_1) \perp U$. Somit gilt für alle $y \in U$, daß $<\Phi(y),x_1> = <y,\Phi^*(x_1)> = 0$ ist, also $\Phi(U) \subset U$. Analog ergibt sich $\Phi^*(U) \subset U$. Damit ist $\Phi|_U$ ein normaler Endomorphismus von U, und nach Induktionsvoraussetzung existiert in U eine ONB $\{x_2,...,x_n\}$ aus Eigenvektoren von $\Phi|_U$, also auch von Φ. Die Basis $\{x_1,...,x_n\}$ ist somit ONB in V. ∎

Im reellen Fall (Satz 13) hatten wir aus der Existenz einer ONB von Eigenvektoren schließen können, daß $\Phi = \Phi^*$ ist. Hier, im komplexen Fall, erhalten wir nur die schwächere Aussage $\Phi \circ \Phi^* = \Phi^* \circ \Phi$. Der Grund ist, daß das analoge Vorgehen zum Beweis von Satz 13 auf die Gleichung $<\Phi(x_i),x_j> = \overline{<x_i,\Phi(x_j)>}$ führt, woraus sich nicht mehr $\Phi = \Phi^*$ folgern läßt.

Bemerkung. Die zu Satz 18 äquivalente Aussage für Matrizen $A \in \mathbb{C}^{n\times n}$ lautet:

Genau dann ist $A \in \mathbb{C}^{n\times n}$ normal, wenn es eine unitäre Matrix S gibt, so daß $\overline{S}^{\mathsf{T}} A\, S$ Diagonalgestalt hat.

Damit ergeben sich speziell die Normalformen hermitescher bzw. schiefhermitescher bzw. unitärer Matrizen. Alle haben Diagonalgestalt, wobei im ersten Fall alle Eigenwerte reell sind, im zweiten Fall haben alle Eigenwerte den Realteil 0 und im dritten Fall haben sie alle Betrag 1 (Beweis als Übungsaufgabe).

Außerdem sei bemerkt, daß sich nun auch die Sätze 3 bis 5 auf hermitesche Matrizen übertragen lassen. Insbesondere gibt es auch für positiv definite hermitesche Matrizen A eine Cholesky-Zerlegung in der Form $A = \overline{B}^{\mathsf{T}} B$ mit einer regulären oberen Dreiecksmatrix $B \in \mathbb{C}^{n\times n}$ (Beweis als Übungsaufgabe).

In euklidischen Vektorräumen ist die Situation anders. Hier gilt:

Satz 19. *Es seien V ein n-dimensionaler euklidischer Vektorraum und $\Phi \in$ Hom (V, V). Genau dann ist Φ normal, wenn es in V eine Orthonormalbasis gibt bezüglich der die Abbildungsmatrix von Φ die Form*

$$(*) \qquad \tilde{A} = \begin{bmatrix} c_1 & & & & & \\ & \ddots & & & O & \\ & & c_k & & & \\ & & & \boxed{\begin{matrix} a_1 & b_1 \\ -b_1 & a_1 \end{matrix}} & & \\ & & & & \ddots & \\ & O & & & & \boxed{\begin{matrix} a_m & b_m \\ -b_m & a_m \end{matrix}} \end{bmatrix}$$

mit c_i, a_j, $b_j \in \mathbb{R}$ hat. Dabei sind $c_1,\dots,c_k$ die reellen Nullstellen und $a_j + i\,b_j$, $a_j - i\,b_j$, $b_j > 0$, $j = 1,\dots,m$, die restlichen komplexen Nullstellen des charakteristischen Polynoms von Φ.

Wir beweisen die zu Satz 19 äquivalente Aussage für reelle Matrizen.

Satz 20. *Genau dann ist $A \in \mathbb{R}^{n\times n}$ normal, wenn es eine orthogonale Matrix $S \in \mathbb{R}^{n\times n}$ gibt, so daß $S^{\mathsf{T}} A\, S$ die Form $(*)$ mit c_i, a_j, $b_j \in \mathbb{R}$ hat. Dabei sind $c_1,\dots,c_k$ die reellen Nullstellen und $a_j + i\,b_j$, $a_j - i\,b_j$, $b_j > 0$, $j = 1,\dots,m$, die restlichen komplexen Nullstellen des charakteristischen Polynoms von A.*

Beweis. Gibt es eine orthogonale Matrix S, so daß $\tilde{A} = S^{\top} A\, S$ die Form (*) hat, so folgt $A\, A^{\top} = S\, \tilde{A}\, \tilde{A}^{\top} S^{\top} = S\, \tilde{A}^{\top} \tilde{A}\, S^{\top} = A^{\top} A$.

Ist umgekehrt A normal, so gibt es nach dem letzten Satz eine ONB von $\mathbb{C}^n$ aus komplexen Eigenvektoren von A und $\mathbb{C}^n$ ist direkte Summe paarweise orthogonaler Eigenräume von A.

Für jeden reellen Eigenwert c konstruieren wir in dem zugehörigen (reellen) Eigenraum E_c eine ONB. Da die komplexen Eigenräume von A orthogonal zueinander sind, sind es auch die reellen. Also erhalten wir ein Orthonormalsystem, das gerade die ersten k Spalten von $\tilde{A}$ liefert.

Für jeden komplexen Eigenwert $c = a + i\, b$, $b \neq 0$, ist auch $\overline{c} = a - i\, b$ ein Eigenwert von A und wegen $c \neq \overline{c}$ gilt $E_c \perp E_{\overline{c}}$. Wir können o.B.d.A. $b > 0$ annehmen. In E_c konstruieren wir eine ONB $(z_1,\ldots,z_q)$. Dann ist $(\overline{z}_1,\ldots,\overline{z}_q)$ eine ONB von $E_{\overline{c}}$ und somit $(z_1\,,\, \overline{z}_1,\ldots,z_q\,,\, \overline{z}_q)$ eine ONB von $E_c \oplus E_{\overline{c}}$. Setzen wir für $j = 1,\ldots,q$

$$x_j = \operatorname{Re} z_j = \frac{1}{2}(z_j + \overline{z}_j)\,,$$

$$y_j = \operatorname{Im} z_j = \frac{1}{2i}(z_j - \overline{z}_j)\,,$$

so gilt $[x_j\,,\, y_j] = [z_j\,,\, \overline{z}_j]$, und wir erhalten die Darstellung

$$E_c \oplus E_{\overline{c}} = [x_1,y_1] \oplus \ldots \oplus [x_q,y_q]$$

mit paarweise orthogonalen zweidimensionalen Unterräumen $[x_j\,,\, y_j]$. Ferner gilt

$$A\, x_j = a\, x_j - b\, y_j$$

$$A\, y_j = b\, x_j + a\, y_j$$

für $j = 1,\ldots,q$. Wegen

$$<z_j + \overline{z}_j\,,\, z_j - \overline{z}_j> \;=\; <z_j\,,\, z_j> + <\overline{z}_j\,,\, z_j> - <z_j\,,\, \overline{z}_j> - <\overline{z}_j\,,\, \overline{z}_j> = \|\, z_j \,\|^2 - \|\, \overline{z}_j \,\|^2 = 0$$

und

$$<z_j + \overline{z}_j\,,\, z_j - \overline{z}_j> \;=\; <2\, x_j\,,\, 2\, i\, y_j> \;=\; -4\, i < x_j\,,\, y_j>$$

folgt $<x_j\,,\, y_j> = 0$ für $j = 1,\ldots,q$. Damit ergibt sich

$$0 = \langle z_j , \bar{z}_j \rangle = \langle x_j + i\, y_j , x_j - i\, y_j \rangle = \| x_j \|^2 - \| y_j \|^2 ,$$

und somit

$$\| x_j \| = \| y_j \| .$$

Also erhalten wir für jedes $j = 1,\ldots,q$ mit

$$\frac{x_j}{\| x_j \|} , \frac{y_j}{\| y_j \|}$$

eine ONB zu dem Zweier–Kästchen

$$\boxed{\begin{matrix} a & b \\ -b & a \end{matrix}}$$

Insgesamt erhalten wir so eine reelle ONB von $\mathbb{C}^n$ bezüglich der A die gewünschte Normalform (*) hat. ■

Bemerkungen. **(a)** Die Matrix $\tilde{A}$ ist die reelle Jordansche Normalform (Satz 4.22) der normalen Matrix A.

(b) Aus Satz 20 ergibt sich für symmetrische Matrizen nochmals Satz 13, für schiefsymmetrische Matrizen $A \in \mathbb{R}^{n \times n}$ erhalten wir als Normalform

$$\tilde{A} = \begin{bmatrix} 0 & & & & & \\ & \ddots & & & O & \\ & & 0 & & & \\ & & & \boxed{\begin{matrix} 0 & b_1 \\ -b_1 & 0 \end{matrix}} & & \\ & & & & \ddots & \\ & O & & & & \boxed{\begin{matrix} 0 & b_m \\ -b_m & 0 \end{matrix}} \end{bmatrix}$$

und für orthogonale Matrizen ergibt sich aus Satz 20 der Satz 16 (vgl. Bem. (c) von

S. 261). Im letzteren Fall gilt ja $c_i = \pm 1$ für $i = 1,...,k$ und $a_j^2 + b_j^2 = 1$, $b_j > 0$, für $j = 1,...,m$. Satz 16 erhält man, indem man $\sin \omega_j = b_j$ setzt und für jedes Zweier–Kästchen die zugehörigen Basisvektoren x_j, y_j vertauscht oder y_j durch $-y_j$ ersetzt (vgl. Bem. (a) von S. 260).

Aus dem oben gesagten ergibt sich damit eine weitere Methode, die Normalform und die Transformationsmatrix einer reellen Isometrie bzw. einer orthogonalen Matrix $A \in \mathbb{R}^{n\times n}$ aufzustellen:

Zu den reellen Eigenwerten 1 bzw. –1 bestimmen wir wie üblich reelle Orthonormalbasen in den zugehörigen Eigenräumen. Zu jedem komplexen Eigenwert $c = a + i\,b$, $b > 0$, sei $z_1,...,z_q$ ein Orthonormalsystem aus komplexen Eigenvektoren. Dann ist $\overline{z}_1,...,\overline{z}_q$ ein Orthonormalsystem aus komplexen Eigenvektoren zu $\overline{z}$. Setzen wir $\sin \omega = b$ und

$$x_j = \frac{\operatorname{Re} z_j}{\|\operatorname{Re} z_j\|}\,, \quad y_j = \frac{\operatorname{Im} z_j}{\|\operatorname{Im} z_j\|}\,, \quad j = 1,...,q\,,$$

so ist $(y_1, x_1, ..., y_q, x_q)$ oder $(x_1, -y_1, ..., x_q, -y_q)$ eine reelle ONB zu den q Zweierkästchen

$$\boxed{\begin{matrix} \cos\omega & -\sin\omega \\ \sin\omega & \cos\omega \end{matrix}}$$

Beispiel. Die Matrix

$$A = \begin{bmatrix} \sqrt{3}/4 + 1/2 & \sqrt{3}/4 - 1/2 & -\sqrt{2}/4 \\ \sqrt{3}/4 - 1/2 & \sqrt{3}/4 + 1/2 & -\sqrt{2}/4 \\ \sqrt{2}/4 & \sqrt{2}/4 & \sqrt{3}/2 \end{bmatrix}$$

ist orthogonal (Übungsaufgabe). Für das charakteristische Polynom p von A gilt

$$p = \det(A - X E_3) = \cdots = -X^3 + (1 + \sqrt{3})\, X^2 - (1 + \sqrt{3})\, X + 1\,.$$

Die Nullstellen sind $c_1 = 1$ sowie $c = \sqrt{3}/2 + i/2$ und $\bar{c} = \sqrt{3}/2 - i/2$. Die Normalform lautet also

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

(Drehwinkel : $\omega = \pi/6$, Drehachse : E_c , Drehebene : $E_{c_1}^{\perp}$) .

Bestimmung der Transformationsmatrix S :

$$v \in E_{c_1} \iff (A - E_3)\, v = o \iff v = \begin{bmatrix} t \\ -t \\ 0 \end{bmatrix}, \; t \in \mathbb{R} .$$

Also wählen wir

$$v := 1/\sqrt{2} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \in E_{c_1} .$$

$$z \in E_c \iff (A - (\sqrt{3}/2 + i/2)\, E_3)\, z = o \iff z = \begin{bmatrix} t \\ t \\ -\sqrt{2}\, i\, t \end{bmatrix}, \; t \in \mathbb{R} .$$

Dann ist

$$z = \begin{bmatrix} 1 \\ 1 \\ -\sqrt{2}\, i \end{bmatrix} \in E_c \;, \quad \bar{z} = \begin{bmatrix} 1 \\ 1 \\ \sqrt{2}\, i \end{bmatrix} \in E_{\bar{c}} ,$$

und das obige Verfahren liefert die orthonormierten Vektoren

$$\frac{\mathrm{Re}\, z}{\|\mathrm{Re}\, z\|} = 1/\sqrt{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \;, \quad \frac{\mathrm{Im}\, z}{\|\mathrm{Im}\, z\|} = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} .$$

Damit ergibt sich als orthogonale Transformationsmatrix

$$S = \begin{bmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & -1 & 0 \end{bmatrix} \quad \text{bzw.} \quad S = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Kapitel 6 Affine und euklidische Geometrie

§ 1 Affine und euklidische Räume

Bisher haben wir die geometrischen Objekte Punkte, Geraden, Ebenen usw. nur als Elemente bzw. Teilmengen eines Vektorraumes V kennengelernt. Die Doppelrolle der Elemente von V als Vektoren und als Punkte bringt aber gelegentlich Schwierigkeiten mit sich und entspricht auch nicht unserer geometrischen Vorstellung in der Zeichenebene oder im Anschauungsraum. Dort unterscheiden wir sehr wohl zwischen den "Punkten" und den zugehörigen "Verbindungsvektoren".

Wir werden daher im folgenden eine neue Struktur betrachten, den affinen Raum, der zur geometrischen Beschreibung des Anschauungsraumes besser geeignet ist, weil in ihm die Begriffe "Punkt" und "Vektor" getrennt behandelt werden. Wenn wir die Beziehungen zwischen affinen Räumen $\mathbb{A}$ und Vektorräumen V geklärt haben, können wir die bisher entwickelten Methoden und Ergebnisse der linearen Algebra auf geometrische Fragestellungen anwenden. Zur Vereinfachung der Darstellung werden wir dabei, nach Auszeichnung eines Ursprungs, in manchen Fällen wieder $\mathbb{A}$ und V identifizieren.

Definition. Gegeben seien eine nichtleere Menge $\mathbb{A}$, deren Elemente *Punkte* heißen, ein $\mathbb{K}$–Vektorraum V, sowie eine Abbildung $f\colon \mathbb{A} \times \mathbb{A} \longrightarrow V$. Das Tripel $(\mathbb{A}, V, f)$ heißt *affiner Punktraum* oder *affiner Raum mit zugehörigem Vektorraum* V, falls folgende Bedingungen erfüllt sind:

(a) Zu jedem $P \in \mathbb{A}$ und jedem $x \in V$ gibt es genau ein $Q \in \mathbb{A}$ mit $f(P,Q) = x$.

(b) Für alle P, Q, $R \in \mathbb{A}$ gilt $f(P,Q) + f(Q,R) = f(P,R)$.

Statt $(\mathbb{A}, V, f)$ schreiben wir kurz $\mathbb{A}$ und sprechen vom affinen Raum $\mathbb{A}$, wenn der zugehörige Vektorraum V und die Abbildung f klar sind.

Unter der *Dimension* von $\mathbb{A}$ verstehen wir die Dimension des zugehörigen Vektorraumes V, Schreibweise: $\dim \mathbb{A}$.

Ist $\mathbb{K} = \mathbb{R}$ bzw. $\mathbb{K} = \mathbb{C}$, so sprechen wir von einem *reellen* bzw. *komplexen* affinen Raum. Ist der Vektorraum V euklidisch, so heißt $\mathbb{A}$ *euklidischer Raum.*

Die Gesamtheit der Definitionen und Sätze, die sich aus den Axiomen für einen affinen Raum herleiten lassen, heißt *affine Geometrie.* Analog sprechen wir auch von *euklidischer Geometrie.* Jeder Satz der affinen Geometrie ist auch ein Satz der euklidischen Geometrie, umgekehrt lassen sich aber viele Sätze der euklidischen Geometrie affin gar nicht formulieren. Dies gilt zum Beispiel für den aus der Schule bekannten Satz, daß sich in jedem Dreieck die Höhen in einem gemeinsamen Punkt schneiden.

In den folgenden Paragraphen werden wir zunächst immer erst den gemeinsamen affinen Anteil der Geometrie darstellen und danach zusätzliche Aussagen im Rahmen der euklidischen Geometrie beweisen. Wir beginnen mit einigen Bemerkungen zum Axiomensystem eines affinen Raumes.

Bemerkungen. **(a)** Die Existenz der Abbildung f bedeutet, daß je zwei Punkte P und Q genau einen *Verbindungsvektor* $f(P,Q)$ besitzen. Um dies auch in der Schreibweise deutlich zu machen, schreiben wir deshalb statt $f(P,Q)$ meistens $\overrightarrow{PQ}$.

Da f im allgemeinen nicht injektiv ist, gibt es mehrere Punktepaare mit dem gleichen Verbindungsvektor. Die Vektoren spielen in dem affinen Raum gerade die Rolle, die die Äquivalenzklassen der Pfeile im Anschauungsraum spielten.

(b) Axiom (a) besagt, daß der Vektor x im Punkt P "abgetragen" zu einem eindeutig bestimmten Punkt Q führt, und Axiom (b) bedeutet, daß das Abtragen von Vektoren an Punkten mit der Vektoraddition verträglich ist.

(c) Einige einfache Folgerungen sind:

$$\overrightarrow{PQ} = o \iff P = Q\,,$$
$$\overrightarrow{PQ} = -\overrightarrow{QP}\,,$$
$$\overrightarrow{PQ} = \overrightarrow{PR} \Longrightarrow Q = R\;;\quad \overrightarrow{QS} = \overrightarrow{RS} \Longrightarrow Q = R\,.$$

Beispiele. **(a)** Wir können den Anschauungsraum als affinen Raum auffassen. Der zugehörige Vektorraum ist dann der auf S.93 betrachtete Vektorraum der Pfeilklassen.

(b) Jeder $\mathbb{K}$-Vektorraum V wird selbst zu einem affinen Raum mit zugehörigem Vektorraum V, wenn wir für beliebige Punkte $x,y \in V$ den Verbindungsvektor durch $\overrightarrow{xy} := y - x$ definieren. Insbesondere ist $\mathbb{K}^n$ ein affiner Raum.

So wie wir jeden Vektorraum als affinen Raum betrachten können, gehört umgekehrt in einem affinen Raum $\mathbb{A}$ nach Auszeichnung eines beliebigen Punktes $O \in \mathbb{A}$ zu jedem Punkt X ein eindeutig bestimmter Vektor $\overrightarrow{OX}$. Es gilt $V = \{\overrightarrow{OX} \mid X \in \mathbb{A}\}$, und die durch $X \longmapsto \overrightarrow{OX}$ erklärte Abbildung von $\mathbb{A}$ auf V ist bijektiv. Wir nennen $\overrightarrow{OX}$ den *Ortsvektor* des Punktes X bezüglich des *Ursprungs* O. Affine Räume und Vektorräume unterscheiden sich also im wesentlichen nur durch die Auszeichnung eines Ursprungs.

Analog wollen wir nun die affinen Unterräume von $\mathbb{A}$ so definieren, daß sie sich nur durch Auszeichnung eines Ursprunges von den Untervektorräumen von V unterscheiden.

Definition. Eine nichtleere Teilmenge L eines affinen Raumes $\mathbb{A}$ heißt *affiner Unterraum* von $\mathbb{A}$, wenn es einen Punkt $P \in L$ gibt, so daß die Menge $\{\overrightarrow{PX} \mid X \in L\}$ ein Untervektorraum von V ist.

Die affinen Unterräume eines euklidischen Raumes heißen *euklidische Unterräume.*

Bemerkungen und Bezeichnungen. **(a)** Ist L affiner Unterraum von $\mathbb{A}$, so ist für jeden Punkt $Q \in L$ die Menge $\{\overrightarrow{QX} \mid X \in L\}$ ein Untervektorraum von V und es gilt

$$\{\overrightarrow{QX} \mid X \in L\} = \{\overrightarrow{XY} \mid X, Y \in L\}.$$

Somit gehört zu L in eindeutiger Weise der Untervektorraum $U_L = \{\overrightarrow{XY} \mid X, Y \in L\}$, der dann *Richtungsraum* oder kurz *Richtung* von L heißt.

Beweis. Nach Definition gibt es einen Punkt $P \in L$, so daß $U = \{\overrightarrow{PX} \mid X \in L\}$ ein Untervektorraum von V ist. Sei $W = \{\overrightarrow{QX} \mid X \in L\}$. Wegen $Q \in L$ ist $\overrightarrow{QP} = -\overrightarrow{PQ} \in U$, also gilt $W = \overrightarrow{QP} + U = U$, und W ist ebenfalls ein Untervektorraum. Daß die Menge $\{\overrightarrow{XY} \mid X, Y \in L\}$ mit dem Untervektorraum W übereinstimmt, ist offensichtlich. ∎

(b) Ist L affiner Unterraum von $\mathbb{A}$ und ist f_L die auf $L \times L$ eingeschränkte Abbildung f, so erfüllt das Tripel (L, U_L, f_L) die Definition eines affinen Raumes. Damit gilt $\dim L = \dim U_L$.

Die nulldimensionalen affinen Unterräume von $\mathbb{A}$ sind genau die Mengen $\{P\}$, $P \in \mathbb{A}$, bestehen also aus einem einzigen Punkt. Man nennt sie deshalb ebenfalls *Punkte*. Die eindimensionalen affinen Unterräume heißen *Geraden*, die zweidimensionalen affinen Unterräume heißen *Ebenen*.

(c) Jeder affine Unterraum $L \subset \mathbb{A}$ läßt sich nach Auszeichnung eines Ursprungs $O \in \mathbb{A}$ und eines Punktes $P \in L$ vektoriell in der Form

$$\{\overrightarrow{OX} \mid X \in L\} = \overrightarrow{OP} + U_L$$

darstellen.

Speziell für $\mathbb{A} = V$ und $O = o$ gilt: Jeder affine Unterraum $L \subset V$ ist von der Form $L = x_0 + U$ mit einen Untervektorraum $U \subset V$ und $x_0 \in L$ (vgl. S. 126).

Wie in § 2.5 ist der Durchschnitt affiner Unterräume entweder leer oder ein affiner Unterraum.

