Teil 5. Selbstadjungierte Endomorphismen. Quadriken.

5A. Selbstadjungierte Endomorphismen.

Sei V ein euklid'scher oder unitärer Vektorräum mit innerem Produkt $\langle -,-\rangle$. Ein Endomorphismus $f\colon V\to V$ heißt *selbstadjungiert,* falls gilt

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

Ist V ein euklid'scher Vektorraum mit einer Orthonormalbasis $v_1,\dots,v_n$, und ist C die Matrizendarstellung eines Endomorphismus $f\colon V\to V$ bezüglich dieser Basis, so ist f *genau dann selbstadjungiert, wenn C eine symmetrische Matrix ist* (wenn also ${}^tC = C$ gilt).

Ist V ein unitärer Vektorraum mit einer Orthonormalbasis $v_1,\dots,v_n$, und ist C die Matrizendarstellung eines Endomorphismus $f\colon V\to V$ bezüglich dieser Basis, so ist f *genau dann selbstadjungiert, wenn C eine hermite'sche Matrix ist* (wenn also ${}^tC = \overline{C}$ gilt).

Beweis im unitären Fall: Sei $C=(c_{ij})_{ij}$. Es ist $f(v_r)=\sum_i c_{ir}v_i$, und $f(v_s)=\sum_i c_{is}v_i$, also

$$c_{sr} = \langle \sum_i c_{ir}v_i, v_s\rangle = \langle f(v_r),v_s\rangle, \quad \text{und} \quad \overline{c}_{rs} = \langle v_r, \sum_i c_{is}v_i\rangle = \langle v_r, f(v_s)\rangle.$$

Ist f selbstadjungiert, so stimmen die rechten Seiten überein, also ist $c_{sr}=\overline{c}_{rs}$ für $1\le r,s\le n$. Ist umgekehrt C hermite'sch, so stimmen die linken Seiten überein, also auch die rechten: Es ist also $\langle f(v_r),v_s\rangle = \langle v_r,f(v_s)\rangle$ für alle $1\le r,s\le n$, und dies überträgt sich auch auf Linearkombinationen, also ist f selbstadjungiert.

Satz (1. Formulierung). *Sei V ein endlich-dimensionaler euklid'scher oder unitärer Vektorraum. Sei $f\colon V\to V$ selbstadjungierter Endomorphismus. Dann gibt es eine Orthonormalbasis von V aus Eigenvektoren zu f und alle Eigenwerte sind reell.*

Satz (2. Formulierung). *Ist C eine symmetrische reelle $(n\times n)$-Matrix, so gibt es eine orthogonale Matrix P, so dass $P^{-1}CP$ eine Diagonalmatrix ist.*

Ist C eine hermite'sche (komplexe) $(n\times n)$-Matrix, so gibt es eine unitäre Matrix P, so dass $P^{-1}CP$ eine Diagonalmatrix mit reellen Koeffizienten ist.

Zusatz. Man kann zusätzlich verlangen, dass $\det P = 1$ gilt. (Im reellen Fall muss man dazu gegebenenfalls zwei Spalten vertauschen; im komplexen Fall multipliziert man P mit der Skalarmatrix mit Skalar $(\det P)^{-1/n}$.)

Beweis. Die beiden Formulierungen besagen das Gleiche. Es genügt also, eine der beiden Formulierungen zu beweisen (wir werden dabei aber hin- und herspringen).

(1) *Alle Eigenwerte einer hermite'schen Matrix sind reell.* Beweis: Sei C hermite'sche $(n\times n)$-Matrix, sei $v\in\mathbb{C}^n$ ein Eigenvektor von f_C mit Eigenwert γ, also $f_C(v)=\gamma v$. Sei $\langle -,-\rangle$ das kanonische innere Produkt auf $\mathbb{C}$.

$$\gamma\langle v,v\rangle = \langle \gamma v,v\rangle = \langle f_C(v),v\rangle = \langle v,f_C(v)\rangle = \langle v,\gamma v\rangle = \overline{\gamma}\langle v,v\rangle.$$

Da v nicht der Nullvektor ist, ist $\langle v, v\rangle \neq 0$, also ist $\gamma = \overline{\gamma}$, und demnach ist γ reell.

(2) Jede reelle Matrix kann man auch als komplexe Matrix auffassen, ist C symmetrische reelle Matrix, so ist C auch hermite'sch. Aus (1) folgt: *Alle komplexen Eigenwerte einer reellen symmetrischen Matrix C sind reell* (das charakteristische Polynom $\chi_C(T)$ hat nur reelle Nullstellen).

(3) *Sei V ein euklid'scher oder unitärer Vektorraum und $f\colon V \to V$ ein selbstadjungierter Endomorphismus. Ist U ein f-invarianter Unterraum von V, so ist auch $U^\perp$ f-invariant.* Beweis: Sei $w \in U^\perp$. Zu zeigen ist $f(w) \in U^\perp$. Ist $u \in U$, so ist

$$\langle f(w), u\rangle = \langle w, f(u)\rangle = 0,$$

denn $w \in U^\perp$ und $u \in U$. Also ist $f(w) \in U^\perp$.