Satz 1. *Es seien $\mathbb{A}$ ein affiner Raum und $\mathcal{M}$ eine nichtleere Menge affiner Unterräume von $\mathbb{A}$. Dann ist der Schnitt*

$$M = \bigcap_{L \in \mathcal{M}} L$$

entweder leer oder ein affiner Unterraum mit Richtungsraum

$$U_M = \bigcap_{L \in \mathcal{M}} U_L .$$

Beweis. Sei $M \neq \emptyset$ und $P \in M$. Dann gilt

$$U_M = \{\overrightarrow{PX} \mid X \in M\} = \bigcap_{L \in \mathcal{M}} \{\overrightarrow{PX} \mid X \in L\} = \bigcap_{L \in \mathcal{M}} U_L \,. \qquad \blacksquare$$

Definition. Ist $C \subset \mathbb{A}$ nichtleer, so heißt der Durchschnitt aller affinen Unterräume L, die C enthalten, die *affine Hülle* von C.

Ist $C = \{P_1,\ldots,P_k\}$, so nennen wir die affine Hülle von C den *Verbindungsraum* der Punkte $P_1,\ldots,P_k$ und schreiben dafür $P_1 \vee\ldots\vee P_k$. Allgemeiner heißt die affine Hülle der Vereinigungsmenge endlich vieler affiner Unterräume $L_1,\ldots,L_k$ der *Verbindungsraum* von $L_1,\ldots,L_k$, Schreibweise: $L_1 \vee\ldots\vee L_k$.

Auch der Verbindungsraum endlich vieler affiner Unterräume läßt sich genauer beschreiben.

Satz 2. *Es seien $L_1,\ldots,L_k \subset \mathbb{A}$ affine Unterräume und $P_i \in L_i$ beliebig gewählt, $i = 1,\ldots,k$. Dann gilt für den Richtungsraum des Verbindungsraumes $L = L_1 \vee\ldots\vee L_k$*

$$U_L = U_{L_1} + \cdots + U_{L_k} + [\overrightarrow{P_1P_2},\ldots,\overrightarrow{P_1P_k}] \,.$$

Ist $L_1 \cap\ldots\cap L_k \neq \emptyset$, so gilt

$$U_L = U_{L_1} + \cdots + U_{L_k} \,.$$

Beweis. Seien $\tilde{U} = U_{L_1} + \cdots + U_{L_k} + [\overrightarrow{P_1P_2},\ldots,\overrightarrow{P_1P_k}]$, $\tilde{L} = \{X \in \mathbb{A} \mid \overrightarrow{P_1X} \in \tilde{U}\}$. Aus $X \in L_i$ folgt $\overrightarrow{P_1X} = \overrightarrow{P_1P_i} + \overrightarrow{P_iX} \in \tilde{U}$, also $X \in \tilde{L}$, $i = 1,\ldots,k$. Somit gilt $L \subset \tilde{L}$ und $U_L \subset \tilde{U}$. Die umgekehrte Inklusion ist offensichtlich.

Ist $L_1 \cap\ldots\cap L_k \neq \emptyset$, so wählen wir $P \in L_1 \cap\ldots\cap L_k$, setzen $P_1 = \ldots = P_k = P$ und erhalten $U_L = U_{L_1} + \cdots + U_{L_k}$. $\blacksquare$

Beispiele. **(a)** Der Verbindungsraum von $k + 1$ Punkten $P_0,\ldots,P_k$ besitzt den Richtungsraum $U = [\overrightarrow{P_0P_1},\ldots,\overrightarrow{P_0P_k}]$, und es gilt $\dim U \leq k$.

Für $k = 1$ und $P_0 \neq P_1$ gilt $\dim U = 1$, also ist $P_0 \vee P_1$ eine Gerade; wir schreiben dann kürzer P_0P_1 statt $P_0 \vee P_1$.

Für $k = 2$ und drei verschiedene Punkte P_0, P_1, P_2, die nicht auf einer Geraden liegen, gilt $\dim U = 2$, also ist der Verbindungsraum $P_0 \vee P_1 \vee P_2$ eine Ebene.

(b) Im $\mathbb{R}^3$ seien der Punkt x und die Gerade L gegeben mit

$$x = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \text{ und } L = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + [\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}].$$

Der Verbindungsraum $x \vee L$ hat die Darstellung

$$x \vee L = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + ([\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}] + [\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}]) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + [\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}].$$

(c) Im $\mathbb{R}^5$ seien die Punkte x_1, x_2, x_3, x_4, x_5 gegeben mit

$$x_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \\ -1 \\ -1 \end{bmatrix}, x_2 = \begin{bmatrix} 3 \\ 8 \\ -5 \\ -6 \\ 1 \end{bmatrix}, x_3 = \begin{bmatrix} 3 \\ -4 \\ 3 \\ 5 \\ -3 \end{bmatrix}, x_4 = \begin{bmatrix} 2 \\ -1 \\ -1 \\ 2 \\ -2 \end{bmatrix}, x_5 = \begin{bmatrix} 0 \\ 5 \\ -3 \\ -4 \\ 0 \end{bmatrix}.$$

Für ihren Verbindungsraum gilt

$$x_1 \vee \cdots \vee x_5 = x_1 + [x_2 - x_1, x_3 - x_1, x_4 - x_1, x_5 - x_1]$$

$$= \begin{bmatrix} 1 \\ 2 \\ -1 \\ -1 \\ -1 \end{bmatrix} + [\begin{bmatrix} 2 \\ 6 \\ -4 \\ -5 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -6 \\ 4 \\ 6 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ -3 \\ 0 \\ 3 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 3 \\ -2 \\ -3 \\ 1 \end{bmatrix}].$$

Wir wollen im Richtungsraum von $x_1 \vee \cdots \vee x_5$ eine Basis angeben. Dazu verwenden wir wieder das Verfahren von S.114 und erhalten

$$x_1 \vee \cdots \vee x_5 = \begin{bmatrix} 1 \\ 2 \\ -1 \\ -1 \\ -1 \end{bmatrix} + [\begin{bmatrix} -1 \\ 3 \\ 0 \\ -3 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 12 \\ 0 \\ -11 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}].$$

Also ist $x_1 \vee \cdots \vee x_5$ ein dreidimensionaler affiner Unterraum von $\mathbb{R}^5$.

Für den Durchschnitt und den Verbindungsraum zweier affiner Unterräume gilt die folgende Dimensionsformel.

Satz 3. *Es seien $\mathbb{A}$ ein affiner Raum und L_1, L_2 affine Unterräume von $\mathbb{A}$ mit zugehörigen Richtungsräumen U_1, U_2. Dann gilt für $L_1 \cap L_2 \neq \emptyset$*

$$\dim L_1 + \dim L_2 = \dim (L_1 \vee L_2) + \dim (L_1 \cap L_2),$$

und für $L_1 \cap L_2 = \emptyset$ ist

$$\dim L_1 + \dim L_2 = \dim (L_1 \vee L_2) + \dim (U_1 \cap U_2) - 1.$$

Beweis. Nach Satz 2.19 gilt $\dim L_1 + \dim L_2 = \dim U_1 + \dim U_2 = \dim (U_1 \cap U_2) + \dim (U_1 + U_2)$. Im Fall $L_1 \cap L_2 \neq \emptyset$ ist nach Satz 2 $U_1 + U_2$ der Richtungsraum von $L_1 \vee L_2$. Da $U_1 \cap U_2$ dann auch der Richtungsraum von $L_1 \cap L_2$ ist, folgt der 1. Teil der Behauptung.

Ist $L_1 \cap L_2 = \emptyset$, so wählen wir $P_1 \in L_1$, $P_2 \in L_2$ und nach Satz 2 gilt $\dim (L_1 \vee L_2) = \dim (U_1 + U_2 + [\overrightarrow{P_1P_2}])$. Wäre $\overrightarrow{P_1P_2} \in U_1 + U_2$, so gäbe es Punkte $Q_1 \in L_1$, $Q_2 \in L_2$ mit $\overrightarrow{P_1P_2} = \overrightarrow{P_1Q_1} + \overrightarrow{P_2Q_2}$. Also wäre $\overrightarrow{P_2Q_2} = \overrightarrow{Q_1P_2}$ und somit $Q_1 \in L_1 \cap L_2$. Daher gilt $\dim (U_1 + U_2 + [\overrightarrow{P_1P_2}]) = \dim (U_1 + U_2) + 1$, woraus die 2. Behauptung folgt. ■

Bezeichnung und Bemerkung. Mit Hilfe des Verbindungsraumes läßt sich eine weitere Klasse affiner Unterräume definieren: Es sei L ein von $\mathbb{A}$ verschiedener affiner Unterraum. Gibt es einen Punkt $P \in \mathbb{A}$, so daß der Verbindungsraum von L und P gleich $\mathbb{A}$ ist, so heißt L eine *Hyperebene* von $\mathbb{A}$. Nach Satz 2 kann P nicht in L liegen. Ist $\mathbb{A}$ n-dimensional, so sind wegen Satz 3 die Hyperebenen genau die affinen Unterräume der Dimension $n-1$.

Affine Unterräume können parallel zueinander sein. Wie auf S. 129 nennen wir zwei affine Unterräume L_1, L_2 *parallel*, Schreibweise $L_1 \parallel L_2$, falls für die zugehörigen

Richtungsräume U_1, U_2 gilt: $U_1 \subset U_2$ oder $U_2 \subset U_1$. Weiterhin heißen zwei Geraden *windschief*, falls sie weder parallel sind noch einen Punkt gemeinsam haben.

Beispiel. In einem n–dimensionalen affinen Raum $\mathbb{A}$ kann eine Gerade g in bezug auf eine Hyperebene L folgende Lagen einnehmen: g ist entweder Teilmenge von L oder g schneidet L in einem Punkt oder g ist parallel zu L und der Schnitt ist leer (Übungsaufgabe).

Als nächstes wollen wir den Begriff der linearen Unabhängigkeit auf affine Räume übertragen.

Definition. Es seien $\mathbb{A}$ ein affiner Raum, $k \in \mathbb{N}_0$ und $P_0,\ldots,P_k \in \mathbb{A}$. Die Punkte $P_0,\ldots,P_k$ heißen *affin unabhängig* oder *in allgemeiner Lage*, wenn $\dim (P_0 \vee \ldots \vee P_k) = k$ ist. Sind die Punkte nicht affin unabhängig, so heißen sie *affin abhängig.*

Eine Teilmenge $C \subset \mathbb{A}$ heißt *affin unabhängig*, wenn für jedes $k \in \mathbb{N}$ alle paarweise verschiedenen Punkte $P_0,\ldots,P_k$ aus C affin unabhängig sind. Ist C nicht affin unabhängig, so heißt C *affin abhängig.*

Beispiele. Jeder Punkt ist affin unabhängig. Zwei Punkte sind genau dann affin unabhängig, wenn sie verschieden sind. Drei Punkte sind genau dann affin unabhängig, wenn sie nicht auf einer gemeinsamen Geraden liegen, d.h. wenn sie nicht kollinear sind, u.s.w.

Satz 4. *Es seien $\mathbb{A}$ ein affiner Raum und $P_0,\ldots,P_k \in \mathbb{A}$. Dann sind folgende Aussagen äquivalent:*

(a) *Die Punkte $P_0,\ldots,P_k$ sind affin unabhängig.*

(b) *Die Vektoren $\overrightarrow{P_0P_1},\ldots,\overrightarrow{P_0P_k}$ sind linear unabhängig.*

Beweis. Für den Richtungsraum U von $P_0 \vee \ldots \vee P_k$ gilt $U = [\overrightarrow{P_0P_1},\ldots,\overrightarrow{P_0P_k}]$. Sind die Punkte $P_0,\ldots,P_k$ affin unabhängig, so gilt $\dim U = k$ und die Vektoren $\overrightarrow{P_0P_1},\ldots,\overrightarrow{P_0P_k}$ sind linear unabhängig. Sind umgekehrt die Vektoren $\overrightarrow{P_0P_1},\ldots,\overrightarrow{P_0P_k}$ linear unabhängig, so ist $U = [\overrightarrow{P_0P_1},\ldots,\overrightarrow{P_0P_k}]$ ein k–dimensionaler Untervektorraum und es gilt somit $\dim(P_0 \vee \ldots \vee P_k) = k$. ■

Bemerkungen. **(a)** Satz 4 gilt ganz entsprechend, wenn der Bezugspunkt P_0 durch irgend einen anderen der Punkte P_i ersetzt wird, wenn man also die Vektoren $\overrightarrow{P_iP_0},\ldots,\overrightarrow{P_iP_{i-1}}, \overrightarrow{P_iP_{i+1}},\ldots,\overrightarrow{P_iP_k}$ betrachtet.

(b) Ist $\mathbb{A} = V$, so gilt außerdem: Die Punkte $x_0,\ldots,x_k$ sind genau dann affin unabhängig, wenn aus $a_0\, x_0 + \cdots + a_k\, x_k = o$ und $a_0 + \cdots + a_k = 0$ stets $a_0 = \ldots = a_k = 0$ folgt.

Beweis. Seien $x_0,\ldots,x_k$ affin unabhängig, also $x_1 - x_0,\ldots,x_k - x_0$ linear unabhängig. Dann folgt aus $a_0\, x_0 + \cdots + a_k\, x_k = o$ und $a_0 + \cdots + a_k = 0$ unmittelbar $a_1 (x_1 - x_0) + \cdots + a_k (x_k - x_0) = o$ und daraus $a_1 = \ldots = a_k = 0$. Dann ist auch $a_0 = 0$.

Umgekehrt folgt aus $a_1 (x_1 - x_0) + \cdots + a_k (x_k - x_0) = o$ zunächst $(-a_1 - \cdots - a_k)\, x_0 + a_1\, x_1 + \cdots + a_k\, x_k = o$ und $a_0 + \cdots + a_k = 0$ mit $a_0 := -a_1 - \cdots - a_k$. Also ist $a_1 = \ldots = a_k = 0$, und die Vektoren $x_1 - x_0,\ldots,x_k - x_0$ somit linear unabhängig, woraus die Behauptung folgt. ■

(c) Sind die Punkte $P_0,\ldots,P_k$ affin unabhängig und ist $\{i_1,\ldots,i_m\} \subset \{0,\ldots,k\}$, so sind auch die Punkte $P_{i_1},\ldots,P_{i_m}$ affin unabhängig.

(d) Sind die Punkte $P_0,\ldots,P_k$ gegeben und sind für $\{i_1,\ldots,i_m\} \subset \{0,..,k\}$ die Punkte $P_{i_1},\ldots,P_{i_m}$ affin abhängig, so sind auch die Punkte $P_0,\ldots,P_k$ affin abhängig.

(e) Sind die Punkte $P_0,\ldots,P_k$ affin abhängig, so gilt $\dim(P_0 \vee \cdots \vee P_k) < k$.

(f) Jeder k-dimensionale affine Unterraum L ist affine Hülle von $k + 1$ affin unabhängigen Punkten.

Beweis. Wir wählen in L einen beliebigen Punkt P_0 und im Richtungsraum U_L eine beliebige Basis $\{x_1,\ldots,x_k\}$. Dann existieren Punkte $P_1,\ldots,P_k$ aus L mit $\overrightarrow{P_0P_i} = x_i$, $i = 1,\ldots,k$. Nach Satz 4 sind die Punkte $P_0,\ldots P_k$ affin unabhängig, und ihre affine Hülle ist L. ■

Mit Hilfe der affinen Unabhängigkeit lassen sich nun auch affine Koordinatensysteme definieren.

Definition. Es seien L ein k–dimensionaler affiner Unterraum von $\mathbb{A}$ und $P_0,\ldots,P_k$ $k+1$ affin unabhängige Punkte in L.

Dann heißt $(P_0;\ \overrightarrow{P_0P_1},\ldots,\overrightarrow{P_0P_k})$ ein *affines Koordinatensystem* von L mit dem *Koordinatenursprung* P_0 , den *Koordinatenachsen* P_0P_i und den *Einheitspunkten* P_i auf diesen Achsen, $i = 1,\ldots,k$.

Für jeden Punkt $P \in L$ heißen die Koordinaten $a_1,\ldots,a_k$ des Vektors $\overrightarrow{P_0P}$ bezüglich der Basis $(\overrightarrow{P_0P_1},\ldots,\overrightarrow{P_0P_k})$ von U_L die *affinen Koordinaten* des Punktes P bezüglich des affinen Koordinatensystems $(P_0;\ \overrightarrow{P_0P_1},\ldots,\overrightarrow{P_0P_k})$. Wir schreiben dann auch $P(a_1,\ldots,a_k)$.

In einem n–dimensionalen affinen Raum $\mathbb{A}$ erhalten wir also für jede Wahl eines Ursprungs $O \in \mathbb{A}$ und jede Wahl einer Basis $(v_1,\ldots,v_n)$ von V ein affines Koordinatensystem $(O;\ v_1,\ldots,v_n)$ von $\mathbb{A}$.

Ist $\mathbb{A}$ ein euklidischer Raum und ist $(v_1,\ldots,v_n)$ eine ONB von V, so sprechen wir von einem *kartesischen Koordinatensystem* $(O;\ v_1,\ldots,v_n)$ von $\mathbb{A}$, und die affinen Koordinaten eines Punktes X heißen dann entsprechend *kartesische Koordinaten.*

Es seien nun wieder L ein k–dimensionaler affiner Unterraum von $\mathbb{A}$ und $(P_0;\ \overrightarrow{P_0P_1},\ldots,\overrightarrow{P_0P_k})$ ein affines Koordinatensystem in L. Dann gilt für die Ortsvektoren der Punkte $P \in L$ bezüglich eines beliebigen Ursprungs $O \in \mathbb{A}$

$$(*)\qquad \overrightarrow{OP} = \overrightarrow{OP_0} + a_1\,\overrightarrow{P_0P_1} + \ldots + a_k\,\overrightarrow{P_0P_k}\,,$$

wobei $a_1,\ldots,a_k$ die affinen Koordinaten von P bezüglich des ausgezeichneten Koordinatensystems sind. Wir nennen (*) eine *Parameterdarstellung 1.Art* von L mit den *Richtungsvektoren* $\overrightarrow{P_0P_1},\ldots,\overrightarrow{P_0P_k}$ und den *Parametern* $a_1,\ldots,a_k$ (vgl. S.127).

In dieser Parameterdarstellung von L ist der Punkt P_0 vor den anderen Punkten $P_1,\ldots,P_k$ ausgezeichnet. Eine symmetrische Darstellung können wir erreichen, indem wir (*) etwas umformen. Wir setzen $a_0 := 1 - a_1 - \ldots - a_k$ und erhalten

$$\overrightarrow{OP} = (a_0 + \ldots + a_k)\,\overrightarrow{OP_0} + a_1\,\overrightarrow{P_0P_1} + \ldots + a_k\,\overrightarrow{P_0P_k}\,.$$

Also gilt

$$(**) \qquad \overrightarrow{OP} = a_0\,\overrightarrow{OP_0} + a_1\,\overrightarrow{OP_1} + \ldots + a_k\,\overrightarrow{OP_k}\ , \ a_i \in \mathbb{K},\ a_0 + \ldots + a_k = 1\,.$$

Wir nennen (**) eine *Parameterdarstellung 2.Art* von L. Die eindeutig bestimmten Skalare $a_0,\ldots,a_k$, die wiederum unabhängig von der Wahl von O sind, heißen *Schwerpunktskoordinaten* oder *baryzentrische Koordinaten* des Punktes P bezüglich $P_0,\ldots,P_k$. Ist char $\mathbb{K}$ kein Teiler von $k+1$, so heißt der Punkt S mit den baryzentrischen Koordinaten $a_0 = \ldots = a_k = \frac{1}{k+1}$ *Schwerpunkt* (*Baryzentrum*) von $P_0,\ldots,P_k$.

Die Darstellung (**) des Ortsvektors $\overrightarrow{OP}$ bezeichnet man auch als *Affinkombination* der Vektoren $\overrightarrow{OP_0},\ldots,\overrightarrow{OP_k}$. Entsprechend heißt $P \in \mathbb{A}$ *Affinkombination* der Punkte $P_0,\ldots,P_k \in \mathbb{A}$, wenn es einen Punkt $O \in \mathbb{A}$ gibt und Zahlen $a_0,\ldots,a_k \in \mathbb{K}$ mit $a_0 + \ldots + a_k = 1$, für die (**) erfüllt ist. In diesem Fall gilt (**) dann für alle $O \in \mathbb{A}$. Die Punkte $P_0,\ldots,P_k$ müssen dabei nicht affin unabhängig sein. Sind sie es aber, so sind die Zahlen $a_0,\ldots,a_k$ die baryzentrischen Koordinaten von P bezüglich $P_0,\ldots,P_k$, also eindeutig bestimmt.

Bemerkungen. **(a)** $P \in \mathbb{A}$ ist genau dann Affinkombination der Punkte $P_0,\ldots,P_k \in \mathbb{A}$, wenn es Zahlen $a_0,\ldots,a_k \in \mathbb{K}$ gibt mit $a_0 + \ldots + a_k = 1$ und

$$a_0\,\overrightarrow{PP_0} + \ldots + a_k\,\overrightarrow{PP_k} = o\,.$$

(b) Sei $P \in \mathbb{A}$ und seien $P_0,\ldots,P_k \in \mathbb{A}$ affin unabhängig. Dann ist P genau dann Affinkombination der Punkte $P_0,\ldots,P_k$, wenn $P_0,\ldots,P_k,P$ affin abhängig sind.

(c) Sei $C \subset \mathbb{A}$ nichtleer. Dann ist die affine Hülle von C die Menge aller Affinkombinationen von Punkten aus C.

(Beweise als Übung)

Als nächstes wollen wir uns mit metrischen Problemen wie Orthogonalität und Abstandsbestimmungen beschäftigen. Hierzu setzen wir voraus, daß $\mathbb{A}$ ein endlich dimensionaler euklidischer Raum ist.

Orthogonale Unterräume

Bei der Definition orthogonaler Unterräume orientieren wir uns an den Verhältnissen im Anschauungsraum. Um Randfälle zu vermeiden, die unserer Anschauung widersprechen, setzen wir voraus, daß die betrachteten Unterräume L_1 und L_2 nicht parallel sind.

Es wäre naheliegend, L_1 und L_2 orthogonal zu nennen, wenn ihre zugehörigen Richtungsräume U_1 und U_2 orthogonal sind. Hierzu müßte $U_1 \cap U_2 = \{o\}$ gelten. Aus geometrischen Gründen wollen wir aber in gewissen Fällen, in denen $U_1 \cap U_2 \neq \{o\}$ gilt, ebenfalls von orthogonalen Unterräumen sprechen. Hierzu betrachten wir die orthogonalen Komplemente $W_1 = (U_1 \cap U_2)^\perp \cap U_1$ bzw. $W_2 = (U_1 \cap U_2)^\perp \cap U_2$ des Schnittes $U_1 \cap U_2$ in U_1 bzw. in U_2. Weil nach Voraussetzung weder U_1 in U_2 noch U_2 in U_1 gelegen ist, gilt $W_i \neq \{o\}$, $i = 1,2$.

Wir nennen nun die Unterräume L_1 und L_2 von $\mathbb{A}$ *orthogonal,* in Zeichen $L_1 \perp L_2$, falls $W_1 \perp W_2$ gilt.

Ist L_1 speziell eine Gerade und $L_1 \cap L_2 \neq \emptyset$, so heißt L_1 auch ein *Lot* auf L_2 .

Beispiel. Im euklidischen Raum $\mathbb{R}^3$ sind bei einem kartesischen Koordinatensystem sowohl die Koordinatenachsen als auch die Koordinatenebenen paarweise orthogonal. Jede Koordinatenachse ist orthogonal zu jeder Koordinatenebene, die durch die beiden anderen Koordinatenachsen aufgespannt wird, und sie ist auch orthogonal zu jeder Geraden in dieser Ebene.

Bemerkungen. **(a)** Für $U_1 \cap U_2 = \{o\}$ ist $W_1 = U_1$ und $W_2 = U_2$, also sind L_1 und L_2 in diesem Fall genau dann orthogonal, wenn es ihre Richtungsräume sind.

(b) Orthogonale Unterräume sind nach Definition nie parallel, müssen sich allerdings auch nicht schneiden. So gibt es, wie wir gesehen haben, im euklidischen $\mathbb{R}^3$ orthogonale windschiefe Geraden.

(c) Jedes Lot g auf einen Unterraum L hat mit diesem genau einen Punkt gemeinsam; dieser Punkt heißt dann der *Lotfußpunkt* von g auf L.

Abstand euklidischer Unterräume

Im euklidischen Fall übertragen sich die metrischen Eigenschaften des Vektorraumes V direkt auf die Punkte von $\mathbb{A}$.

Definition. Unter dem *Abstand* der Punkte $X, Y \in \mathbb{A}$ verstehen wir die Länge ihres Verbindungsvektors,

$$d(X,Y) := \| \overrightarrow{XY} \| .$$

Der euklidische Raum $\mathbb{A}$ wird damit ein metrischer Raum. Ist $\mathbb{A}$ n-dimensional und sind $(x_1,\ldots,x_n)$ und $(y_1,\ldots,y_n)$ die kartesischen Koordinaten der Punkte X und Y bezüglich eines kartesischen Koordinatensystems $(O;\, v_1,\ldots,v_n)$, so gilt

$$d(X,Y) = \Big(\sum_{i=1}^{n} (y_i - x_i)^2 \Big)^{1/2} .$$

Wir wollen die Definition des Abstandes auf beliebige Teilmengen M_1, M_2 von $\mathbb{A}$ übertragen.

Definition. Unter dem *Abstand* der Teilmengen M_1, M_2 des euklidischen Raumes $\mathbb{A}$ verstehen wir die Zahl

$$d(M_1,M_2) := \inf \{ d(X_1,X_2) \mid X_1 \in M_1 ,\, X_2 \in M_2 \}.$$

Speziell für Unterräume eines endlich dimensionalen euklidischen Raumes läßt sich der Abstand mit Hilfe von Orthogonalprojektionen konkret angeben.

Satz 5. *Es seien L_1, L_2 euklidische Unterräume mit zugehörigen Richtungsräumen U_1, U_2, und $P_1 \in L_1$ sowie $P_2 \in L_2$ seien beliebig gewählt. Dann gilt*

$$d(L_1,L_2) = d(\overrightarrow{P_1P_2} , U_1 + U_2) = \| \overrightarrow{P_1P_2} - \pi_{U_1+U_2} (\overrightarrow{P_1P_2}) \| .$$

Beweis. Für alle Punkte $X_1 \in L_1$ und $X_2 \in L_2$ gilt

$$d(X_1,X_2) = \| \overrightarrow{X_1X_2} \| = \| \overrightarrow{X_1P_1} + \overrightarrow{P_1P_2} + \overrightarrow{P_2X_2} \| = \| \overrightarrow{P_1P_2} - \overrightarrow{P_1X_1} + \overrightarrow{P_2X_2} \| ,$$

also

$$d(L_1,L_2) = \inf_{\substack{x_1 \in U_1 \\ x_2 \in U_2}} \| \overrightarrow{P_1P_2} - x_1 + x_2 \| = \| \overrightarrow{P_1P_2} - \pi_{U_1+U_2}(\overrightarrow{P_1P_2}) \| . \qquad ■$$

Bemerkungen. **(a)** Die Formel von Satz 5 gibt nur den Abstand zwischen L_1 und L_2 an, aber keine Punkte $Q_1 \in L_1$ und $Q_2 \in L_2$, für die der Abstand angenommen wird, für die also $d(Q_1,Q_2) = d(L_1,L_2)$ gilt. Diese erhalten wir folgendermaßen:

Ist $\pi_{U_1+U_2}(\overrightarrow{P_1P_2}) = x_1 + x_2$, $x_1 \in U_1$, $x_2 \in U_2$, so seien Q_1 bzw. Q_2 die eindeutig bestimmten Punkte mit $x_1 = \overrightarrow{P_1Q_1}$ bzw. $-x_2 = \overrightarrow{P_2Q_2}$. Es gilt dann $d(L_1,L_2) = \| \overrightarrow{P_1P_2} - \overrightarrow{P_1Q_1} + \overrightarrow{P_2Q_2} \| = \| \overrightarrow{Q_1Q_2} \|$.

(b) Sind die Unterrräume L_1 und L_2 disjunkt, so ist $Q_1 \neq Q_2$, und die Gerade Q_1Q_2 ist ein gemeinsames Lot von L_1 und L_2:

Aus $d(Q_1,Q_2) = d(L_1,L_2)$ folgt nämlich mit Satz 5 und dem Satz von Pythagoras

$$\| \overrightarrow{Q_1Q_2} \|^2 = \| \overrightarrow{Q_1Q_2} - \pi_{U_1+U_2}(\overrightarrow{Q_1Q_2}) \|^2 = \| \overrightarrow{Q_1Q_2} \|^2 - \| \pi_{U_1+U_2}(\overrightarrow{Q_1Q_2}) \|^2 .$$

Also ist $\pi_{U_1+U_2}(\overrightarrow{Q_1Q_2}) = o$ und somit $\overrightarrow{Q_1Q_2} \perp U_1 + U_2$. Daraus folgt $Q_1Q_2 \perp L_1$ und $Q_1Q_2 \perp L_2$.

Umgekehrt gilt für die Lotfußpunkte Q_1 und Q_2 eines beliebigen gemeinsamen Lotes von L_1 und L_2 offensichtlich $d(Q_1,Q_2) = d(L_1,L_2)$.

(c) Das Punktepaar $(Q_1,Q_2) \in L_1 \times L_2$ der Lotfußpunkte ist genau dann eindeutig bestimmt, wenn die Summe $U_1 + U_2$ direkt ist. Dies wiederum ist genau dann der Fall, wenn L_1 und L_2 keine parallelen Geraden enthalten (Übungsaufgabe).

(d) Speziell für den Abstand eines Punktes X von einem Unterraum L erhalten wir nach Wahl eines beliebigen Punktes $P \in L$

$$d(X,L) = \| \overrightarrow{PX} - \pi_{U_L}(\overrightarrow{PX}) \| .$$

Beispiel. Im euklidischen $\mathbb{R}^5$ seien die Gerade $g = x_0 + [x_1]$ und die Ebene

$L = y_0 + [x_2,x_3]$ gegeben mit

$$x_0 = \begin{bmatrix} 2 \\ -1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \; y_0 = \begin{bmatrix} 3 \\ 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \; x_1 = \begin{bmatrix} 1 \\ -2 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \; x_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \\ 2 \end{bmatrix}, \; x_3 = \begin{bmatrix} 0 \\ 2 \\ 0 \\ -1 \\ 1 \end{bmatrix}.$$

Wir wollen den Abstand $d(g,L)$ berechnen sowie Punkte $x \in g$ und $y \in L$ bestimmen mit $d(g,L) = d(x,y) = \| x - y \|$.

Die Verbindungsgerade von x und y muß notwendigerweise orthogonal zu g und zu L sein, also muß $x - y \in [x_1,x_2,x_3]^{\perp}$ sein. Für die gesuchten Punkte x und y gilt $x = x_0 + a_1x_1$ und $y = y_0 + a_2x_2 + a_3x_3$ mit $a_1,a_2,a_3 \in \mathbb{R}$. Die Bedingung $x - y \in [x_1,x_2,x_3]^{\perp}$ ist dann äquivalent zu

$$<x_0 - y_0 + a_1x_1 - a_2x_2 - a_3x_3 \, , x_i> \; = \; 0. \qquad (i = 1,2,3).$$

Dies ist ein inhomogenes LGS mit den Unbekannten a_1,a_2,a_3 und der erweiterten Matrix

$$\left[\begin{array}{rrr|r} 6 & -1 & 4 & -2 \\ 1 & -6 & -3 & 1 \\ -4 & -3 & -6 & 2 \end{array}\right]$$

Als Lösung erhalten wir $a_1 = -\frac{1}{2}$, $a_2 = -\frac{1}{3}$, $a_3 = \frac{1}{6}$. Damit ergibt sich

$$x = \frac{1}{2}\begin{bmatrix} 3 \\ 0 \\ 1 \\ -2 \\ 0 \end{bmatrix}, \; y = \frac{1}{6}\begin{bmatrix} 16 \\ 2 \\ 6 \\ -5 \\ -3 \end{bmatrix},$$

also $d(g,L) = \sqrt{2}$.

Besonders einfach läßt sich der Abstand eines Punktes X von einer Hyperebene L berechnen, wenn diese in Hessescher Normalform gegeben ist. Darunter verstehen wir folgendes:

Wir wählen einen beliebigen Ursprung O in $\mathbb{A}$ und in L einen beliebigen Punkt P. Der Orthogonalraum zu dem Richtungsraum U von L ist eindimensional, also gibt es in $U^\perp$ genau zwei Vektoren der Länge 1. Wir wählen einen davon aus, $y \in U^\perp$, $\| y \| = 1$, und erhalten

$$L = \{X \in \mathbb{A} \mid \overrightarrow{OX} = \overrightarrow{OP} + u \,,\, u \in U\} = \{X \in \mathbb{A} \mid <\overrightarrow{OX},y> = <\overrightarrow{OP},y> =: a\} .$$

Die letztere Darstellung von L heißt *Hessesche Normalform* der Hyperebene L, und y heißt *Normalenvektor* von L.

Ist nun $X \in \mathbb{A}$ ein beliebiger Punkt, so ist der Vektor $\overrightarrow{PX} - \pi_U(\overrightarrow{PX})$ ein Vielfaches des Normalenvektors y von L,

$$\overrightarrow{PX} - \pi_U(\overrightarrow{PX}) = b\, y \,,\, b \in \mathbb{R} .$$

Daraus folgt $<\overrightarrow{PX},y> = b$ und wegen $<\overrightarrow{OX},y> = <\overrightarrow{OP},y> + <\overrightarrow{PX},y> = a + b$ erhalten wir schließlich für den Abstand des Punktes X zur Hyperebene L:

$$d(X,L) = |b| = |<\overrightarrow{OX},y> - a| .$$

Beispiel. Im euklidischen Raum $\mathbb{R}^4$ sei die Hyperebene L durch die Gleichung

$$2\, x_1 + x_3 + 2\, x_4 = 6$$

mit den Variablen $x_1,\ldots,x_4$ gegeben. Wir wollen den Abstand des Punktes

$$x = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 0 \end{bmatrix}$$

von L bestimmen und stellen dazu L in der Hesseschen Normalform dar:

$$2\, x_1 + x_3 + 2\, x_4 = 6 \iff < \begin{bmatrix} 2 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} > = 6 .$$

Somit besitzt L die Hessesche Normalform $L = \{x' \mid <x',y> = a\}$ mit

$$y = \frac{1}{3}\begin{bmatrix} 2 \\ 0 \\ 1 \\ 2 \end{bmatrix} \ , \quad a = 2 \ .$$

Daraus folgt für den gesuchen Abstand: $d\,(x,L) = \mid <x,y> - a \mid = \frac{5}{3}$.

Speziell in einem dreidimensionalen euklidischen Raum gibt es noch eine einfache Methode um den Abstand zweier windschiefer Geraden zu bestimmen, bei der das sogenannte *vektorielle Produkt* eine Rolle spielt. Dieses spezielle Produkt ist auch für andere Anwendungen, etwa in der Physik, nützlich. Wir geben deshalb zum Abschluß dieses Paragraphen eine knappe Einführung in das Vektorprodukt.

Das Vektorprodukt

In einem dreidimensionalen euklidischen Raum $\mathbb{A}$ mit zugehörigem Vektorraum V besitzt ein Parallelogramm $OPQR$ (mit $\overrightarrow{OP} = \overrightarrow{RQ}$ und $\overrightarrow{OR} = \overrightarrow{PQ}$) den Flächeninhalt $\| \overrightarrow{OP} \| \cdot \| \overrightarrow{OR} \| \cdot |\sin \omega|$, wobei ω der Winkel zwischen $\overrightarrow{OP}$ und $\overrightarrow{OR}$ ist. Wir wollen die Parallelogrammfläche durch einen Vektor beschreiben, der zu $\overrightarrow{OP}$ und $\overrightarrow{OR}$ orthogonal ist und dessen Länge mit dem Flächeninhalt übereinstimmt.

Seien nun x und y beliebige Vektoren aus V und (e_1,e_2,e_3) eine feste ONB von V. Dann wird durch

$$z \longmapsto \Delta\,(x,y,z) \ := \ \det(\hat{x} \mid \hat{y} \mid \hat{z})$$

auf V eine Linearform definiert. Nach Satz 5.11 gibt es einen eindeutig bestimmten Vektor $z_0 \in V$ mit

$$(*) \qquad \Delta\,(x,y,z) \ = \ <z_0,z>$$

für alle $z \in V$. Dieser Vektor z_0 heißt *Vektorprodukt* (*Kreuzprodukt, äußeres Produkt*) der Vektoren x und y ; wir schreiben dafür $z_0 = x \times y$.

Bemerkung. Das Vektorprodukt ist eine Abbildung $\times : V \times V \longrightarrow V$ mit folgenden Eigenschaften:

(a) $x \times y = -y \times x$,

(b) $(x + x') \times y = x \times y + x' \times y$,

(c) $(a\,x) \times y = a\,(x \times y)$,

(d) $<x \times y, x> = <x \times y, y> = 0$,

(e) x, y linear abhängig $\iff x \times y = o$,

(f) $e_1 \times e_2 = e_3$, $e_2 \times e_3 = e_1$, $e_3 \times e_1 = e_2$.

(g) Seien $x = x_1\,e_1 + x_2\,e_2 + x_3\,e_3$ und $y = y_1\,e_1 + y_2\,e_2 + y_3\,e_3$. Dann gilt:

$$x \times y = (x_2\,y_3 - x_3\,y_2)\,e_1 + (x_3\,y_1 - x_1\,y_3)\,e_2 + (x_1\,y_2 - x_2\,y_1)\,e_3 .$$

Merkschema: Man "entwickle" folgenden Ausdruck formal wie eine Determinante nach der ersten Zeile:

$$\begin{vmatrix} e_1 & e_2 & e_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix}$$

(h) $\| x \times y \| = \sqrt{(\| x \|^2 \; \| y \|^2 - <x,y>^2)} = \| x \| \cdot \| y \| \cdot |\sin \omega|$,

wo ω der Winkel zwischen x und y ist ($x \neq o$, $y \neq o$).

Die Beweise überlassen wir als Übungsaufgabe.

Der Flächeninhalt des Parallelogramms $OPQR$ ist nach (h) also $\| \overrightarrow{OP} \times \overrightarrow{OR} \|$.

Nun wollen wir noch, wie angekündigt, mit Hilfe des Vektorproduktes den Abstand zweier windschiefer Geraden berechnen:

Es seien g und h windschiefe Geraden. Ihre Richtungen seien $U_g = [x]$ und $U_h = [y]$. Dann ist $x \times y$ der Richtungsvektor des gemeinsamen Lotes, und für die Lotfußpunkte $P_1 \in g$ und $Q_1 \in h$ gilt $\overrightarrow{P_1Q_1} = a\,(x \times y)$, $a \in \mathbb{R}$. Wir wollen a bestimmen. Hierzu wählen wir beliebige Punkte $P \in g$ und $Q \in h$ und erhalten

$$\overrightarrow{P_1Q_1} = \overrightarrow{P_1P} + \overrightarrow{PQ} + \overrightarrow{QQ_1} .$$

Daraus folgt

$$a \| x \times y \|^2 = <\overrightarrow{P_1 Q_1}, x \times y> = <\overrightarrow{PQ}, x \times y> .$$

Also ist

$$d(g,h) = \| \overrightarrow{P_1 Q_1} \| = | a | \cdot \| x \times y \| = \frac{|<\overrightarrow{PQ}, x \times y>|}{\| x \times y \|} .$$

Beispiel. Im $\mathbb{R}^3$ seien die windschiefen Geraden

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

gegeben, ihr Abstand soll berechnet werden. In diesem Fall ist

$$\overrightarrow{PQ} = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}$$

und

$$x \times y = (2 \cdot 3 - 1 \cdot 0)\, e_1 + (1 \cdot 2 - (-1) \cdot 3)\, e_2 + ((-1) \cdot 0 - 2 \cdot 2)\, e_3 = \begin{bmatrix} 6 \\ 5 \\ -4 \end{bmatrix} .$$

Also folgt

$$\| x \times y \| = \sqrt{77} \ , \ < \overrightarrow{PQ}, x \times y> = 16 ,$$

und wir erhalten $d(g,h) = \dfrac{16}{\sqrt{77}}$.

§ 2 Affine Abbildungen und Bewegungen

Nun wollen wir auch affine Abbildungen in allgemeinen affinen Räumen betrachten.

Definition. Es seien $\mathbb{A}$ und $\mathbb{B}$ affine Räume mit zugehörigen $\mathbb{K}$–Vektorräumen V und W. Dann heißt $\varphi : \mathbb{A} \longrightarrow \mathbb{B}$ eine *affine Abbildung*, falls eine lineare Abbildung Φ aus Hom (V, W) existiert mit

$$\Phi(\overrightarrow{PQ}) = \overrightarrow{\varphi(P)\varphi(Q)}$$

für alle Punkte $P, Q \in \mathbb{A}$. Jede bijektive affine Selbstabbildung $\varphi : \mathbb{A} \longrightarrow \mathbb{A}$ heißt *Affinität.*

Bemerkung. Nach Auszeichnung eines Ursprungs $O \in \mathbb{A}$ und eines Ursprungs $O' \in \mathbb{B}$ gilt für alle Punkte $X \in \mathbb{A}$:

$$\overrightarrow{O'\varphi(X)} = \overrightarrow{\varphi(O)\varphi(X)} + \overrightarrow{O'\varphi(O)} = \Phi(\overrightarrow{OX}) + w$$

mit $w = \overrightarrow{O'\varphi(O)} \in W$. Damit genügt es meist, den Fall $\mathbb{A} = V$, $\mathbb{B} = W$ und

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

$\Phi \in \mathrm{Hom}(V, W)$, $w \in W$, zu betrachten (vgl. § 3.4).

w heißt *Translationsvektor* zur affinen Abbildung φ. Er ist ebenso wie die lineare Abbildung Φ durch φ eindeutig bestimmt.

Beispiele affiner Abbildungen haben wir schon in § 3.4 kennengelernt:

Die Translationen sind diejenigen Abbildungen $\tau : \mathbb{A} \longrightarrow \mathbb{A}$, für die $\overrightarrow{\tau(P)\tau(Q)} = \overrightarrow{PQ}$ für alle $P, Q \in \mathbb{A}$ gilt. Die zugehörige lineare Abbildung ist die Identität.

Die Streckungen mit Zentrum P sind diejenigen Abbildungen $\delta : \mathbb{A} \longrightarrow \mathbb{A}$, für die es ein $c \in \mathbb{K}$, $c \neq 0$, gibt mit $\overrightarrow{P\delta(Q)} = c\,\overrightarrow{PQ}$ für alle $Q \in \mathbb{A}$. Die zugehörigen linearen Abbildungen sind also von der Form $\Phi = c\,\mathrm{id}_V$, $c \neq 0$. Für char $\mathbb{K} \neq 2$ und

$c = -1$ ist δ die Punktspiegelung an P.

Translationen und Streckungen haben die Eigenschaft, daß sie jede Gerade auf eine dazu parallele Gerade abbilden. Solche Affinitäten nennt man auch *Homothetien*. Umgekehrt läßt sich zeigen, daß jede Homothetie, die von der Identität verschieden ist, entweder eine Translation oder eine Streckung ist (Beweis als Übungsaufgabe).

Affine Abbildungen sind durch ein affines Koordinatensystem eindeutig festgelegt.

Satz 6. *Es seien* $\mathbb{A},\mathbb{B}$ *affine Räume über demselben Körper und* $\dim \mathbb{A} = n$. *Weiter seien* $P_0,\ldots,P_n$ *affin unabhängige Punkte in* $\mathbb{A}$ *sowie* $Q_0,\ldots,Q_n$ *beliebige Punkte in* $\mathbb{B}$. *Dann gibt es genau eine affine Abbildung* $\varphi : \mathbb{A} \longrightarrow \mathbb{B}$ *mit* $\varphi(P_i) = Q_i$ *für* $i = 0,\ldots,n$.

Beweis. Seien V bzw. W die zu $\mathbb{A}$ bzw. $\mathbb{B}$ gehörenden Vektorräume. Wegen der affinen Unabhängigkeit der Punkte $P_0,\ldots,P_n$ sind die Vektoren $\overrightarrow{P_0P_i} \in V$, $i = 1,\ldots,n$ linear unabhängig, bilden also eine Basis von V. Nach Satz 3.4 gibt es genau eine lineare Abbildung $\Phi : V \longrightarrow W$ mit $\Phi(\overrightarrow{P_0P_i}) = \overrightarrow{Q_0Q_i}$, $i = 1,\ldots,n$. Die durch $X \longmapsto \varphi(X)$ mit $\overrightarrow{Q_0\varphi(X)} = \Phi(\overrightarrow{P_0X})$, $X \in \mathbb{A}$, erklärte Abbildung $\varphi : \mathbb{A} \longrightarrow \mathbb{B}$ ist affin und bildet P_i auf Q_i ab für $i = 0,\ldots,n$. Die Abbildung φ ist eindeutig bestimmt. Aus $\varphi(P_i) = \varphi'(P_i)$ folgt nämlich $\Phi(\overrightarrow{P_0P_i}) = \Phi'(\overrightarrow{P_0P_i})$ für $i = 1,\ldots,n$. Daraus erhalten wir $\Phi = \Phi'$ und somit $\overrightarrow{Q_0\varphi(X)} = \Phi(\overrightarrow{P_0X}) = \Phi'(\overrightarrow{P_0X}) = \overrightarrow{Q_0\varphi'(X)}$ für alle $X \in \mathbb{A}$, also ergibt sich $\varphi = \varphi'$. ∎

Beispiel. Im $\mathbb{R}^3$ seien die Punkte $x_0,\ldots,x_3$ gegeben und im $\mathbb{R}^4$ die Punkte $y_0,\ldots,y_3$. Dabei seien

$$x_0 = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}, x_1 = \begin{bmatrix} 4 \\ 3 \\ 3 \end{bmatrix}, x_2 = \begin{bmatrix} 3 \\ 2 \\ 3 \end{bmatrix}, x_3 = \begin{bmatrix} 0 \\ 3 \\ 4 \end{bmatrix},$$

$$y_0 = \begin{bmatrix} 2 \\ 2 \\ 3 \\ 1 \end{bmatrix}, y_1 = \begin{bmatrix} 5 \\ 4 \\ 4 \\ 5 \end{bmatrix}, y_2 = \begin{bmatrix} 4 \\ 3 \\ 4 \\ 4 \end{bmatrix}, y_3 = \begin{bmatrix} 1 \\ 4 \\ 0 \\ -3 \end{bmatrix}.$$

Wir suchen eine affine Abbildung $\varphi : \mathbb{R}^3 \longrightarrow \mathbb{R}^4$ mit $\varphi(x_i) = y_i$, $i = 0,\ldots,3$. Zunächst stellen wir fest, daß die Punkte $x_0,\ldots,x_3$ affin unabhängig sind, da die Matrix

$$S = (x_1 - x_0 \mid x_2 - x_0 \mid x_3 - x_0) = \begin{bmatrix} 3 & 2 & -1 \\ 2 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$

den Rang 3 hat.

Die gesuchte affine Abbildung hat die Form $x \longmapsto A\,x + w$ mit $A \in \mathbb{R}^{4\times 3}$ und $w \in \mathbb{R}^4$. Es gilt

$$\varphi(x_i) - \varphi(x_0) = A\,(x_i - x_0)\ ,\ i = 1,2,3 \iff y_i - y_0 = A\,(x_i - x_0)\ ,\ i = 1,2,3$$

$$\iff (y_1 - y_0 \mid y_2 - y_0 \mid y_3 - y_0) = A\,S \iff A = (y_1 - y_0 \mid y_2 - y_0 \mid y_3 - y_0) \cdot S^{-1}$$

$$= \begin{bmatrix} 3 & 2 & -1 \\ 2 & 1 & 2 \\ 1 & 1 & -3 \\ 4 & 3 & -4 \end{bmatrix} \begin{bmatrix} -1 & 2 & -5 \\ 2 & -3 & 8 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & -1 & 0 \\ 2 & -1 & 0 \end{bmatrix}$$

Für den Translationsvektor w gilt $w = y_0 - A\,x_0$, also

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

Aus Satz 6 folgt speziell, daß in $\mathbb{A}$ jeder Wechsel eines affinen Koordinatensystems durch eine Affinität beschrieben wird.

Umgekehrt bildet jede Affinität $\varphi : \mathbb{A} \longrightarrow \mathbb{A}$ mit zugehörigem Vektorraumisomorphismus Φ ein Koordinatensystem $(O;\, v_1,\ldots,v_n)$ auf ein neues Koordinatensystem $(\varphi(O);\, \Phi(v_1),\ldots,\Phi(v_n))$ ab. Für jeden Punkt $X \in \mathbb{A}$ stimmen dann die affinen Koordinaten von X im alten und die von $\varphi(X)$ im neuen Koordinatensystem überein. Aus

$$\overrightarrow{OX} = x_1\,v_1 + \ldots + x_n\,v_n$$

folgt nämlich

$$\overrightarrow{\varphi(O)\varphi(X)} = \Phi(\overrightarrow{OX}) = x_1\, \Phi(v_1) + \ldots + x_n\, \Phi(v_n).$$

Insbesondere stimmen also die affinen Koordinaten von X im neuen System mit denen von $\varphi^{-1}(X)$ im alten System überein.

Affine Abbildungen lassen die geometrische Struktur des affinen Raumes invariant, denn sie bilden affine Unterräume auf affine Unterräume ab und erhalten die Parallelität (vgl. § 3.4). Bijektive affine Abbildungen erhalten darüber hinaus auch die Dimension. Sie bilden also Geraden auf Geraden ab, insbesonderere parallele Geraden auf ebensolche. Man spricht dann von der *Geradentreue* und der *Parallelentreue* dieser Abbildungen.

Umgekehrt reichen diese Eigenschaften aber nicht aus, um die bijektiven affinen Abbildungen zu charakterisieren.

Beispiel. In der affinen Ebene $\mathbb{C}^2$ ist die Abbildung

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \longmapsto \begin{bmatrix} \overline{z}_1 \\ \overline{z}_2 \end{bmatrix}$$

bijektiv, geradentreu und parallelentreu, aber nicht affin.

Zur geometrischen Beschreibung affiner Abbildungen benötigen wir noch einen weiteren wichtigen Begriff der affinen Geometrie, das Teilverhältnis.

Definition. Sind P, Q, R drei kollineare Punkte (d.h. Punkte, die auf einer Geraden liegen) und ist $P \neq Q$, so heißt die Zahl $c \in \mathbb{K}$ mit $\overrightarrow{PR} = c\, \overrightarrow{PQ}$ das *Teilverhältnis* von R bezüglich P und Q, Schreibweise: $\mathrm{TV}\,(P,Q\,;R) = c$.

Bemerkungen. (a) Für $c = 0$ ist $R = P$, für $c = 1$ ist $R = Q$ und für $\operatorname{char} \mathbb{K} \neq 2$ und $c = 1/2$ ist R der *Mittelpunkt* von P und Q.

(b) Affine Abbildungen $\varphi : \mathbb{A} \longrightarrow \mathbb{B}$ lassen das Teilverhältnis invariant, d.h. es gilt $\mathrm{TV}\,(P,Q;\,R) = \mathrm{TV}\,(\varphi(P),\varphi(Q);\varphi(R))$, falls $\varphi(P) \neq \varphi(Q)$ ist.

Satz 7. *Jede bijektive Abbildung $\varphi : \mathbb{A} \longrightarrow \mathbb{A}$, die geradentreu, parallelentreu und teilverhältnistreu ist, ist eine Affinität.*

Beweis. Wir betrachten zunächst den Fall, daß $\dim \mathbb{A} = 1$ gilt, $\mathbb{A}$ also eine Gerade ist. Wir wählen in $\mathbb{A}$ zwei verschiedene Punkte O und P. Dann sind auch $\varphi(O)$ und $\varphi(P)$ verschieden und es gibt genau eine Affinität $\alpha : \mathbb{A} \longrightarrow \mathbb{A}$ mit $\alpha(\varphi(O)) = O$ und $\alpha(\varphi(P)) = P$. Aus der Teilverhältnistreue von $\alpha \circ \varphi$ folgt nun sofort $\alpha \circ \varphi = \mathrm{id}_{\mathbb{A}}$, also ist $\varphi = \alpha^{-1}$ eine Affinität.

Sei nun $\dim \mathbb{A} \geq 2$. Da φ bijektiv, geradentreu und parallelentreu ist, bildet φ nichtausgeartete Parallelogramme auf ebensolche ab. Für alle $P,Q,R,S \in \mathbb{A}$ folgt somit aus $\overrightarrow{PQ} = \overrightarrow{RS}$ auch $\overrightarrow{\varphi(P)\varphi(Q)} = \overrightarrow{\varphi(R)\varphi(S)}$. Setzen wir also für beliebiges $x = \overrightarrow{PQ} \in V$ als Bildvektor $\Phi(x) := \overrightarrow{\varphi(P)\varphi(Q)}$, so wird dadurch eine bijektive Abbildung $\Phi : V \longrightarrow V$ definiert. Φ ist linear:

Die Homogenität von Φ ergibt sich unmittelbar aus der Teilverhältnistreue von φ: Seien $x = \overrightarrow{PQ} \in V$ und $c \in \mathbb{K}$ beliebig. Dann gilt mit $\overrightarrow{PR} = c\,\overrightarrow{PQ}$

$$\Phi(c\,x) = \Phi(c\,\overrightarrow{PQ}) = \Phi(\overrightarrow{PR}) = \overrightarrow{\varphi(P)\varphi(R)} = c\,\overrightarrow{\varphi(P)\varphi(Q)} = c\,\Phi(\overrightarrow{PQ}) = c\,\Phi(x)\,.$$

Die Additivität von Φ folgt für linear abhängige $x,y \in V$ aus der Homogenität von Φ, für linear unabhängige $x,y \in V$ folgt sie aus der Eigenschaft von φ, echte Parallelogramme in ebensolche abzubilden. ■

Die Voraussetzungen von Satz 7 lassen sich abschwächen. Wir verweisen hierzu auf die Literatur (etwa: H. Schaal: Lineare Algebra und Analytische Geometrie, Teil I).

Die Affinitäten von $\mathbb{A}$ bilden eine Gruppe. Mit ihrer Hilfe lassen sich die Teilmengen von $\mathbb{A}$ so in Klassen einteilen, daß sich Elemente aus derselben Klasse bezüglich ihrer affingeometrischen Eigenschaften nicht mehr unterscheiden lassen.

Wir nennen M_1, $M_2 \subset \mathbb{A}$ *affin-äquivalent*, wenn es eine Affinität φ von $\mathbb{A}$ gibt mit $\varphi(M_1) = M_2$.

Dieses Einteilungsprinzip wird bei der affinen Klassifikation der Quadriken eine wichtige Rolle spielen.

Zum Schluß betrachten wir nun auch noch affine Abbildungen in euklidischen

Räumen.

Definition. Eine affine Abbildung $\varphi : \mathbb{A} \longrightarrow \mathbb{B}$ zweier euklidischer Räume $\mathbb{A}$ und $\mathbb{B}$ heißt *Isometrie*, falls die zugehörige lineare Abbildung Φ eine Isometrie ist.

Für $\mathbb{A} = \mathbb{B}$ heißt eine Isometrie φ auch *Bewegung*, und zwar *eigentliche Bewegung*, wenn $\det \Phi = 1$ und *uneigentliche Bewegung*, wenn $\det \Phi = -1$ ist.

Isometrien lassen den Abstand unverändert, denn für alle $X, Y \in \mathbb{A}$ gilt

$$d(\varphi(X),\varphi(Y)) = \| \overrightarrow{\varphi(X)\varphi(Y)} \| = \| \Phi(\overrightarrow{XY}) \| = \| \overrightarrow{XY} \| = d(X,Y).$$

Diese Eigenschaft genügt nun auch schon, die Isometrien zu charakterisieren.

Satz 8. *Es seien $\mathbb{A}$ und $\mathbb{B}$ euklidische Räume und $\varphi : \mathbb{A} \longrightarrow \mathbb{B}$ eine Abbildung. Genau dann ist φ eine Isometrie, wenn*

$$d(\varphi(X),\varphi(Y)) = d(X,Y)$$

für alle Punkte $X, Y \in \mathbb{A}$ gilt.

Beweis. Zu zeigen ist nur noch die eine Richtung. Seien o.B.d.A. $\mathbb{A} = V$ und $\mathbb{B} = W$, und gelte $d(\varphi(x),\varphi(y)) = d(x,y)$ für alle $x,y \in V$. Wir betrachten die Abbildung $\Phi : V \longrightarrow W$, die durch

$$\Phi(x) := \varphi(x) - \varphi(o) \quad , x \in V,$$

erklärt ist. Es gilt dann:

$$\text{(i)} \qquad \| \Phi(x) - \Phi(y) \| = \| x - y \| \quad \text{für alle } x,y \in V,$$

insbesondere wegen $\Phi(o) = o$

$$\text{(ii)} \qquad \| \Phi(x) \| = \| x \| \quad \text{für alle } x \in V.$$

Daraus folgt wegen

$$\| x - y \|^2 - \| x \|^2 - \| y \|^2 = -2 <x,y>$$

und

$$\| \Phi(x) - \Phi(y) \|^2 - \| \Phi(x) \|^2 - \| \Phi(y) \|^2 = -2 <\Phi(x),\Phi(y)>$$

schließlich

(iii) $$<\Phi(x),\Phi(y)> \;=\; <x,y> \text{ für alle } x,y \in V.$$

Weiter gilt:

(iv) $$\Phi(x+y) \;=\; \Phi(x) + \Phi(y) \text{ für alle } x,y \in V,$$

denn $\| \Phi(x+y) - \Phi(x) - \Phi(y) \|^2 \;=\; \| \Phi(x+y) \|^2 + \| \Phi(x) \|^2 + \| \Phi(y) \|^2 - 2 <\Phi(x+y),\Phi(x)> - 2 <\Phi(x+y),\Phi(y)> + 2 <\Phi(x),\Phi(y)> \;=\; \| x+y \|^2 + \| x \|^2 + \| y \|^2 - 2 <x+y,x> - 2 <x+y,y> + 2 <x,y> \;=\; \| (x+y) - x - y \| = 0$.

(v) $$\Phi(a\,x) = a\,\Phi(x) \text{ für alle } x \in V,\ a \in \mathbb{R},$$

denn $\| \Phi(a\,x) - a\,\Phi(x) \|^2 \;=\; \| \Phi(a\,x) \|^2 + a^2 \| \Phi(x) \|^2 - 2\,a <\Phi(a\,x),\Phi(x)> = \| a\,x \|^2 + a^2 \| x \|^2 - 2\,a <a\,x,x> \;=\; 2\,a^2 \| x \|^2 - 2\,a^2 \| x \|^2 \;=\; 0.$

Aus (iv) und (v) folgt nun, daß Φ linear ist. Wegen (iii) ist Φ sogar Isometrie. Also ist auch die Abbildung $\varphi : \mathbb{A} \longrightarrow \mathbb{B}$ wegen $\varphi(x) = \Phi(x) + w$ mit $w = \varphi(o)$ eine Isometrie. ■

Die Bewegungen eines euklidischen Raumes $\mathbb{A}$ bilden ebenfalls eine Gruppe. Sie ist eine Untergruppe der Gruppe der Affinitäten, weshalb sich mit ihrer Hilfe die Einteilung der Teilmengen von $\mathbb{A}$ in affine Äquivalenzklassen weiter verfeinern läßt.

Wir nennen die Teilmengen M_1, $M_2 \subset \mathbb{A}$ *kongruent*, wenn es eine Bewegung φ von $\mathbb{A}$ gibt mit $\varphi(M_1) = M_2$.

Dieses Einteilungsprinzip wird in § 6.4 bei der euklidischen Klassifikation der Quadriken angewendet werden.

§ 3 Quadriken in affinen Räumen

Bisher haben wir in affinen Räumen als Punktmengen hauptsächlich die affinen Unterräume untersucht. Sie lassen sich bei gegebenem Koordinatensystem durch lineare Gleichungssysteme beschreiben. Nun wollen wir einen Schritt weitergehen und Punktmengen betrachten, die sich durch die einfachsten nichtlinearen Gleichungen, also durch quadratische Gleichungen beschreiben lassen; dies sind die Quadriken.

Zu den Quadriken zählen die elementaren geometrischen Gebilde, die man vielleicht noch von der Schule her kennt, wie Kreis, Ellipse, Hyperbel, Kugel, Ellipsoid, Paraboloid usw. Auf solche Gebilde stößt man vielfach, wenn man Abstandsaufgaben in der Ebene oder im Raum lösen will. Dadurch spielen Quadriken auch bei vielen technischen Konstruktionen eine Rolle. Quadriken erhält man auch, wenn differenzierbare Flächen lokal approximiert werden sollen. In erster Näherung ergibt sich die Tangentialhyperebene, in zweiter Näherung eine Quadrik, die die Krümmung beschreibt.

Quadriken können in beliebigen affinen Räumen definiert werden, im Hinblick auf die obengenannten Anwendungen beschränken wir uns aber hier auf reelle affine Räume und beginnen mit dem $\mathbb{R}^n$.

Die allgemeinste Form einer quadratischen Gleichung in $\mathbb{R}^n$ ist

$$\sum_{i,\,j=1}^{n} a_{ij} x_i x_j + 2 \sum_{j=1}^{n} b_j x_j + c = 0$$

mit reellen Variablen x_j und reellen Koeffizienten a_{ij}, b_j, c, wobei mindestens einer der Ausdrücke $a_{ij} + a_{ji}$ von Null verschieden ist. Setzen wir

$$\tilde{a}_{ij} := (a_{ij} + a_{ji})/\,2 \ , \ i,j = 1,\dots,n \,,$$

und

$$b = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}, \ x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix},$$

so läßt sich die obige Gleichung auch in der Form

$$(*) \qquad x^{\mathsf{T}} A\, x \;+\; 2\, b^{\mathsf{T}}\, x \;+\; c \;=\; 0$$

schreiben, wobei $A = (\!(\tilde{a}_{ij})\!)$ eine symmetrische Matrix ist und $A \neq O$ gilt.

Die allgemeine quadratische Gleichung im $\mathbb{R}^n$ besteht also aus einem rein-quadratischen Anteil, gegeben durch die von der Nullabbildung verschiedene symmetrische Bilinearform $\beta : (x,y) \longmapsto x^{\mathsf{T}} A\, y$, aus einem linearen Anteil, gegeben durch die Linearform $\Phi : x \longmapsto b^{\mathsf{T}}\, x$, und aus einer Konstanten $c \in \mathbb{R}$. Damit können wir die Gleichung (*) auch so schreiben:

$$(**) \qquad \beta(x,x) \;+\; 2\,\Phi(x) \;+\; c \;=\; 0\,.$$

Interpretieren wir die Lösungen von (*) bzw. (**) als Punkte des affinen Standardraumes $\mathbb{R}^n$, so haben wir stillschweigend das affine Standardkoordinatensystem $(o;e_1,\dots,e_n)$ ausgezeichnet und fassen $x_1,\dots,x_n$ als Koordinaten des Punktes x bezüglich dieses Koordinatensystems auf. Ist die so beschriebene Punktmenge nicht leer, so sprechen wir von einer Quadrik in $\mathbb{R}^n$.

Im allgemeinen Fall eines beliebigen n–dimensionalen reellen affinen Raumes $\mathbb{A}$ mit zugehörigem Vektorraum V legen wir entsprechend eine Gleichung der Form (**) zugrunde und fassen jede Lösung $x \in V$ als Ortsvektor $x = \overrightarrow{OX}$ eines Punktes $X \in \mathbb{A}$ bezüglich eines ausgezeichneten Ursprungs $O \in \mathbb{A}$ auf. Damit wird durch (**) eine Punktmenge in $\mathbb{A}$ beschrieben, die wir dann ebenfalls eine Quadrik nennen.

Definition. Es seien $\mathbb{A}$ ein reeller n–dimensionaler affiner Raum mit zugehörigem Vektorraum V, $\beta : V \times V \longrightarrow \mathbb{R}$ eine von der Nullabbildung verschiedene symmetrische Bilinearform, $\Phi : V \longrightarrow \mathbb{R}$ eine Linearform und $c \in \mathbb{R}$ eine Konstante. Ferner sei ein Punkt $O \in \mathbb{A}$ als Ursprung ausgezeichnet. Ist die Punktmenge

$$Q \;=\; \{X \in \mathbb{A} \mid \overrightarrow{OX} = x\,,\, \beta(x,x) \;+\; 2\,\Phi(x) \;+\; c \;=0\,\}$$

nicht leer, so heißt sie eine *Quadrik.* In der affinen Ebene heißt eine Quadrik auch *Kegelschnitt.* Ist die affine Hülle von Q der gesamte Raum $\mathbb{A}$, so heißt Q eine *eigentliche Quadrik.*

Wir werden uns hauptsächlich mit den eigentlichen Quadriken beschäftigen. Quadriken, die nicht eigentlich sind, liegen in einer Hyperebene von $\mathbb{A}$, es wird sich sogar herausstellen, daß sie selbst affine Unterräume von $\mathbb{A}$ sind, also durch lineare Gleichungen beschrieben werden können.

Bemerkung. Die definierende Gleichung eine Quadrik Q ist nicht eindeutig bestimmt. Multiplizieren wir sie etwa mit $a \neq 0$, so erhalten wir dieselbe Punktmenge.

Zeichnen wir in V eine Basis $(v_1,\dots,v_n)$ aus, so erfüllen die Koordinaten der Punkte von Q bezüglich des affinen Koordinatensystems $(O;\ v_1,\dots,v_n)$ eine Gleichung der Form (*) mit $A = (\!(\beta(v_i,v_j))\!)$ und $b^{\mathsf{T}} = (\Phi(v_1)\cdots\Phi(v_n))$. Diesen Zusammenhang werden wir im folgenden häufig ausnutzen.

Das weitere Programm ist nun vorgezeichnet. Wir suchen alle Äquivalenzklassen affin–äquivalenter Quadriken in $\mathbb{A}$ und in jeder Äquivalenzklasse einen Repräsentanten, der durch eine möglichst einfache quadratische Gleichung beschrieben wird. Diese nennen wir dann die affine Normalform von Q.

Wir beginnen mit dem Nachweis, daß Quadriken unter bijektiven affinen Abbildungen in Quadriken übergehen.

Satz 9. *Es seien $Q \subset \mathbb{A}$ eine Quadrik und $\varphi : \mathbb{A} \longrightarrow \mathbb{A}$ eine Affinität. Dann ist auch $\varphi\,(Q)$ eine Quadrik.*

Beweis. Q werde bezüglich des Ursprungs O durch die quadratische Gleichung

$$\beta(x,x) \;+\; 2\,\Phi(x) \;+\; c \;=0$$

beschrieben und für die Affinität φ gelte für alle $X \in \mathbb{A}$

$$\overrightarrow{O\varphi(X)} \;=\; \Psi(\overrightarrow{OX}) + \;v$$

mit einem Isomorphismus $\Psi \in \operatorname{Hom}(V,V)$ und $v \in V$. Dann gilt:

$$Y = \varphi(X) \in \varphi(Q) \iff X = \varphi^{-1}(Y) \in Q\,.$$

Durch Einsetzen von $x = \overrightarrow{OX} = \Psi^{-1}(\overrightarrow{OY}) - \Psi^{-1}(v)$ in die obige Gleichung erhalten wir für die Ortsvektoren $y = \overrightarrow{OY}$ der Punkte Y von $\varphi(Q)$

$$\beta(\Psi^{-1}(y) - \Psi^{-1}(v), \Psi^{-1}(y) - \Psi^{-1}(v)) + 2\,\Phi(\Psi^{-1}(y) - \Psi^{-1}(v)) + c = 0\,,$$

also

$$\beta(\Psi^{-1}(y), \Psi^{-1}(y)) + 2\,\{\Phi \circ \Psi^{-1}(y) - \beta(\Psi^{-1}(y), \Psi^{-1}(v))\}$$
$$+ \beta(\Psi^{-1}(v), \Psi^{-1}(v)) - 2\,\Phi \circ \Psi^{-1}(v) + c = 0\,.$$

Dies ist eine quadratische Gleichung, also ist $\varphi\,(Q)$ eine Quadrik. ■

Bemerkung. Ist $\hat{x}^{\mathsf{T}} A\, \hat{x} + 2\, b^{\mathsf{T}}\, \hat{x} + c = 0$ die Matrixdarstellung von Q bezüglich eines affinen Koordinatensystems $(O;\, v_1,\ldots,v_n)$ mit dem ausgezeichneten Punkt O als Ursprung und sind entsprechend

$$\hat{x} \longmapsto B\,\hat{x} + \hat{v}\ , \quad \hat{x} \in \mathbb{R}^n,$$

bzw.

$$\hat{y} \longmapsto B^{-1}\,\hat{y} - B^{-1}\,\hat{v}\ , \quad \hat{y} \in \mathbb{R}^n,$$

die Matrixdarstellungen von φ bzw. φ^{-1}, so gilt für die Gleichung von $\varphi\,(Q)$ bezüglich $(O;\, v_1,\ldots,v_n)$

(***) $\quad \hat{y}^{\mathsf{T}} C^{\mathsf{T}} A\, C\, \hat{y} + 2\,(b^{\mathsf{T}} C + \hat{u}^{\mathsf{T}} A\, C)\,\hat{y} + (\hat{u}^{\mathsf{T}} A\,\hat{u} + 2\, b^{\mathsf{T}}\,\hat{u} + c) = 0\,,$

wobei $C = B^{-1}$ und $\hat{u} = -B^{-1}\,\hat{v}$ ist.

Fassen wir die Affinität φ als Koordinatenwechsel auf, so wird die Quadrik Q bezüglich des neuen Koordinatensystems $(\varphi^{-1}(O);\, \Psi^{-1}(v_1),\ldots,\Psi^{-1}(v_n))$ durch die Gleichung (***) dargestellt (vgl. S.298 unten).

Ist φ insbesondere eine Translation, also $\Psi = \mathrm{id}_V$ und damit $B = C = E_n$, so hat (***) die Form

$$\hat{y}^{\mathsf{T}} A\,\hat{y} + 2\,(b^{\mathsf{T}} + \hat{u}^{\mathsf{T}} A)\,\hat{y} + (\hat{u}^{\mathsf{T}} A\,\hat{u} + 2\, b^{\mathsf{T}}\,\hat{u} + c) = 0\,,$$

d.h. bei einer Verschiebung des Urprungs ändert sich an dem quadratischen Teil der Gleichung von Q nichts, wohl aber an dem linearen Teil und an der Konstanten.

Bleibt der Ursprung fest, so bleibt die Konstante unverändert.

Unser Ziel wird nun sein, den Koordinatenwechsel so vorzunehmen, daß die Gleichung von Q eine möglichst einfache Gestalt annimmt, oder anders interpretiert, eine Affinität φ so zu wählen, daß die Gleichung von $\varphi(Q)$ möglichst einfach wird.

Satz 10 (Satz über die affine Hauptachsentransformation von Quadriken). *Es sei $\mathbb{A}$ ein n-dimensionaler reeller affiner Raum. Dann läßt sich jede Quadrik $Q \subset \mathbb{A}$ bezüglich eines geeigneten affinen Koordinatensystems durch eine der folgenden Gleichungen beschreiben:*

(I) $$x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_r^2 = 0 \;, \qquad 0 < p \leq r \leq n,\ r-p \leq p\,,$$

(II) $$x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_r^2 = 1 \;, \qquad 0 < p \leq r \leq n\,,$$

(III) $$x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_r^2 = 2\,x_n \;, \qquad 0 < p \leq r < n,\ r-p \leq p\,.$$

Beweis. Die Quadrik Q werde nach Auszeichnung eines Koordinatensystems $(O; v_1,\dots,v_n)$ durch die Gleichung

$$\hat{x}^{\mathsf{T}} A\, \hat{x} + 2\, b^{\mathsf{T}}\, \hat{x} + c = 0$$

dargestellt, wo $A = (\!(\beta(v_i,v_j))\!)$ eine symmetrische Matrix ist. Nach Satz 5.13 gibt es eine orthogonale Matrix S, so daß $A' = S^{\mathsf{T}} A\, S$ Diagonalgestalt hat. In der Diagonalen stehen dann die Eigenwerte $c_1,\dots,c_n$ von A.

Für den Vektorraum V bedeutet dies die Existenz einer Basis $(v_1',\dots,v_n')$, so daß die Matrix $A' = (\!(\beta(v_i',v_j'))\!)$ Diagonalgestalt hat mit $\beta(v_i',v_i') = c_i$, $i = 1,\dots,n$. Wegen $\beta \neq O$ ist mindestens einer der Eigenwerte c_i von Null verschieden. Sei o.B.d.A. $c_1,\dots,c_p > 0$, $c_{p+1},\dots,c_r < 0$ und $c_{r+1} = \dots = c_n = 0$. Es ist $r = \operatorname{Rg} A' = \operatorname{Rg} A > 0$, die Zahl r hängt also nicht von der Wahl der Basis ab, sondern nur von der Bilinearform β. Ersetzen wir die Basisvektoren v_i' durch neue Vektoren $\overline{v}_i$,

$$\overline{v}_i = \frac{1}{\sqrt{|c_i|}}\, v_i' \;, \quad i = 1,\dots,r\,,$$

$$\overline{v}_i = v_i' \;, \quad i = r+1,\dots,n\,,$$

so nimmt die Gleichung von Q bezüglich des neuen Koordinatensystems $(O; \overline{v}_1,\dots,\overline{v}_n)$ die Form

$$\overline{x}_1^2 + \dots + \overline{x}_p^2 - \overline{x}_{p+1}^2 - \dots - \overline{x}_r^2 + 2\,b_1\,\overline{x}_1 + \dots + 2\,b_n\,\overline{x}_n + c = 0$$

an mit neuen Konstanten $b_1,\dots,b_k$ und den (neuen) Variablen $\overline{x}_j$. Durch quadratische Ergänzung erhalten wir weiter

$$(\overline{x}_1 + b_1)^2 + \dots + (\overline{x}_p + b_p)^2 - (\overline{x}_{p+1} - b_{p+1})^2 - \dots - (\overline{x}_r - b_r)^2$$
$$+ 2\,b_{r+1}\,\overline{x}_{r+1} + \dots + 2\,b_n\,\overline{x}_n + c = 0,$$

wobei sich jetzt auch die Konstante c verändert haben kann. Nun verschieben wir den Ursprung O durch eine Translation, die bezüglich $(O; \overline{v}_1, \dots,\overline{v}_n)$ durch die folgenden Gleichungen definiert ist:

$$y_i = \overline{x}_i + b_i\,, \quad i = 1,\dots,p\,,$$
$$y_i = \overline{x}_i - b_i\,, \quad i = p+1,\dots,r\,,$$
$$y_i = \overline{x}_i\,, \quad i = r+1,\dots,n\,.$$

Wir erhalten

$$y_1^2 + \dots + y_p^2 - y_{p+1}^2 - \dots - y_r^2 + 2\,b_{r+1}\,y_{r+1} + \dots + 2\,b_n\,y_n + c = 0\,.$$

1. Fall: $b_{r+1} = \dots = b_n = c = 0$. Dann besitzt Q bezüglich des zuletzt betrachteten affinen Koordinatensystems die Gleichung

$$\text{(I)} \qquad y_1^2 + \dots + y_p^2 - y_{p+1}^2 - \dots - y_r^2 = 0\,.$$

Dabei gilt o.B.d.A. $r - p \leq p$, denn andernfalls multiplizieren wir die Gleichung mit -1.

2. Fall: $b_{r+1} = \dots = b_n = 0$, $c \neq 0$. Hier können wir entsprechend annehmen, daß $c < 0$ gilt. Wir dividieren durch $-c$ und nach einer einer weiteren Koordinatentransformation

$$z_i = (-c)^{-\frac{1}{2}}\,y_i\,, \quad i = 1,\dots,r\,,$$
$$z_i = y_i\,, \quad i = r+1,\dots,n\,,$$

erhalten wir die Gleichung

(II) $$z_1^2 + \cdots + z_p^2 - z_{p+1}^2 - \cdots - z_r^2 = 1 .$$

Wegen $Q \neq \emptyset$ ist dann $p > 0$.

3. Fall: Mindestens einer der Koeffizienten b_i , $i = r + 1,\ldots,n$, ist von 0 verschieden. Dann ist $r < n$. Wir können annehmen, daß $b_n \neq 0$ gilt, andernfalls numerieren wir um, was wiederum einem Wechsel des Koordinatensystems entspricht. Setzen wir jetzt

$$z_i = y_i \ , \ i = 1,\ldots,n-1 \ ,$$

$$z_n = -(b_{r+1}\, y_{r+1} + \ldots + b_n\, y_n) - \frac{c}{2} \ ,$$

so erhalten wir die Gleichung

(III) $$z_1^2 + \cdots + z_p^2 - z_{p+1}^2 - \cdots - z_r^2 = 2\, z_n$$

mit $0 < r < n$. Weiter können wir erreichen, daß $r - p \leq p$ gilt, indem wir gegebenenfalls die Gleichung mit -1 multiplizieren und

$$\bar{z}_i = z_i \ , \ i = 1,\ldots,n-1 \ ,$$

$$\bar{z}_n = -z_n$$

setzen. ∎

Bezeichnungen und Bemerkungen. **(a)** Die Gleichungen (I), (II), (III) heißen, im Vorgriff auf Satz 15, die *affinen Normalformen* einer Quadrik. Zwei Quadriken Q und $\tilde{Q}$, die bezüglich geeigneter Koordinatensysteme dieselbe Normalform besitzen, sind offensichtlich affin äquivalent.

(b) Die Zahl r hängt, wie wir gesehen haben, nur von β ab. Für den Untervektorraum

$$\mathrm{Rad}\,\beta := \{x \in V \mid \beta(x,y) = 0 \text{ für alle } y \in V\},$$

der das *Radikal* oder der *Ausartungsraum* von β heißt, gilt $\dim \mathrm{Rad}\,\beta = n - r$.

(c) Mit Hilfe der affinen Normalformen können wir nun leicht klären, welche der Quadriken eigentlich sind und welche nicht. Der Einfachheit wegen bezeichnen wir das ausgezeichnete Koordinatensystem in allen drei Fällen mit $(O;\, v_1,\dots,v_n)$.

Fall (I): Hier ist Q für $p = r$ ein affiner Unterraum der Dimension $n - r$. Für $p < r$ ist Q eine eigentliche Quadrik, denn auf Q liegen die $n + 1$ affin unabhängigen Punkte $O, Y_1,\dots,Y_n$, wobei $Y_1,\dots,Y_n$ durch

$$\begin{aligned}
\overrightarrow{OY_i} &= v_i - v_r\ , \quad i = 1,\dots,p\ ,\\
\overrightarrow{OY_i} &= v_1 + v_i\ , \quad i = p+1,\dots,r\ ,\\
\overrightarrow{OY_i} &= v_i\ , \quad i = r+1,\dots,n\ .
\end{aligned}$$

gegeben sind.

Fall (II): Q ist hier stets eine eigentliche Quadrik, denn es liegen auf Q die $n + 1$ affin unabhängigen Punkte $Y_0,\dots,Y_n$ mit den Ortsvektoren

$$\begin{aligned}
\overrightarrow{OY_0} &= -v_1\ ,\\
\overrightarrow{OY_i} &= v_i\ , \quad i = 1,\dots,p\ ,\\
\overrightarrow{OY_i} &= \sqrt{2}\, v_1 + v_i\ , \quad i = p+1,\dots,r\ ,\\
\overrightarrow{OY_i} &= v_1 + v_i\ , \quad i = r+1,\dots,n\ .
\end{aligned}$$

Fall (III): Auch in diesem Fall ist Q immer eigentlich, denn die $n + 1$ Punkte $O, Y_1,\dots,Y_n$ mit

$$\begin{aligned}
\overrightarrow{OY_i} &= v_i + \tfrac{1}{2} v_n\ , \quad i = 1,\dots,p\ ,\\
\overrightarrow{OY_i} &= v_i - \tfrac{1}{2} v_n\ , \quad i = p+1,\dots,r\ ,\\
\overrightarrow{OY_i} &= v_i\ , \quad i = r+1,\dots,n-1\ ,\\
\overrightarrow{OY_n} &= 2\,(v_1 + v_n)\ .
\end{aligned}$$

liegen auf Q und sind affin unabhängig.

Beispiele. Für $n = 1,2,3$ ergeben sich aus Satz 10 die folgenden affinen Normalformen von Quadriken:

$n = 1:$	(I)	$x_1^2 = 0\ ,$	(*Punkt*, $r = 1,\ p = 1$)
	(II)	$x_1^2 = 1\ .$	(*Punktepaar*, $r = 1,\ p = 1$)

$n = 2$:

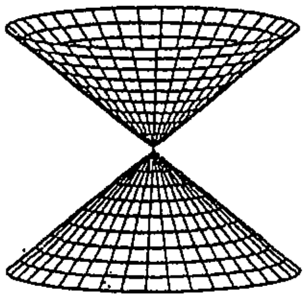
(I)	$x_1^2 + x_2^2 = 0$,	(*Punkt* , $r = 2$, $p = 2$)
	$x_1^2 - x_2^2 = 0$,	(*Paar sich schneidender Geraden* , $r = 2$, $p = 1$)
	$x_1^2 = 0$,	(*Gerade* , $r = 1$, $p = 1$)
(II)	$x_1^2 + x_2^2 = 1$,	(*Ellipse* , $r = 2$, $p = 2$)
	$x_1^2 - x_2^2 = 1$,	(*Hyperbel* , $r = 2$, $p = 1$)
	$x_1^2 = 1$,	(*Paar paralleler Geraden* , $r = 1$, $p = 1$)
(III)	$x_1^2 = 2\,x_2$.	(*Parabel* , $r = 1$, $p = 1$)

$n = 3$:

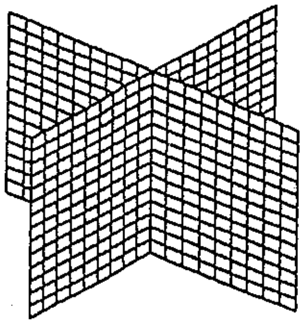
(I)	$x_1^2 + x_2^2 + x_3^2 = 0$,	(*Punkt* , $r = 3$, $p = 3$)
	$x_1^2 + x_2^2 - x_3^2 = 0$,	(*Kegel* , $r = 3$, $p = 2$)
	$x_1^2 + x_2^2 = 0$,	(*Gerade* , $r = 2$, $p = 2$)
	$x_1^2 - x_2^2 = 0$,	(*Paar sich schneidender Ebenen* , $r = 2$, $p = 1$)
	$x_1^2 = 0$,	(*Ebene* , $r = 1$, $p = 1$)
(II)	$x_1^2 + x_2^2 + x_3^2 = 1$,	(*Ellipsoid* , $r = 3$, $p = 3$)
	$x_1^2 + x_2^2 - x_3^2 = 1$,	(*einschaliges Hyperboloid* , $r = 3$, $p = 2$)
	$x_1^2 - x_2^2 - x_3^2 = 1$,	(*zweischaliges Hyperboloid* , $r = 3$, $p = 1$)
	$x_1^2 + x_2^2 = 1$,	(*elliptischer Zylinder* , $r = 2$, $p = 2$)
	$x_1^2 - x_2^2 = 1$,	(*hyperbolischer Zylinder* , $r = 2$, $p = 1$)
	$x_1^2 = 1$,	(*Paar paralleler Ebenen* , $r = 1$, $p = 1$)
(III)	$x_1^2 + x_2^2 = 2\,x_3$,	(*elliptisches Paraboloid* , $r = 2$, $p = 2$)
	$x_1^2 - x_2^2 = 2\,x_3$,	(*hyperbolisches Paraboloid* , $r = 2$, $p = 1$)
	$x_1^2 = 2\,x_3$.	(*parabolischer Zylinder* , $r = 1$, $p = 1$)

Eigentliche Quadriken im dreidimensionalen reellen affinen Raum.

I Kegel

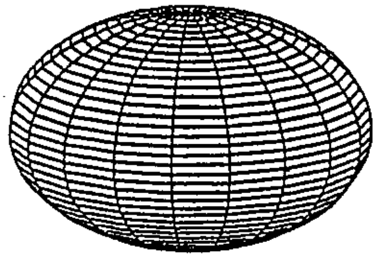


Kegel

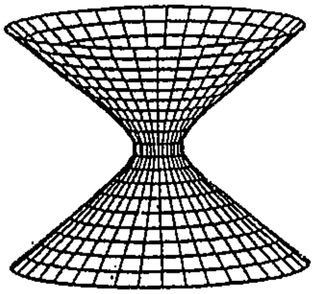


Paar sich schneidender Ebenen

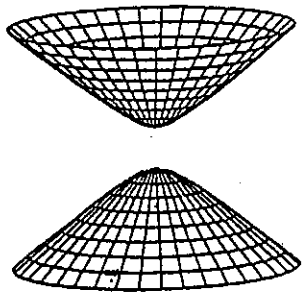
II Echte Mittelpunktsquadriken



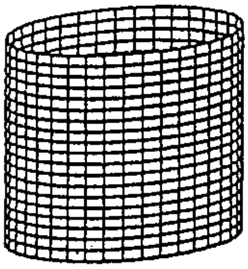
Ellipsoid



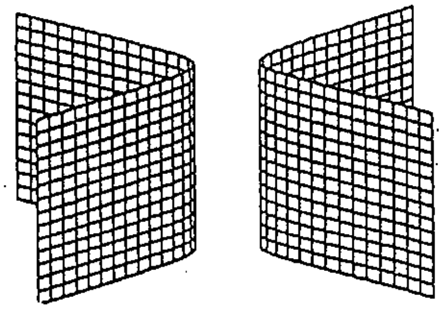
einschaliges Hyperboloid



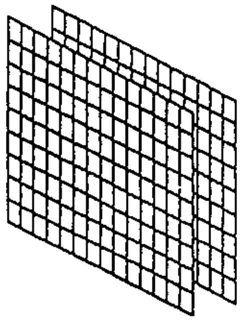
zweischaliges Hyperboloid



elliptischer Zylinder

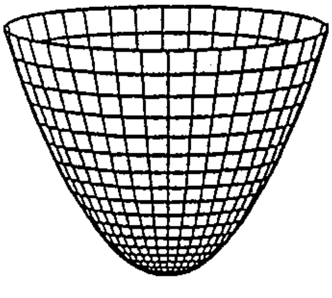


hyperbolischer Zylinder

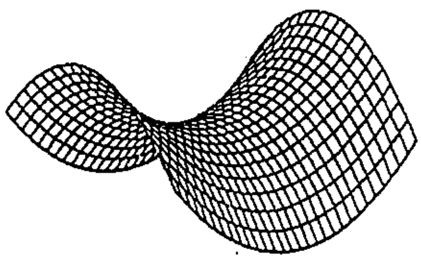


Paar paralleler Ebenen

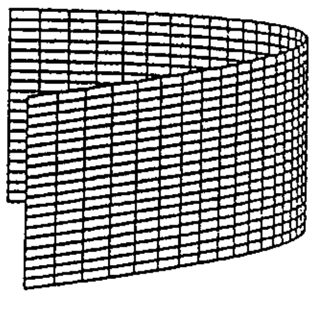
III Paraboloide



elliptisches Paraboloid



hyperbolisches Paraboloid



parabolischer Zylinder

Für die praktische Herleitung der Gleichungen (I), (II), (III) gibt es ein einfaches Verfahren, das mit quadratischer Ergänzung arbeitet. Statt dieses Verfahren allgemein zu behandeln, wollen wir es an zwei Beispielen demonstrieren.

Beispiele. **(a)** Im affinen Raum $\mathbb{R}^3$ sei bezüglich eines affinen Koordinatensystems $(O; v_1, v_2, v_3)$ für jedes $t \in \mathbb{R}$ eine Quadrik Q_t durch die Gleichung

$$2\,x_1^2 + 4\,x_1\,x_2 - 6\,x_1\,x_3 + 3\,x_2^2 - 8\,x_2\,x_3 + \tfrac{5}{2}\,x_3^2 + 2\,t\,x_3 + t^2 \;=\; 0$$

gegeben. Wir beginnen mit $2\,x_1^2$ und betrachten alle anderen Summanden, in denen x_1 als Faktor auftritt. Durch quadratische Ergänzung erhalten wir

$$2\,[x_1^2 + 2\,x_1\,(x_2 - \tfrac{3}{2}\,x_3) + (x_2 - \tfrac{3}{2}\,x_3)^2] - 2\,(x_2 - \tfrac{3}{2}\,x_3)^2$$
$$+\,3\,x_2^2 - 8\,x_2\,x_3 + \tfrac{5}{2}\,x_3^2 + 2\,t\,x_3 + t^2 \;=\; 0$$

und somit

$$2\,(x_1 + x_2 - \tfrac{3}{2}\,x_3)^2 + x_2^2 - 2\,x_2\,x_3 - 2\,x_3^2 + 2\,t\,x_3 + t^2 = 0\,.$$

In einer weiteren quadratischen Ergänzung berücksichtigen wir x_2^2 und alle Summanden, in denen x_2 als Faktor vorkommt:

$$2\,(x_1 + x_2 - \tfrac{3}{2}\,x_3)^2 + (x_2^2 - 2\,x_2\,x_3 + x_3^2) - 3\,x_3^2 + 2\,t\,x_3 + t^2 = 0\,.$$

Eine letzte quadratische Ergänzung ergibt dann

$$2\,(x_1 + x_2 - \tfrac{3}{2}\,x_3)^2 + (x_2 - x_3)^2 - 3\,(x_3^2 - \tfrac{2}{3}\,t\,x_3 + \tfrac{1}{9}\,t^2) + \tfrac{1}{3}\,t^2 + t^2 \;=\; 0\,.$$

Setzen wir

$$y_1 \;=\; \sqrt{2}\,(x_1 + x_2 - \tfrac{3}{2}\,x_3)\,,$$

$$y_2 \;=\; x_2 - x_3\,,$$

$$y_3 \;=\; \sqrt{3}\,(x_3 - \tfrac{1}{3}\,t)\,,$$

so erhalten wir für Q_t die Gleichung

$$y_1^2 + y_2^2 - y_3^2 + \tfrac{4}{3}t^2 = 0 .$$

Für $t = 0$ hat diese Gleichung Normalgestalt (I). Es handelt sich um einen Kegel mit $p = 2$ und $r = 3$.

Für $t \neq 0$ dividieren wir durch $-\frac{4}{3}t^2$, setzen

$$z_1 = \frac{\sqrt{3}}{2t}y_3 \ , \ z_2 = \frac{\sqrt{3}}{2t}y_2 \ , \ z_3 = \frac{\sqrt{3}}{2t}y_1$$

und erhalten die Normalform

$$z_1^2 - z_2^2 - z_3^2 = 1 .$$

Dies ist eine Quadrik vom Typ (II), $p = 1$, $r = 3$, also ein zweischaliges Hyperboloid.

Die obige Rechnung liefert uns für jeden Fall auch gleichzeitig die Affinität φ, für die $\varphi(Q)$ Normalform hat, bzw. das neue affine Koordinatensystem $(O';v_1',v_2',v_3')$, bezüglich dessen die Gleichung von Q Normalform annimmt.

Für $t = 0$ ist die gesuchte affine Abbildung φ durch die Gleichungen

$$\begin{aligned} z_1 &= \sqrt{2}\,x_1 + \sqrt{2}\,x_2 - \frac{3}{\sqrt{2}}x_3 , \\ z_2 &= x_2 - x_3 , \\ z_3 &= \sqrt{3}\,x_3 \end{aligned}$$

gegeben, hat also bezüglich des Koordinatensystems $(O;\, v_1,v_2,v_3)$ die Matrixdarstellung

$$\hat{x} \longmapsto \begin{bmatrix} \sqrt{2} & \sqrt{2} & -3/\sqrt{2} \\ 0 & 1 & -1 \\ 0 & 0 & \sqrt{3} \end{bmatrix} \hat{x} \ , \ \hat{x} \in \mathbb{R}^3 .$$

Dann hat φ^{-1} die Form

$$\hat{x} \longmapsto \begin{bmatrix} 1/\sqrt{2} & -1 & 1/2\sqrt{3} \\ 0 & 1 & 1/\sqrt{3} \\ 0 & 0 & 1/\sqrt{3} \end{bmatrix} \hat{x} \ , \ \hat{x} \in \mathbb{R}^3 .$$

Für das neue Koordinatensystem $(O';\, v_1',v_2',v_3')$ gilt somit $O' = \varphi^{-1}(O) = O$ und

$$\begin{aligned} v_1' &= 1/\sqrt{2}\, v_1 \ , \\ v_2' &= -v_1 + v_2 \ , \\ v_3' &= 1/2\sqrt{3}\, v_1 + 1/\sqrt{3}\, v_2 + 1/\sqrt{3}\, v_3 \, . \end{aligned}$$

Für $t \neq 0$ ist φ durch die Gleichungen

$$\begin{aligned} v_1' &= \frac{3}{2t}\, v_3 - \frac{1}{2}, \\ v_2' &= \frac{\sqrt{3}}{2t}\, v_2 - \frac{\sqrt{3}}{2t}\, v_3 \, , \\ v_3' &= \frac{\sqrt{6}}{2t}\, v_1 + \frac{\sqrt{6}}{2t}\, v_2 - \frac{3\sqrt{6}}{4t}\, v_3 \end{aligned}$$

gegeben. Hier hat φ bezüglich $(O;\, v_1,v_2,v_3)$ die Matrixdarstellung

$$\hat{x} \longmapsto \frac{\sqrt{3}}{2t} \begin{bmatrix} 0 & 0 & \sqrt{3} \\ 0 & 1 & -1 \\ \sqrt{2} & \sqrt{2} & -3/\sqrt{2} \end{bmatrix} \hat{x} + \begin{bmatrix} -1/2 \\ 0 \\ 0 \end{bmatrix} , \quad \hat{x} \in \mathbb{R}^3.$$

Für φ^{-1} erhalten wir somit die Darstellung

$$\hat{x} \longmapsto \frac{2t}{\sqrt{3}} \begin{bmatrix} 1/2\sqrt{3} & -1 & 1/\sqrt{2} \\ 1/\sqrt{3} & 1 & 0 \\ 1/\sqrt{3} & 0 & 0 \end{bmatrix} \hat{x} + \begin{bmatrix} t/6 \\ t/3 \\ t/3 \end{bmatrix} , \quad \hat{x} \in \mathbb{R}^3$$

Also ist $\overrightarrow{OO'} = \frac{t}{6}\, v_1 + \frac{t}{3}\, v_2 + \frac{t}{3}\, v_3$ und

$$\begin{aligned} v_1' &= \frac{t}{3}\, v_1 + \frac{2t}{3}\, v_2 + \frac{2t}{3}\, v_3, \\ v_2' &= -\frac{2t}{\sqrt{3}}\, v_1 + \frac{2t}{\sqrt{3}}\, v_2 \ , \\ v_3' &= \frac{2t}{\sqrt{6}}\, v_1 \, . \end{aligned}$$

(b) Im affinen Raum $\mathbb{R}^3$ sei bezüglich eines affinen Koordinatensystems eine Quadrik Q durch die Gleichung

$$x_1 x_2 + x_1 x_3 + x_2 x_3 + 2 x_1 - 1 = 0$$

gegeben. Um hier mit quadratischer Ergänzung arbeiten zu können, müssen wir uns zunächst geeignete quadratische Terme verschaffen. Dies geschieht mit Hilfe der folgenden Koordinatentransformation

$$\begin{aligned} x_1 &= y_1 - y_2 \\ x_2 &= y_1 + y_2 \\ x_3 &= y_3 \end{aligned}$$

Die neue Gleichung für Q hat dann die Form

$$y_1^2 - y_2^2 + 2 y_1 y_3 + 2 y_1 - 2 y_2 - 1 = 0 ,$$

auf die wir nun das Verfahren der quadratischen Ergänzung anwenden können. Es ergibt sich als Normalform

$$z_1^2 + z_2^2 - z_3^2 = 0,$$

also ist Q ein Kegel.

§ 4 Affine Klassifikation der Quadriken

Mit Satz 10 ist die affine Klassifikation der Quadriken noch nicht vollständig. Dazu muß noch gezeigt werden, daß zwei Quadriken genau dann affin äquivalent sind, wenn ihre Gleichungen dieselbe Normalform besitzen.

Die eine Richtung ist einfach. Besitzen Q und $\tilde{Q}$ bezüglich geeigneter affiner Koordinatensysteme dieselbe Normalform, so bildet die Affinität φ, die diese Koordinatensysteme aufeinander abbildet, auch Q auf $\tilde{Q}$ ab oder umgekehrt $\tilde{Q}$ auf Q.

Die andere Richtung ist komplizierter. Gilt $\varphi(Q) = \tilde{Q}$ mit einer Affinität φ, so müssen wir zeigen, daß die Normalformen von Q und von $\tilde{Q}$ zur gleichen Klasse (I), (II) oder (III) gehören und daß innerhalb dieser Klassen die Normalformen auch noch in den Größen p und r übereinstimmen.

Für die Zahlen r und $\tilde{r}$ haben wir die Gleichheit schon bewiesen, denn es sind die Ränge der symmetrischen Matrizen A bzw. $\tilde{A}$, welche die quadratischen Anteile von Q bzw. $\tilde{Q}$ beschreiben. Nach (***), S. 306, gilt $\tilde{A} = C^T A\ C$ mit einer regulären Matrix C, also ist $\tilde{r} = r$. Da weiterhin $\operatorname{Rg} A = \operatorname{Rg}(a\ A)$ für alle $a \neq 0$ gilt, ist r auch invariant gegenüber Multiplikationen der definierenden Gleichung von Q mit Zahlen $a \neq 0$.

Für die beiden anderen Behauptungen genügt es nicht mehr, die Gleichungen nur algebraisch umzuformen, vielmehr müssen jetzt auch geometrische Eigenschaften der Quadriken berücksichtigt werden, die invariant sind gegenüber affinen Abbildungen. Eine solche Eigenschaft ist z.B. die, einen Mittelpunkt zu besitzen.

Definition. $M \in \mathbb{A}$ heißt *Mittelpunkt* einer Quadrik Q, wenn die Punktspiegelung an M die Quadrik Q auf sich abbildet. Eine Quadrik Q, die einen Mittelpunkt besitzt, heißt *Mittelpunktsquadrik*. Eine Mittelpunktsquadrik Q, bei der kein Mittelpunkt auf Q liegt, heißt *echte Mittelpunktsquadrik*.

Besitzt Q einen Mittelpunkt $M \in Q$, so heißt dieser eine *Spitze* oder ein *Doppelpunkt* und Q heißt *Hyperkegel* oder kurz *Kegel*.

Quadriken ohne Mittelpunkt heißen *parabolisch.*

Beispiel. Jeder affine Unterraum L ist ein Kegel und genau die Punkte von L sind die Spitzen dieser nichteigentlichen Quadrik.

Für eigentliche Quadriken beantwortet der folgende Satz die Frage nach der Existenz von Mittelpunkten.

Satz 11. *Es sei* $Q = \{X \in \mathbb{A} \mid \overrightarrow{OX} = x\ ,\ \beta(x,x) + 2\ \Phi(x) + c = 0\}$ *eine eigentliche Quadrik. Dann gilt:*

(a) $M \in \mathbb{A}$ *ist genau dann Mittelpunkt von* Q, *wenn* $\beta(\cdot,\overrightarrow{OM}) + \Phi$ *die Nullform ist. Die Menge der Mittelpunkte von* Q *ist entweder leer oder ein affiner Unterraum von* $\mathbb{A}$ *mit Richtungsraum* Rad β.

(b) $M \in \mathbb{A}$ *ist genau dann eine Spitze von* Q, *wenn* $\beta(\cdot,\overrightarrow{OM}) + \Phi$ *die Nullform ist und* $\Phi(\overrightarrow{OM}) + c = 0$ *gilt. In diesem Fall ist jeder Mittelpunkt von* Q *eine Spitze.*

Beweis. (a) $M \in \mathbb{A}$ ist genau dann Mittelpunkt von Q, wenn für jedes $X \in Q$ auch der an M gespiegelte Punkt $\sigma(X)$ zu Q gehört. Der Ortsvektor von $\sigma(X)$ ist $\overrightarrow{O\sigma(X)} = -\overrightarrow{OX} + 2\ \overrightarrow{OM}$. Somit ist M genau dann Mittelpunkt von Q, wenn für alle Punkte $X \in \mathbb{A}$ aus

$$\beta(\overrightarrow{OX},\overrightarrow{OX}) + 2\ \Phi(\overrightarrow{OX}) + c = 0$$

stets auch

$$\beta(-\overrightarrow{OX} + 2\ \overrightarrow{OM},-\overrightarrow{OX} + 2\ \overrightarrow{OM}) + 2\ \Phi(-\overrightarrow{OX} + 2\ \overrightarrow{OM}) + c = 0$$

folgt. Dazu äquivalent ist, daß für alle $X \in Q$

$$(\#) \qquad \beta(\overrightarrow{MX},\overrightarrow{OM}) + \Phi(\overrightarrow{MX}) = 0$$

gilt.

Da Q eine eigentliche Quadrik ist, gibt es $n + 1$ affin unabhängige Punkte $Y_0,\dots,Y_n$ auf Q. Für einen Mittelpunkt M von Q gilt dann notwendigerweise

$$\beta(\overrightarrow{MY_i},\overrightarrow{OM}) + \Phi(\overrightarrow{MY_i}) = 0\quad ,\quad i = 0,\dots,n\ .$$

Durch Subtraktion erhalten wir daraus

$$\beta(\overrightarrow{Y_0Y_i},\overrightarrow{OM}) + \Phi(\overrightarrow{Y_0Y_i}) = 0 \quad , \quad i = 1,\ldots,n\,.$$

Da $(\overrightarrow{Y_0Y_1},\ldots,\overrightarrow{Y_0Y_n})$ eine Basis von V ist, muß also die Linearform $\beta(\cdot,\overrightarrow{OM}) + \Phi$ die Nullform sein. Ist umgekehrt dies der Fall, so ist M wegen (#) auch tatsächlich ein Mittelpunkt von $\mathcal{Q}$.

Sei nun L die Menge der Mittelpunkte von $\mathcal{Q}$. Existiert ein Mittelpunkt M von $\mathcal{Q}$, so gilt für jeden anderen Mittelpunkt M' entsprechend $\beta(\cdot,\overrightarrow{OM'}) + \Phi = O$. Durch Subtraktion erhalten wir $\beta(\cdot,\overrightarrow{MM'}) = O$, also $\overrightarrow{MM'} \in \operatorname{Rad}\beta$. Somit gilt

$$L \subset \{X \in \mathbb{A} \mid \overrightarrow{MX} \in \operatorname{Rad}\beta\}.$$

Umgekehrt ist jeder Punkt X mit $\overrightarrow{MX} \in \operatorname{Rad}\beta$ wegen

$$\beta(\cdot,\overrightarrow{OX}) + \Phi \;=\; \beta(\cdot,\overrightarrow{OM}) + \beta(\cdot,\overrightarrow{MX}) + \Phi \;=\; O$$

ein Mittelpunkt von $\mathcal{Q}$, also gilt Gleichheit, und L ist somit ein affiner Unterraum mit Richtungsraum $\operatorname{Rad}\beta$.

(b) Der Mittelpunkt M liegt genau dann auf $\mathcal{Q}$, wenn $\beta(\overrightarrow{OM},\overrightarrow{OM}) + 2\,\Phi(\overrightarrow{OM}) + c = 0$ gilt. Dies ist wegen $\beta(\overrightarrow{OM},\overrightarrow{OM}) + \Phi(\overrightarrow{OM}) = 0$ gleichwertig mit $\Phi(\overrightarrow{OM}) + c = 0$.

Ist M eine Spitze von $\mathcal{Q}$, so gilt für jeden anderen Mittelpunkt S wegen $M \in \mathcal{Q}$ und $\overrightarrow{MS} \in \operatorname{Rad}\beta$:

$$\begin{aligned}&\beta(\overrightarrow{OS},\overrightarrow{OS}) + 2\,\Phi(\overrightarrow{OS}) + c\\ &= (\beta(\overrightarrow{OM},\overrightarrow{OM}) + 2\,\Phi(\overrightarrow{OM}) + c) + 2\,(\beta(\overrightarrow{MS},\overrightarrow{OM}) + \Phi(\overrightarrow{MS})) + \beta(\overrightarrow{MS},\overrightarrow{MS}) \;=\; 0,\end{aligned}$$

also $S \in \mathcal{Q}$. ∎

Bemerkung. Hat die eigentliche Quadrik $\mathcal{Q}$ bezüglich eines affinen Koordinatensystems die Darstellung

$$\hat{x}^{\mathsf{T}} A\,\hat{x} \;+\; 2\,b^{\mathsf{T}}\,\hat{x} \;+\; c \;=\; 0\,,$$

so ist M genau dann Mittelpunkt von $\mathcal{Q}$, wenn sein Koordinatenvektor $\hat{m} \in \mathbb{R}^n$ eine Lösung des inhomogenen LGS $A\,\hat{x} = -b$ ist, und M ist eine Spitze, wenn außerdem $b^{\mathsf{T}}\,\hat{m} + c = 0$ gilt. Ist die Lösungsmenge von $A\,\hat{x} = -b$ nicht leer, so hat sie nach

Satz 11 die Dimension $d = \dim \operatorname{Rad} \beta = n - \operatorname{Rg} A$.

Beispiel. Im affinen Raum $\mathbb{R}^3$ sei bezüglich eines affinen Koordinatensystems für jedes $t \in \mathbb{R}$ eine Quadrik Q_t gegeben durch die Gleichung

$$x_1^2 - 8\,x_1\,x_2 + 6\,x_1\,x_3 - 9\,x_2^2 - 24\,x_2\,x_3 + t\,x_3^2 + 20\,x_2 + 3\,x_3 + 1 \;=\; 0\,.$$

In Matrixschreibweise lautet diese Gleichung

$$x^{\mathsf{T}} A\, x \;+\; 2\, b^{\mathsf{T}}\, x \;+\; c \;=\; 0$$

mit

$$x \;=\; \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \text{ und } A \;=\; \begin{bmatrix} 1 & -4 & 3 \\ -4 & -9 & -12 \\ 3 & -12 & t \end{bmatrix}, \; b \;=\; \frac{1}{2}\begin{bmatrix} 0 \\ 20 \\ 3 \end{bmatrix}, \; c = 1\,.$$

Q_t besitzt genau dann einen Mittelpunkt mit Ortsvektor m, wenn m Lösung des inhomogenen linearen Gleichungssystems $A\,x = -\,b$ ist.

Im vorliegenden Fall lautet die zugehörige erweiterte Matrix

$$\left[\begin{array}{ccc|c} 1 & -4 & 3 & 0 \\ -4 & -9 & -12 & -10 \\ 3 & -12 & t & -3/2 \end{array}\right]$$

Dieses LGS ist genau für $t \neq 9$ lösbar. Also ist Q_t für $t = 9$ eine Quadrik ohne Mittelpunkt, d.h. parabolisch. Durch das Verfahren der quadratischen Ergänzung erhalten wir in diesem Fall die Normalform

$$z_1^2 - z_2^2 \;=\; 2\,z_3\,,$$

mit $p = 1$ und $r = 2$. Q_9 ist ein hyperbolisches Paraboloid.

Für $t \neq 9$ besitzt Q_t einen Mittelpunkt und die Menge der Mittelpunkte ergibt sich als Lösung des obigen linearen Gleichungssystems. Wegen Rang $A = 3$ gibt es für jedes $t \neq 9$ genau einen Mittelpunkt mit zugehörigem Ortsvektor

$$m = \begin{bmatrix} (16t - 99)/10(t-9) \\ 2/5 \\ -3/2(t-9) \end{bmatrix}.$$

Genau für $t = \frac{189}{20}$ ist der Mittelpunkt eine Spitze der Quadrik. In diesem Fall ist Q ein Kegel mit der Normalform

$$z_1^2 + z_2^2 - z_3^2 = 0.$$

Für $t \neq 9$, $t \neq \frac{189}{20}$ erhalten wir echte Mittelpunktsquadriken, und zwar für $9 < t < \frac{189}{20}$ zweischalige Hyperboloide und für $t < 9$ oder $\frac{189}{20} < t$ einschalige Hyperboloide (Beweis als Übungsaufgabe).

Bemerkung. Die Eigenschaft, Mittelpunkt zu sein, ist invariant unter Affinitäten φ.

Beweis. Ist σ die Punktspiegelung an M, so ist $\varphi\sigma\varphi^{-1}$ die Punktspiegelung an $\varphi(M)$. Nun ist M genau dann Mittelpunkt von Q, wenn aus $X \in Q$ stets auch $\sigma(X) \in Q$ folgt. Gleichwertig hierzu ist die Bedingung, daß aus $\varphi(X) \in \varphi(Q)$ stets $\varphi\sigma(X) \in \varphi(Q)$ folgt, d.h. $\varphi\sigma\varphi^{-1}(\varphi(X)) \in \varphi(Q)$ gilt. Ist also M Mittelpunkt von Q, so ist $\varphi(M)$ Mittelpunkt von $\varphi(Q)$ und umgekehrt. ∎

Wir wenden nun Satz 11 bzw. die anschließende Bemerkung auf die Quadriken mit den Gleichungen der Form (I), (II) oder (III) an.

Fall (I): Hier gilt Rg $A = r$, $b = o$, $c = 0$ und der Ursprung ist eine Spitze, es handelt sich also um Kegel. Sind die Quadriken eigentlich, so ist nach Satz 11 die Menge der Spitzen jeweils ein affiner Unterraum der Dimension $n-r$. Sind sie nicht eigentlich, was nur für $p = r$ eintreten kann, so handelt es sich um affine Unterräume der Dimension $n - r$, und jeder Punkt des Unterraumes ist eine Spitze.

Der Name Kegel rührt daher, daß jede Spitze M die Eigenschaft hat, daß die Verbindungsgerade mit einen beliebigen Punkt $X \in Q$, $X \neq M$, ganz auf Q liegt. Eine solche Gerade heißt dann eine *Mantellinie* von Q.

Umgekehrt ist jede Quadrik Q, die einen Punkt M enthält, der $MX \subset Q$ für alle

$X \in Q$, $X \neq M$, erfüllt, ein Kegel (Beweis als Übungsaufgabe).

Fall (II): Hier gilt Rg $A = r$, $b = o$, $c = 1$ und der Ursprung ist ein Mittelpunkt, aber keine Spitze, also handelt es sich um echte Mittelpunktsquadriken.

Fall (III): Hier gilt Rg $A = r < n$, $b = -e_n$ und $c = 0$. Alle Quadriken sind eigentlich. Damit ein Mittelpunkt existiert, muß das LGS $A\,\hat{x} = -b = e_n$ lösbar sein. Wegen Rg $(A \mid e_n) = r + 1 \neq$ Rg A ist dies aber nicht der Fall, also existiert kein Mittelpunkt, die Quadriken sind somit parabolisch.

Als weiteres Teilergebnis können wir somit festhalten, daß affin äquivalente Quadriken Q und $\tilde{Q}$ aufgrund ihres Mittelpunktverhaltens zur selben Klasse (I), (II) oder (III) gehören.

Um schließlich die Gleichheit von p und $\tilde{p}$ zu beweisen, zeigen wir, daß mit jeder Quadrik affine Unterräume verbunden sind, deren Dimension im wesentlichen durch diese Zahlen gegeben sind.

Satz 12. *Besitzt die Quadrik $Q \subset \mathbb{A}$ die Normalform*

(I) $$x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_r^2 = 0$$

mit $0 < p \leq r \leq n$, $r-p \leq p$, so ist $n-p$ die maximale Dimension aller ganz auf Q liegenden affinen Unterräume.

Beweis. Für $p = r$ ist Q selbst ein $(n-p)$-dimensionaler affiner Unterraum und die Behauptung ist bewiesen. Seien nun $p < r$ und L ein affiner Unterraum auf Q mit zugehörigem Richtungsraum U_L. Wir betrachten den affinen Unterraum $L' = L \vee O$ und zeigen, daß auch L' auf Q liegt:

Weil der Ursprung O Spitze von Q ist, liegen alle Verbindungsgeraden $O \vee X$, $X \in L$, ganz auf Q. Weiterhin liegt jeder Punkt X aus dem zu L parallelen Unterraum durch O auf Q. Wählen wir nämlich in L einen beliebigen Punkt P, so liegt die Gerade durch P mit Richtung $\overrightarrow{OX}$ wegen $\overrightarrow{OX} \in U_L$ in L, also auf Q. Somit gilt für alle $t \in \mathbb{R}$

$$0 = \beta(\overrightarrow{OP} + t\,\overrightarrow{OX}\,,\,\overrightarrow{OP} + t\,\overrightarrow{OX}) = 2\,t\,\beta(\overrightarrow{OP},\overrightarrow{OX}) + t^2\,\beta(\overrightarrow{OX},\overrightarrow{OX})\,.$$

Daraus folgt $\beta(\overrightarrow{OX},\overrightarrow{OX}) = 0$, also ist $X \in Q$ und es gilt somit $L' \subset Q$.

Wir schneiden L' mit dem affinen Unterraum L_1, der durch das LGS

$$x_{p+1} = \ldots = x_n = 0$$

gegeben ist. Es gilt $\dim L_1 = p$ und $L_1 \cap Q = \{O\}$. Damit gilt auch $L_1 \cap L' = \{O\}$. Aus der Dimensionsformel für affine Unterräume folgt nun

$$\dim L' + \dim L_1 \;=\; \dim(L' \vee L_1) + \dim(L' \cap L_1) \;\le\; n\,,$$

also

$$\dim L \;\le\; \dim L' \;\le\; n - p\,.$$

Daß die Dimension $n - p$ auch tatsächlich vorkommt, zeigt das Beispiel des affinen Unterraumes $L \subset Q$, der durch das LGS

$$\begin{array}{lllll}
x_1 & & + x_{p+1} & & = 0 \\
& \ddots & & \ddots & \\
& x_{r-p} & & + x_r & = 0 \\
& & x_{r-p+1} & & = 0 \\
& & & \ddots & \\
& & & x_p & = 0
\end{array}$$

gegeben ist. ∎

Satz 13. *Besitzt die Quadrik $Q \subset \mathbb{A}$ die Normalform*

(III) $$x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_r^2 \;=\; 2\,x_n$$

mit $0 < p \le r < n$, $r-p \le p$, *so ist* $n-p-1$ *die maximale Dimension aller ganz auf* Q *liegenden affinen Unterräume.*

Beweis. Sei L ein affiner Unterraum, der ganz auf Q gelegen ist. Wir schneiden L mit dem Unterraum L_1 , der durch das LGS

$$x_{p+1} = \ldots = x_r = 0\;,\; x_n = -1$$

gegeben ist. Es gilt $\dim L_1 = n + p - r - 1$ und $L_1 \cap Q = \emptyset$. Damit gilt auch $L_1 \cap L = \emptyset$ und wir erhalten aus der Dimensionsformel

$$\dim L + \dim L_1 = \dim(L \vee L_1) + \dim(U_L \cap U_{L_1}) - 1$$

die Ungleichung

$$\dim L \leq r - p + \dim(U_L \cap U_{L_1}) .$$

Sei nun $Y \in L$ fest gewählt. Dann liegt für jeden Vektor $x \in U_L \cap U_{L_1}$, $x \neq o$, die Gerade durch Y mit der Richtung $[x]$ in L und somit ganz auf Q. Für die Ortsvektoren der Punkte X auf dieser Gerade gilt $\overrightarrow{OX} = \overrightarrow{OY} + \overrightarrow{YX} = \overrightarrow{OY} + t\,x$, $t \in \mathbb{R}$.

Weil der Richtungsraum U_{L_1} durch das homogene LGS

$$x_{p+1} = \ldots = x_r = 0 \ , \ x_n = 0$$

gegeben ist, hat der Koordinatenvektor $\hat{x}$ des Vektors $x \in U_L \cap U_{L_1}$ die Form

$$\hat{x} = (x_1 ,\ldots, x_p , 0 ,\ldots, 0 , x_{r+1} ,\ldots, x_{n-1} , 0) .$$

Somit erfüllt für jedes $t \in \mathbb{R}$ das n–Tupel

$$(y_1 + t\,x_1 ,\ldots, y_p + t\,x_p , y_{p+1} ,\ldots, y_r , y_{r+1} + t\,x_{r+1} ,\ldots, y_{n-1} + t\,x_{n-1} , y_n)$$

die Gleichung von Q. Speziell für $t = 1$ und $t = -1$ erhalten wir nach Addition die Gleichung

$$x_1^2 + \ldots + x_p^2 = 0 ,$$

also $x_1 = \ldots = x_p = 0$. Damit gilt

$$\dim(U_L \cap U_{L_1}) \leq n - (r + 1) = n - r - 1$$

und somit

$$\dim L \leq r - p + (n - r - 1) = n - p - 1.$$

Daß die Dimension $n - p - 1$ tatsächlich erreicht wird, zeigt das Beispiel des affinen Unterraumes $L \subset Q$, der durch das LGS

$$
\begin{array}{llllll}
x_1 & & & + x_{p+1} & & = 0 \\
& \ddots & & & \ddots & \\
& & x_{r-p} & & + x_r & = 0 \\
& & x_{r-p+1} & & & = 0 \\
& & & \ddots & & \\
& & & x_p & & = 0 \\
& & & x_n & & = 0
\end{array}
$$

gegeben ist. ■

In den Fällen (I) und (III) folgt somit aus der Äquivalenz der Quadriken Q und $\tilde{Q}$, daß $p = \tilde{p}$ gilt.

Im Fall (II) läßt sich dieser Schluß nicht durchführen. Ohne Beweis bemerken wir, daß hier für die maximale Dimension d der auf Q gelegenen affinen Unterräume

$$d = n - p \quad \text{für} \quad p > r - p$$

gilt bzw.

$$d = n - (r - p) - 1 \quad \text{für} \quad p \leq r - p\,.$$

Aus der Gleichheit der maximalen Dimensionen d und $\tilde{d}$ für Q und $\tilde{Q}$ können wir nun nicht mehr allgemein auf $p = \tilde{p}$ schließen. So gilt zum Beispiel im $\mathbb{R}^2$ sowohl für Ellipsen als auch für Hyperbeln $d = 0$, aber im ersten Fall ist $p = 2$, im zweiten $p = 1$.

Hier müssen wir andere affine Unterräume betrachten, die mit Q affin invariant verbunden sind. Am Beispiel der Ellipsen und Hyperbeln erkennen wir, daß die maximale Dimension der Unterräume durch den Mittelpunkt O, die Q nicht treffen, 0 bzw. 1 ist, und daß diese Zahlen jeweils mit der Zahl $n - p$ übereinstimmen.

Satz 14. *Besitzt die Quadrik $Q \subset \mathbb{A}$ die Normalform*

(II) $$x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_r^2 = 1$$

mit $0 < p \leq r \leq n$, so ist $n-p$ die maximale Dimension aller affinen Unterräume, die den Mittelpunkt O enthalten und mit Q keinen Punkt gemeinsam haben.

Beweis. Der affine Unterraum L, der durch das LGS $x_1 = \ldots = x_p = 0$ definiert ist, hat die Dimension $n - p$, enthält den Ursprung O und hat mit Q keinen Punkt gemeinsam.

Sei nun L_1 ein affiner Unterraum durch O mit $\dim L_1 > n - p$. Wir schneiden L_1 mit dem Unterraum L_2, der durch das LGS $x_{p+1} = \dots = x_n = 0$ gegeben ist. Es gilt $O \in L_2$ und $\dim L_2 = p$. Wegen $L_1 \cap L_2 \neq \emptyset$ folgt für die Dimension des Schnittes

$$\dim (L_1 \cap L_2) = \dim L_1 + \dim L_2 - \dim (L_1 \vee L_2)$$
$$\geq \dim L_1 + \dim L_2 - n > (n-p) + p - n = 0.$$

Es gibt also in $L_1 \cap L_2$ eine Gerade OX. Für die Koordinaten von X gilt dann

$$x_{p+1} = \dots = x_n = 0 \text{ und } x_1^2 + \dots + x_p^2 > 0.$$

Somit liegt der Punkt Y mit dem Ortsvektor

$$\overrightarrow{OY} = 1/(x_1^2 + \dots + x_p^2)^{\frac{1}{2}} \; \overrightarrow{OX}$$

sowohl in L_1 als auch auf Q, also ist $L_1 \cap Q \neq \emptyset$. Die Dimension $n - p$ von L ist daher maximal. ∎

Wegen Satz 14 folgt nun auch im verbliebenen Fall (II) für affin äquivalente Quadriken die Gleichheit von p und $\tilde{p}$. Wir fassen zusammen:

Satz 15. *Zwei Quadriken sind genau dann affin äquivalent, wenn ihre affinen Normalformen vom gleichen Typ* (I), (II) *oder* (III) *sind und außerdem in den Zahlen p und r übereinstimmen.*

Zum Abschluß dieses Paragraphen wollen wir noch die Frage klären, ob es möglich ist, eine Quadrik bezüglich eines festen Ursprungs durch verschiedene Gleichungen zu beschreiben, die sich nicht nur um einen Faktor $a \neq 0$ unterscheiden. Der folgende Satz zeigt, daß für eigentliche Quadriken diese Situation nicht eintreten kann. Er wird im letzen Abschnitt dieses Kapitels noch eine wichtige Rolle spielen.

Satz 16. *Es sei $Q \subset \mathbb{A}$ eine eigentliche Quadrik. Wird Q bezüglich eines festen Ursprungs $O \in \mathbb{A}$ jeweils durch die Gleichungen*

$$\beta(x,x) + 2\,\Phi(x) + c = 0 \quad und \quad \beta'(x,x) + 2\,\Phi'(x) + c' = 0$$

dargestellt, so gibt es eine reelle Zahl $a \neq 0$ mit $\beta' = a\,\beta$, $\Phi' = a\,\Phi$ und $c' = a\,c$.

Beweis. Wir unterscheiden 3 Fälle, je nachdem ob die affine Normalform der ersten Gleichung von Q vom Typ (I), (II) oder (III) ist.

1.Fall: Bezüglich eines geeigenten affinen Koordinatensystems $(O;v_1,\dots,v_n)$ hat die erste Gleichung die Normalform

(I) $$x_1^2 + \dots + x_p^2 - x_{p+1}^2 - \dots - x_r^2 = 0\,, \qquad 0 < p \le r \le n\,,\ r-p \le p\,,$$

und die zweite Gleichung geht über in die Form

$$\hat{x}^{\mathsf{T}} A\,\hat{x} + 2\,b^{\mathsf{T}}\hat{x} + c = 0\,.$$

Der Ursprung O ist Mittelpunkt von Q, also ist o Lösung des inhomogenen LGS $A\,\hat{x} = -b$ (vgl. S.319). Somit gilt $b = o$. Weil O außerdem auf Q liegt, folgt auch $c = 0$. Wir müssen noch zeigen, daß A eine Diagonalmatrix ist mit $a_{ii} = a \neq 0$ für $i = 1,\dots,p$, $a_{ii} = -a$ für $i = p+1,\dots,r$ und $a_{ii} = 0$ für $i = r+1,\dots,n$.

Da die Mittelpunkte von Q einen affinen Unterraum von $\mathbb{A}$ bilden, der ganz auf Q liegt und dessen Richtungsraum $[v_{r+1},\dots,v_n]$ ist, gilt $A\,\hat{v}_i = o$ für $i = r+1,\dots,n$, also $a_{ij} = 0$ für $i > r$ oder $j > r$. Der weitere Beweis erfolgt dadurch, daß wir die Koordinaten genügend vieler geeigneter Punkte von Q in die quadratische Gleichung $\hat{x}^{\mathsf{T}} A\,\hat{x} = 0$ einsetzen.

Sei $1 \le i \le p$, $p + 1 \le j \le r$. Dann liegen die Punkte Y mit $\overrightarrow{OY} = v_i \pm v_j$ auf Q. Also ist $(\hat{v}_i \pm \hat{v}_j)^{\mathsf{T}} A\,(\hat{v}_i \pm \hat{v}_j) = 0$, woraus durch Subtraktion zunächst $a_{ij} = \hat{v}_i^{\mathsf{T}} A\,\hat{v}_j = 0$ und dann $a_{ii} = \hat{v}_i^{\mathsf{T}} A\,\hat{v}_i = -\,\hat{v}_j^{\mathsf{T}} A\,\hat{v}_j = a_{jj}$ folgen. Somit gilt $a_{11} = \dots = a_{pp} = a$ und $a_{p+1,p+1} = \dots = a_{rr} = -a$ mit $a \neq 0$.

Für $1 \le i < j \le p$ liegen die Punkte Y mit $\overrightarrow{OY} = v_i + v_j + \sqrt{2}\,v_r$ auf Q. Daraus erhalten wir unter Verwendung der schon bekannten a_{ij} in diesen Fällen $a_{ij} = 0$.

Analog zeigen wir für $p + 1 \leq i < j \leq r$ mit Hilfe der durch die Ortsvektoren $\overrightarrow{OY} = \sqrt{2}\, v_1 + v_i + v_j$ gegebenen Punkte $Y \in Q$, daß auch hier $a_{ij} = \hat{v}_i^{\mathsf{T}} A\, \hat{v}_j = 0$ gilt.

Damit ist A eine Diagonalmatrix von der gewünschten Form, und die Behauptung des Satzes ist in diesem Fall bewiesen.

2.Fall: Q ist eine echte Mittelpunktsquadrik und die erste Gleichung hat nun die Normalform

$$\text{(II)} \qquad x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_r^2 = 1 \,, \qquad 0 < p \leq r \leq n\,.$$

Dieser Fall wird mit entsprechenden Methoden bewiesen. Zunächst folgt wieder $b = o$ und $a_{ij} = 0$ für $i > r$ oder $j > r$. Weil hier der Ursprung O nicht auf Q liegt, ist jetzt $c' \neq 0$ und wir können o.B.d.A. annehmen, daß die zweite Gleichung von Q bezüglich des ausgezeichneten Koordinatensystems $(O;v_1,\ldots,v_n)$ die Form $\hat{x}^{\mathsf{T}} A\, \hat{x} = 1$ hat.

Sei $1 \leq i \leq j \leq p$. Für diese i liegen die Einheitspunkte des gegebenen Koordinatensystems auf Q, also gilt $a_{ii} = 1$. Für $i \neq j$ betrachten wir die Punkte $Y \in Q$ mit den Ortsvektoren $y = 1/\sqrt{2}\ (v_i + v_j)$ und erhalten aus $\hat{y}^{\mathsf{T}}\ A\ \hat{y} = 1$ und den schon bekannten Ergebnissen $a_{ij} = 0$.

Für $1 \leq i \leq p$ und $p + 1 \leq j \leq r$ wählen wir die Punkte $Y \in Q$ mit $\overrightarrow{OY} = \sqrt{2}\, v_i \pm v_j$ und erhalten $a_{ij} = \hat{v}_i^{\mathsf{T}}\ A\ \hat{v}_j = 0$ und $a_{jj} = \hat{v}_j^{\mathsf{T}}\ A\ \hat{v}_j = -1$, und für $p + 1 \leq i < j \leq r$ wählen wir $Y \in Q$ mit $\overrightarrow{OY} = \sqrt{3}\, v_1 + v_i + v_j$ und erhalten $a_{ij} = \hat{v}_i^{\mathsf{T}}\ A\ \hat{v}_j = 0$.

Damit ist die Behauptung auch in diesem Fall bewiesen.

3. Fall: Die Quadrik ist ein Paraboloid. Die erste Gleichung besitzt die Normalform

$$\text{(III)} \qquad x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_r^2 = 2\, x_n \,, \qquad 0 < p \leq r < n,\ r-p \leq p\,,$$

und die zweite Gleichung geht wieder über in die Form

$$\hat{x}^{\mathsf{T}} A\, \hat{x} \ +\ 2\, b^{\mathsf{T}} \hat{x} \ +\ c = 0\,.$$

Wegen $O \in Q$ ist $c = 0$. Weil Q keine Mittelpunktsquadrik ist, muß $b \neq o$ sein.

Sei nun $X \neq O$ mit zugehörigem Koordinatenvektor $\hat{x} = (x_1,\ldots,x_{n-1},0)$ beliebig gewählt und sei

$$s := x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_r^2 .$$

Ist $s \neq 0$, so liegen alle Punkte Y mit den Koordinatenvektoren

$$\hat{y} = t\,\hat{x} + (s\,t^2)/2\;\hat{v}_n\,,\; t \in \mathbb{R},$$

auf Q, also gilt $\hat{y}^\top A\,\hat{y} + 2\,b^\top\,\hat{y} = 0$. Somit erhalten wir für alle $t \in \mathbb{R}$:

$$\frac{1}{4}\,s^2\,t^4\,\hat{v}_n^\top\,A\,\hat{v}_n + s\,t^3\,\hat{x}^\top A\,\hat{v}_n + t^2\,(\hat{x}^\top A\,\hat{x} + s\,b^\top\,\hat{v}_n) + 2\,t\,b^\top\,\hat{x} = 0\,.$$

Daraus folgt

$$a_{nn} = 0\;,\;\; \hat{x}^\top A\,\hat{v}_n = 0\;,\;\; (\hat{x}^\top A\,\hat{x} + s\,b^\top\,\hat{v}_n) = 0\;,\;\; b^\top\,\hat{x} = 0.$$

Ist $s = 0$, so liegt die Gerade OX ganz auf Q, woraus

$$t^2\,\hat{x}^\top A\,\hat{x} + 2\,t\,b^\top\,\hat{x} = 0$$

für alle $t \in \mathbb{R}$ folgt. Also muß $\hat{x}^\top A\,\hat{x} = b^\top\,\hat{x} = 0$ sein.

In beiden Fällen ist $b^\top\,\hat{x} = 0$, also ist $b = a\,v_n$, $a \neq 0$.

Wäre $\hat{x}^\top A\,\hat{v}_n \neq 0$ für $s = 0$, so würde

$$\hat{y} = \hat{v}_n + \frac{a}{\hat{x}^\top A\;\hat{v}_n}\,\hat{x}$$

die Gleichung $\hat{y}^\top A\,\hat{y} + 2\,b^\top\,\hat{y} = 0$ erfüllen, der zugehörige Punkt läge also auf Q. Dies kann aber nicht sein, da $\hat{y}$ nicht die Normalform von Q erfüllt, also ist auch in diesem Fall $\hat{x}^\top A\,\hat{v}_n = 0$.

Somit gilt für alle $\hat{x} = (x_1,\ldots,x_{n-1},0) \neq o$:

$$\hat{x}^\top A\,\hat{v}_n = 0 \quad\text{und}\quad \hat{x}^\top A\,\hat{x} = -a\,s = -a\,(x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_r^2)\,.$$

Ist schließlich $\hat{x} = (x_1,\ldots,x_n)$ beliebig, so schreiben wir $\hat{x}$ in der Form $\hat{x} = (x_1,\ldots,x_{n-1},0) + x_n\,\hat{v}_n$ und erhalten ebenfalls

$$\hat{x}^\top A\,\hat{x} = -a\,(x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_r^2)\,.$$

Somit unterscheiden sich die beiden definierenden Gleichungen von Q nur um einen Faktor $-a \neq 0$. ∎

§ 5 Quadriken in euklidischen Räumen

In einem euklidischen Raum spielt die affine Äquivalenz der Quadriken keine Rolle mehr, da bei affinen Abbildungen metrische Eigenschaften im allgemeinen verlorengehen; so ist zum Beispiel die Ellipse affin äquivalent zum Kreis und das Ellipsoid ist affin äquivalent zur Kugel. Da jetzt die Gestalt einer Quadrik unverändert bleiben soll, darf man unter den affinen Abbildungen nur noch die Bewegungen zulassen. Die affine Einteilung der Quadriken wird dadurch verfeinert. Wir erhalten innerhalb der Typen (I), (II), (III) eine weitere Aufteilung in Kongruenzklassen.

Die zugehörigen euklidischen Normalformen ergeben sich in der gleichen Weise wie beim Beweis von Satz 10. Wir müssen nur feststellen, welche der dort verwendeten Koordinatentransformationen Bewegungen entsprechen.

Satz 17 (Satz über die euklidische Hauptachsentransformation von Quadriken). *Es sei $\mathbb{A}$ ein n-dimensionaler euklidischer Raum. Dann läßt sich jede Quadrik $Q \subset \mathbb{A}$ bezüglich eines geeigneten kartesischen Koordinatensystems durch eine der folgenden Gleichungen beschreiben:*

(I) $$\frac{x_1^2}{a_1^2} + \cdots + \frac{x_p^2}{a_p^2} - \frac{x_{p+1}^2}{a_{p+1}^2} - \cdots - \frac{x_r^2}{a_r^2} = 0 \ , \qquad 0 < p \leq r \leq n \ , \ p - r \leq p \ ,$$

(II) $$\frac{x_1^2}{a_1^2} + \cdots + \frac{x_p^2}{a_p^2} - \frac{x_{p+1}^2}{a_{p+1}^2} - \cdots - \frac{x_r^2}{a_r^2} = 1 \ , \qquad 0 < p \leq r \leq n \ ,$$

(III) $$\frac{x_1^2}{a_1^2} + \cdots + \frac{x_p^2}{a_p^2} - \frac{x_{p+1}^2}{a_{p+1}^2} - \cdots - \frac{x_r^2}{a_r^2} = 2\,x_n \ , \qquad 0 < p \leq r < n, \ r-p \leq n \ ,$$

mit positiven reellen Zahlen $a_1, \ldots, a_r$.

Beweis. Die Quadrik Q sei durch die Gleichung $\beta(x,x) + 2\,\Phi(x) + c = 0$ gegeben. Wie im Beweis von Satz 10 beginnen wir mit dem quadratischen Anteil der Gleichung von Q, beschrieben durch die Bilinearform β. Nach Wahl einer ONB in V wird β durch eine symmetrische Matrix A beschrieben. Diese läßt sich wieder mit Hilfe einer orthogonalen Matrix auf Diagonalgestalt transformieren. In der Diagonalen stehen dann die Eigenwerte $c_1,\dots,c_n$ von A.

Für den Vektorraum V bedeutet dies die Existenz einer ONB $(v_1,\dots,v_n)$, so daß die Matrix $\overline{A} = (\!(\beta(v_i,v_j))\!)$ Diagonalgestalt hat mit $\beta(v_i,v_i) = c_i$, $i = 1,\dots,n$. Wir können o.B.d.A. annehmen, daß $c_1,\dots,c_p > 0$, $c_{p+1},\dots,c_r < 0$ und $c_{r+1} = \dots = c_n = 0$ gilt. Dabei ist $r = \operatorname{Rg} A = n - \dim \operatorname{Rad} \beta > 0$.

Indem wir nun wieder den Ursprung O durch eine Translation geeignet verschieben, erhalten wir ein kartesisches Koordinatensystem $(O'; v_1,\dots,v_n)$, bezüglich dessen die Gleichung von Q die Form

$$c_1\, y_1^2 + \dots + c_r\, y_r^2 + 2\, b_{r+1}\, y_{r+1} + \dots + 2\, b_n\, y_n + c = 0$$

annimmt mit geeigneten $b_{r+1},\dots,b_n$, einer neuen Konstanten c und Variablen $y_1,\dots y_n$.

1. Fall: $b_{r+1} = \dots = b_n = 0$. Dann besitzt Q bezüglich des zuletzt betrachteten kartesischen Koordinatensystems die Gleichung

$$c_1\, y_1^2 + \dots + c_r\, y_r^2 + c = 0\,.$$

2. Fall: Einer der Koeffizienten $b_{r+1},\dots,b_n$ ist von 0 verschieden. Dann ist $r < n$. Wir können o.E. annehmen, daß $b_n \neq 0$ gilt, andernfalls numerieren wir um. Jetzt wechseln wir die Basis in dem Untervektorraum $[v_{r+1},\dots,v_n]$ und konstruieren mit Hilfe des Schmidtschen Orthogonalisierungsverfahrens, ausgehend von dem Vektor

$$v_n' := -\frac{b_{r+1}}{b'}\, v_{r+1} - \dots - \frac{b_n}{b'}\, v_n$$

mit $b' := (b_{r+1}^2 + \dots + b_n^2)^{\frac{1}{2}} \neq 0$, eine weitere Orthonormalbasis $(v_{r+1}',\dots,v_n')$ von $[v_{r+1},\dots,v_n]$. Dann gilt für $i = r+1,\dots,n-1$

$$v'_i = a_{ir+1} v_{r+1} + \cdots + a_{in} v_n$$

und für $i = n$

$$v'_n = -\frac{b_{r+1}}{b'} v_{r+1} - \cdots - \frac{b_n}{b'} v_n .$$

Die zugehörige Koeffizientenmatrix ist orthogonal. Setzen wir noch $v'_i = v_i$ für $i = 1,\ldots,r$, so ist insgesamt $(v'_1,\ldots,v'_n)$ eine ONB von V und die Gleichung von Q hat bezüglich des neuen Koordinatensystems $(O'; v'_1,\ldots,v'_n)$ die Gestalt

$$c_1 z_1^2 + \cdots + c_r z_r^2 - 2 z_n + c = 0 .$$

Verschieben wir jetzt noch den Ursprung O' um den Vektor $\frac{c}{2} v'_n$, so erhalten wir schließlich die Gleichung

$$c_1 u_1^2 + \cdots + c_r u_r^2 = 2 u_n .$$

Um die Normalformen (I), (II) und (III) zu erhalten, setzen wir im 1. Fall (für $c = 0$) bzw. im 2. Fall

$$a_i = \sqrt{1/c_i} \quad \text{für } 1 \leq i \leq p \quad \text{und} \quad a_i = \sqrt{-1/c_i} \quad \text{für } p + 1 \leq i \leq r$$

und erhalten

$$\text{(I)} \qquad \frac{y_1^2}{a_1^2} + \cdots + \frac{y_p^2}{a_p^2} - \frac{y_{p+1}^2}{a_{p+1}^2} - \cdots - \frac{y_r^2}{a_r^2} = 0 , \qquad 0 < p \leq r \leq n ,$$

bzw.

$$\text{(III)} \qquad \frac{u_1^2}{a_1^2} + \cdots + \frac{u_p^2}{a_p^2} - \frac{u_{p+1}^2}{a_{p+1}^2} - \cdots - \frac{u_r^2}{a_r^2} = 2 u_n , \qquad 0 < p \leq r < n ,$$

wobei wir $p - r \leq p$ annehmen können.

Ist im 1. Fall $c \neq 0$, so können wir wieder annehmen, daß $c < 0$ gilt. Wir dividieren durch $-c$, setzen

$$a_i = (-c / |c_i|)^{\frac{1}{2}} \quad \text{für } i = 1,\ldots,r$$

und erhalten die Normalform (II). ∎

Bemerkungen und Bezeichnungen. Die Zahlen $a_1,\dots,a_r$ sind in den Fällen (II) und (III) eindeutig bestimmt, im Fall (I) erreichen wir die Eindeutigkeit durch die Normierung $a_r = 1$. Durch Vertauschen der Koordinatenachsen können wir sogar noch erreichen, daß

$$a_1 \geq a_2 \geq \cdots \geq a_p > 0 ,$$

$$0 < a_{p+1} \leq a_{p+2} \leq \cdots \leq a_r$$

gilt. Die Koordinatenachsen des kartesischen Koordinatensystems, für das die Gleichung von Q Normalform hat, heißen die *Hauptachsen* von Q. Im Fall (II) geben die Zahlen $2\,a_i$ die *Hauptachsenlängen* an.

Für $n = 2,3$ ergeben sich aus Satz 16 die folgenden euklidischen Normalformen von Quadriken.

$\boldsymbol{n = 2}$: In der euklidischen Ebene gibt es fünf Klassen von eigentlichen Kegelschnitten.

(I) $\dfrac{x_1^2}{a_1^2} - \dfrac{x_2^2}{a_2^2} = 0$, (*Paar sich schneidender Geraden*)

(II) $\dfrac{x_1^2}{a_1^2} + \dfrac{x_2^2}{a_2^2} = 1$, (*Ellipse, Kreis für* $a_1 = a_2$)

$\dfrac{x_1^2}{a_1^2} - \dfrac{x_2^2}{a_2^2} = 1$, (*Hyperbel*)

$\dfrac{x_1^2}{a_1^2} = 1$, (*Paar paralleler Geraden*)

(III) $\dfrac{x_1^2}{a_1^2} = 2\,x_2$. (*Parabel*)

$n = 3$: Im dreidimensionalen euklidischen Raum gibt es elf Klassen von Quadriken, die keine Punkte, Geraden oder Ebenen sind.

(I) $$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} - \frac{x_3^2}{a_3^2} = 0\,, \qquad \textit{(Kegel)}$$

$$\frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} = 0\,, \qquad \textit{(Paar sich schneidender Ebenen)}$$

(II) $$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} + \frac{x_3^2}{a_3^2} = 1\,, \qquad \textit{(Ellipsoid, Kugel für } a_1 = a_2 = a_3\textit{)}$$

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} - \frac{x_3^2}{a_3^2} = 1\,, \qquad \textit{(einschaliges Hyperboloid)}$$

$$\frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} - \frac{x_3^2}{a_3^2} = 1\,, \qquad \textit{(zweischaliges Hyperboloid)}$$

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 1\,, \qquad \textit{(elliptischer Zylinder)}$$

$$\frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} = 1\,, \qquad \textit{(hyperbolischer Zylinder)}$$

$$\frac{x_1^2}{a_1^2} = 1\,, \qquad \textit{(Paar paralleler Ebenen)}$$

(III) $$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 2\,x_3\,, \qquad \textit{(elliptisches Paraboloid)}$$

$$\frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} = 2\,x_3\,, \qquad \textit{(hyperbolisches Paraboloid)}$$

$$\frac{x_1^2}{a_1^2} = 2\,x_3 \ . \qquad\qquad \textit{(parabolischer Zylinder)}$$

Beispiel. In einem dreidimensionalen euklidischen Raum sei bezüglich eines kartesischen Koordinatensystems eine Quadrik Q durch die Gleichung

$$8\,x_1^2 + 4\,x_1\,x_2 + 4\,x_1\,x_3 + 5\,x_2^2 - 8\,x_2\,x_3 + 5\,x_3^2 + 54\,x_1 + 72 = 0$$

gegeben. Wir wollen die euklidische Normalform von Q bestimmen.

Die Gleichung von Q lautet in Matrixschreibweise

$$\hat{x}^{\mathrm{T}} A\,\hat{x} + 2\,b^{\mathrm{T}}\,\hat{x} + c = 0$$

mit

$$A = \begin{bmatrix} 8 & 2 & 2 \\ 2 & 5 & -4 \\ 2 & -4 & 5 \end{bmatrix}, \ b = \begin{bmatrix} 27 \\ 0 \\ 0 \end{bmatrix}, \ c = 72 \ .$$

Das charakteristische Polynom von A ist $p = -\,X\,(9 - X)^2$. Die Eigenwerte von A sind somit $c_1 = c_2 = 9$, $c_3 = 0$.

In den Eigenräumen E_{c_1} und E_{c_3} bestimmen wir jeweils eine ONB. Für E_{c_1} bilden die Vektoren

$$\hat{v}_1 = \frac{1}{3}\begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}, \ \hat{v}_2 = \frac{1}{3}\begin{bmatrix} 2 \\ -1 \\ 2 \end{bmatrix},$$

in E_{c_3} der Vektor

$$\hat{v}_3 = \frac{1}{3}\begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}$$

eine ONB. Dann ist $S = (\hat{v}_1 | \hat{v}_2 | \hat{v}_3)$ eine orthogonale Matrix, die A auf Diagonalgestalt transformiert.

In den neuen Koordinaten y_1, y_2, y_3 , die mit den alten Koordinaten x_1, x_2, x_3 durch die Gleichungen

$$x_1 = \tfrac{2}{3}\,y_1 + \tfrac{2}{3}\,y_2 - \tfrac{1}{3}\,y_3\,,$$

$$x_2 = \tfrac{2}{3}\,y_1 - \tfrac{1}{3}\,y_2 + \tfrac{2}{3}\,y_3\,,$$

$$x_3 = -\tfrac{1}{3}\,y_1 + \tfrac{2}{3}\,y_2 + \tfrac{2}{3}\,y_3$$

verbunden sind, hat Q somit die Gleichung

$$9\,(y_1^2 + y_2^2 + 4\,y_1 + 4\,y_2 - 2\,y_3 + 8) = 0\,.$$

Mit Hilfe der Translation

$$z_1 = y_1 + 2\;,\; z_2 = y_2 + 2\;,\; z_3 = y_3$$

erhalten wir die euklidische Normalform der Gleichung von Q :

$$z_1^2 + z_2^2 = 2\,z_3\,.$$

Q ist also ein elliptisches Paraboloid.

Euklidische Klassifikation der Quadriken

Im Gegensatz zu den affinen Normalformen, gibt es im euklidischen Fall wegen der Nebenbedingungen an die Hauptachsenlängen unendlich viele Typen in jeder Klasse. Außerdem können uneigentliche Quadriken, d.h. euklidische Unterräume, durch verschiedene Normalformen der Klasse (I) dargestellt werden. So beschreiben die Normalformen

$$x_1^2 + x_2^2 = 0 \quad , \quad \frac{x_1^2}{2} + x_2^2 = 0$$

bezüglich eines festen kartesischen Koordinatensystems im $\mathbb{R}^3$ dieselbe Gerade.

Beschränken wir uns auf eigentliche Quadriken, so gilt analog zu Satz 15 der folgende Satz.

Satz 18. *In einem euklidischen Raum sind zwei eigentliche Quadriken genau dann kongruent, wenn ihre euklidischen Normalformen vom gleichen Typ* (I), (II) *oder* (III) *sind und außerdem in den Zahlen p, r und a_i übereinstimmen, wobei im Fall* (I) $a_r = 1$ *gilt.*

Beweis. Besitzen zwei Quadriken Q und $\tilde{Q}$ dieselbe Normalform, so stimmen auch die Größen a_i überein, also bildet die Bewegung φ, die die zugehörigen kartesischen Koordinatensysteme aufeinander abbildet, auch Q auf $\tilde{Q}$ ab oder umgekehrt $\tilde{Q}$ auf Q.

Gilt $\varphi(Q) = \tilde{Q}$ mit einer Bewegung φ, so gehören Q und $\tilde{Q}$ nach Satz 15 zunächst zur selben Klasse (I), (II) oder (III) und stimmen auch bezüglich p und r überein. Es bleibt also nur noch zu zeigen, daß auch die Größen a_i und $\tilde{a}_i$ gleich sind. Dies ist aber nach Definition von a_i und $\tilde{a}_i$ genau dann der Fall, wenn die Eigenwerte c_i und $\tilde{c}_i$ der zu β und $\tilde{\beta}$ gehörenden Matrizen A und $\tilde{A}$ dieselben sind. Nach S. 306 gilt $\tilde{A} = C^T A\ C$, wobei jetzt C sogar orthogonal ist, also sind A und $\tilde{A}$ orthogonal äquivalent und besitzen somit dieselben Eigenwerte. ∎

§ 6 Tangenten und Tangentialhyperebenen von Quadriken

In diesem letzten Abschnitt wollen wir die geometrischen Eigenschaften der Quadriken noch etwas genauer studieren. Hierzu untersuchen wir zunächst die möglichen Lagen, die eine Gerade in bezug auf eine Quadrik Q einnehmen kann. Da der Fall, daß Q ein affiner Unterraum ist, schon früher behandelt worden ist, können wir hier ohne Einschränkung annehmen, daß Q eine eigentliche Quadrik ist. Der zugrunde liegende affine Raum kann auch euklidisch sein.

Eine Gerade g hat entweder mit der Quadrik Q keinen, einen oder zwei Punkte gemeinsam, oder sie ist ganz auf Q gelegen. Alle Fälle sind möglich, wie man an den Kegelschnitten im $\mathbb{R}^2$ sieht.

Die Schnittmenge $g \cap Q$ läßt sich explizit bestimmen, indem man eine Parameterdarstellung $x = x_0 + t\,y$, $t \in \mathbb{R}$, von g in die definierende Gleichung

$$\beta(x,x) + 2\,\Phi(x) + c = 0$$

von Q einsetzt. Man erhält so für $g \cap Q$ eine quadratische Gleichung in t,

$$(*) \qquad t^2\beta(y,y) + 2\,t\,\{\beta(x_0,y) + \Phi(y)\} + \beta(x_0,x_0) + 2\,\Phi(x_0) + c = 0\,,$$

die entweder keine oder eine oder zwei oder unendlich viele Lösungen besitzt.

Der Fall $g \cap Q \neq \emptyset$

Ist $g \cap Q \neq \emptyset$, so können wir die Parameterdarstellung von g so wählen, daß der Punkt X_0 mit dem Ortsvektor x_0 auf Q liegt. Dann nimmt die Schnittgleichung (*) die folgende einfache Form an

$$t^2\beta(y,y) + 2\,t\,\{\beta(x_0,y) + \Phi(y)\} = 0\,.$$

Für $\beta(x_0,y) + \Phi(y) \neq 0$ und $\beta(y,y) \neq 0$ gibt es zwei Lösungen. Die Gerade g hat dann mit Q genau zwei Punkte gemeinsam, sie ist eine *Sekante*.

Für $\beta(x_0,y) + \Phi(y) \neq 0$ und $\beta(y,y) = 0$ gibt es nur die Lösung $t = 0$ mit zugehörigem Punkt X_0. In diesem Fall heißt g eine *Einfach-Schneidende.*

Für $\beta(x_0,y) + \Phi(y) = 0$ und $\beta(y,y) \neq 0$ ist wiederum $t = 0$ die einzige Lösung (jetzt aber mit Vielfachheit 2). In diesem Fall heißt g eine *Tangente* an Q und der gemeinsame Punkt X_0 heißt der *Berührpunkt* der Tangente.

Für $\beta(x_0,y) + \Phi(y) = 0$ und $\beta(y,y) = 0$ erfüllt jedes $t \in \mathbb{R}$ die Schnittgleichung. Dies bedeutet, daß g ganz auf Q liegt. Wir nennen die Gerade g dann eine *Erzeugende* von Q.

Bemerkung. Diese Definitionen sind von der speziellen Darstellung von Q unabhängig, da Q eine eigentliche Quadrik ist und die Gleichungen von Q sich dann nach Satz 16 nur um einen von Null verschiedenen Faktor unterscheiden.

Nun betrachten wir die Vereinigung $T_{X_0}(Q)$ aller Tangenten an Q mit Berührpunkt X_0 und aller Erzeugenden von Q, die X_0 enthalten. Dann gilt:

$X \in T_{X_0}(Q)$

$\Longleftrightarrow$ X liegt auf einer Tangente oder auf einer Erzeugenden durch X_0

$\Longleftrightarrow$ Der Ortsvektor $x = \overrightarrow{OX}$ erfüllt die lineare Gleichung

$$\beta(x_0, x - x_0) + \Phi(x - x_0) = 0,$$

die wegen $X_0 \in Q$ äquivalent ist zu der linearen Gleichung

$$(**) \qquad \beta(x_0,x) + \Phi(x) + \Phi(x_0) + c = 0\,.$$

Also ist $T_{X_0}(Q)$ ein affiner (euklidischer) Unterraum. Er heißt der *Tangentialraum* von Q im Punkt X_0.

Ist die Linearform $\beta(x_0,\cdot) + \Phi$ in (**) nicht die Nullform, so ist der Tangentialraum eine Hyperebene. Sie heißt die *Tangentialhyperebene* von Q in X_0 und der Punkt X_0 heißt ein *regulärer Punkt* der Quadrik Q.

Ist die Linearform $\beta(x_0,\cdot) + \Phi$ in (**) die Nullform, so ist der Tangentialraum $T_{X_0}(Q)$ der gesamte Raum $\mathbb{A}$, denn dann ist $\Phi(x_0) + c = \beta(x_0,x_0) + \Phi(x_0) + \Phi(x_0) + c = 0$, und jeder Vektor $x \in V$ erfüllt die Gleichung (**). Der Punkt X_0 wird dann *singulärer Punkt* von Q genannt.

Singuläre Punkte können nur bei Kegeln auftreten, es sind die Spitzen der Kegel. Quadriken vom Typ (II) oder (III) dagegen besitzen nur reguläre Punkte.

Beispiele. **(a)** Wir betrachten in der euklidischen Ebene eine Ellipse Q mit der Normalform

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} = 1 .$$

Dann ist

$$\beta(x,y) = \frac{x_1\, y_1}{a_1^2} + \frac{x_2\, y_2}{a_2^2} , \quad \Phi = O , \quad c = -1 .$$

Die Gleichung der Tangente im Punkt $X' \in Q$ lautet somit

$$\frac{x_1'\, x_1}{a_1^2} + \frac{x_2'\, x_2}{a_2^2} = 1 .$$

(b) Q sei eine Parabel in der euklidischen Ebene mit der Normalform

$$\frac{x_1^{\,2}}{a_1^2} = 2\, x_2 .$$

Dann gilt

$$\beta(x,y) = \frac{x_1\, y_1}{a_1^2} , \quad \Phi(x) = -\, x_2 , \quad c = 0 ,$$

und die Gleichung der Tangente in $X' \in Q$ ist

$$\frac{x_1' \, x_1}{a_1^2} = x_2 + x_2' .$$

Der Fall $g \cap Q = \emptyset$

Hier darf die Schnittgleichung keine Lösung haben. Für $\beta(y,y) \neq 0$ ist dies nur möglich, wenn die Diskriminante der quadratischen Gleichung (*) negativ ist. In diesem Fall nennen wir die Gerade g eine *Passante.*

Für $\beta(y,y) = 0$ muß notwendigerweise

$$\beta(x_0,y) + \Phi(y) = 0 \text{ und } \beta(x_0,x_0) + 2\,\Phi(x_0) + c \neq 0$$

gelten. Dies ist dann auch hinreichend dafür, daß g und Q keinen Punkt gemeinsam haben. Die Gerade g erfüllt dann die Tangentenbedingung $\beta(x_0,y) + \Phi(y) = 0$, hat aber mit Q keinen Punkt gemeinsam; sie heißt *Asymptote* von Q. Asymptoten haben die Eigenschaft, daß im euklidischen Raum ihr Abstand zur Quadrik beliebig klein wird. Wir verzichten hier auf einen Beweis dieser Grenzwerteigenschaft.

Ist Q eine echte Mittelpunktsquadrik und ist X_0 ein Mittelpunkt von Q, so ist $\beta(x_0,\cdot) + \Phi$ die Nullform, also ist eine Gerade $X_0 Y$ genau dann eine Asymptote von Q, wenn $\beta(\overrightarrow{X_0 Y}, \overrightarrow{X_0 Y}) = 0$ gilt. Die Vereinigungsmenge aller Asymptoten durch X_0 ist also selbst wieder eine Quadrik, nämlich ein Kegel. Er wird auch *Asymptotenkegel* von Q im Punkt X_0 genannt. Gibt es nur einen Mittelpunkt von Q, so spricht man von *dem* Asymptotenkegel der Quadrik Q.

Beispiel. In der euklidischen Ebene besitzt die Hyperbel Q mit der Normalform

$$\frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2} = 1$$

den Asymptotenkegel

$$[\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}] \cup [\begin{bmatrix} a_1 \\ -a_2 \end{bmatrix}] .$$

Bemerkung. Die Gleichung (**) von S. 339 ist auch für $X_0 \notin Q$ sinnvoll. Sie beschreibt genau dann eine Hyperebene in $\mathbb{A}$, wenn X_0 kein Mittelpunkt von Q ist.

Definition. Es seien $Q = \{X \mid \overrightarrow{OX} = x \, , \, \beta(x,x) + 2\,\Phi(x) + c = 0\}$ eine eigentliche Quadrik und X_0 kein Mittelpunkt von Q. Dann heißt die Hyperebene

$$\{X \mid \overrightarrow{OX} = x \, , \, \beta(x_0,x) + \Phi(x) + \Phi(x_0) + c = 0\}$$

die *Polarhyperebene* von X_0 bezüglich Q und X_0 heißt der *Pol* dieser Hyperebene.

Für $X_0 \in Q$ ist die Polarhyperebene auch Tangentialhyperebene und der Pol X_0 ist der Berührpunkt. Hat die Polarhyperebene von X_0 mit der Quadrik Q einen nichtleeren Schnitt, so ist jede Verbindungsgerade von X_0 mit einem Punkt Y aus dieser Schnittmenge eine Tangente an Q mit Berührpunkt Y.

Satz 19 (Hauptsatz der Polarentheorie). *Es sei Q eine eigentliche Quadrik und die Punkte $X, Y \in \mathbb{A}$ seien keine Mittelpunkte von Q. Dann gilt:*

X liegt genau dann in der Polarhyperebene von Y, wenn Y in der Polarhyperebene von X liegt.

Beweis. Der Beweis ist offensichtlich, da die Gleichung (**) symmetrisch in x_0 und x ist. ∎

Literaturverzeichnis

Artmann, B. : Lineare Algebra. Birkhäuser Skripten, 1989 (2. Auflage).

Beiglböck, W.D. : Lineare Algebra. Springer, 1983.

Beutelspacher, A. : Lineare Algebra. Vieweg, 2003 (6., durchgesehene Auflage).

Bosch, S. : Lineare Algebra. Springer, 2006 (3. Auflage).

Brieskorn, E. : Lineare Algebra und anlytische Geometrie I,II. Vieweg, 1983,1985.

Bröcker, Th. : Lineare Algebra und anlytische Geometrie. Birkhäuser, 2004 (2., korrigierte Auflage).

Fischer, G. : Lineare Algebra. Vieweg, 2005 (15., verbesserte Auflage).

Fischer, G. : Analytische Geometrie. Vieweg, 2001 (7., durchgesehene Auflage).

Grauert, H. , Grunau, H.-C. : Lineare Algebra und analytische Geometrie. Oldenbourg, 1999.

Greub, W. : Lineare Algebra. Springer, Heidelberger TB, 1976.

Heinhold, J. , Riedmüller, B. : Lineare Algebra und Analytische Geometrie I,II. Hanser, 1973, 1980 (1. bzw. 3. Auflage).

Heinhold, J. , Riedmüller, B. : Aufgaben und Lösungen zur Linearen Algebra und Analytischen Geometrie I,II. Hanser, 1977 (2. bzw. 3. Auflage).

Jänich, K. : Lineare Algebra. Springer, 2004 (10. Auflage).

Klingenberg, W. : Lineare Algebra und Geometrie. Springer, 1992 (3. Auflage).

Koecher, M. : Lineare Algebra und Analytische Geometrie. Springer, 1997 (4., erg. u. aktualisierte Aufllage).

Kowalsky, H.-J. , Michler, G.O. : Lineare Algebra. de Gruyter, 1998 (11. Auflage).

Lamprecht, E. : Lineare Algebra I,II. Birkhäuser, Unitaschenbücher, 1980, 1983.

Lorenz, F. : Lineare Algebra I,II. BI-Hochschultaschenbuch, 1992 (3. Auflage).

Schaal, H. : Lineare Algebra und Analytische Geometrie I,II. Vieweg, 1976.

Schaal, H. , Glässner, E. : Lineare Algebra und Analytische Geometrie III. Aufgaben mit Lösungen. Vieweg, 1981 (2. Auflage).

Storch, U. , Wiebe, H. : Lehrbuch der Mathematik, Bd II: Lineare Algebra. BI, 1990.

Strang, G. : Lineare Algebra. Springer, 2003.

Stroth, G. : Lineare Algebra. Heldermann, 1995

Wagner, R. : Grundzüge der linearen Algebra. Teubner, MLG-Reihe, 1981.

Walter, R. : Einführung in die lineare Algebra. Vieweg, 1990 (3., verb. Auflage).

Walter, R. : Lineare Algebra und analytische Geometrie. Vieweg, 1993 (2., durchgesehene Auflage).

Symbolverzeichnis

Neben den üblichen lateinischen Buchstaben werden auch Buchstaben anderer Schriften verwendet.

Griechische Buchstaben:

β (Beta) , δ und Δ (Delta) , φ und Φ (Phi) , Ψ (Psi) ,π (Pi) , σ (Sigma) , τ (Tau) , ω (Omega).

Skriptbuchstaben: $\mathscr{M}$, $\mathscr{P}$, $\mathscr{Q}$.

Stichwortverzeichnis

K

L

M

N

O

P

Q

U

V

W

Z