(4) Beweis des Satzes mit Induktion. Sei V ein euklid'scher oder unitärer Vektorraum und $f\colon V \to V$ ein selbstadjungierter Endomorphismus. Wähle einen Eigenvektor v_1 von f. Wir können annehmen, dass $||v_1|| = 1$ gilt. Setze $V' = (\operatorname{span}\{v_1\})^\perp$. Wegen (3) wird V' unter f in sich abgebildet und die Einschränkung von f auf V' ist natürlich ebenfalls ein selbstadjungierter Endomorphismus. Induktion zeigt, dass es in V' eine Orthonormalbasis aus Eigenvektoren gibt, etwa $v_2, \dots, v_n$. Dann ist $v_1, \dots, v_n$ eine Orthonormalbasis von V aus Eigenvektoren.

Folgerung. *Ist A eine reelle symmetrische $(n \times n)$-Matrix, und sind $\gamma \neq \gamma'$ Eigenwerte von A, so sind die Eigenräume* $\operatorname{Eig}(A, \gamma)$ *und* $\operatorname{Eig}(A, \gamma')$ *zueinander orthogonal* (bezüglich des kanonischen inneren Produkts des $\mathbb{R}^n$).

Hier noch einmal der **Algorithmus,** wie man zur symmetrischen Matrix C eine orthogonale Matrix P findet mit $P^{-1}CP$:

- Bestimme des charakteristische Polynom χ_C der Matrix C.
  [Hier ist also eine Determinante zu berechnen.]
- Bestimme die Nullstellen $\gamma_1, \dots, \gamma_t$ von χ_C (auf diese Weise erhält man die Eigenwerte von C).
  [Dies ist meist der schwierigste Schritt. Die Theorie besagt, dass χ_C Produkt von Linearfaktoren ist — sie zu berechnen kann aber schwierig sein.]
- Zu jedem γ_i bestimmt man eine Basis $v_{i1}, \dots, v_{i,j_i}$ von $\operatorname{Eig}(C, \gamma_i)$.
  [Hier handelt es scih um das Lösen eines linearen Gleichungssystems, nämlich dem homogenene Gleichungssystem mit Koeffizientenmatrix $C - \gamma_i I_n$. Dafür gibt es zum Beispiel den Gauß-Algorithmus.]
- Orthonormiere die Folge $v_{i1}, \dots, v_{i,j_i}$, man erhält auf diese Weise eine Orthonormalbasis $w_{i1}, \dots, w_{i,j_i}$ von $\operatorname{Eig}(C, \gamma_i)$.
  [Hierfür gibt es das Gram-Schmidt-Verfahren.]
- Die Matrix P hat als Spalten die Vektoren

$$w_{11}, \dots, w_{1,j_1}, w_{21}, \dots, w_{2,j_2}, \quad \dots \quad , w_{t1}, \dots, w_{t,j_t}.w_{i1}, \dots, w_{i,j_i}$$

(Beachte: Diese Folge ist eine Orthonormalbasis von $\mathbb{R}^n$, daher ist P orthogonale Matrix. Will man zusätzlich haben, dass $\det P$ gilt, so muss man gegebenfalls zwei Spalten vertauschen.)

Bemerkung: Ist C eine symmetrische reelle $(n \times n)$-Matrix und P eine orthogonale $(n \times n)$-Matrix, so ist $P^{-1}SP$ wieder symmetrisch. Ist C eine hermite'sche $(n \times n)$-Matrix und P eine unitäre $(n \times n)$-Matrix, so ist $P^{-1}SP$ wieder hermite'sch.

Es gilt also: *Die Menge der symmetrischen Matrizen in $M(n \times n, \mathbb{R})$ ist die Menge der Matrizen der Form $P^{-1}DP$, wobei P eine orthogonale Matrix und D eine reelle Diagonalmatrix ist.*

Die Menge der hermite'schen $(n \times n)$-Matrizen ist die Menge der Matrizen der Form $P^{-1}DP$, wobei P eine unitäre Matrix und D eine reelle Diagonalmatrix ist.

Folgerung. *Die einzige Matrix in $M(n \times n, \mathbb{R})$, die symmetrisch und nilpotent ist, ist die Nullmatrix.*

Beweis: Sei $A \in M(n \times n, \mathbb{R})$ symmetrisch und nilpotent. Da A symmetrisch ist, gibt es eine orthogonale Matrix P, sodass $D = P^{-1}AP$ eine Diagonalmatrix ist. Mit A ist auch $P^{-1}AP$ symmetrisch. Andererseits ist mit A auch $P^{-1}AP$ nilpotent. Es ist also D eine symmetrische nilpotente Diagonalmatrix. Also ist D die Nullmatrix. Dann ist aber auch $A = PDP^{-1} = 0$.

Wir haben gesehen, dass symmetrische Matrizen in $M(n \times n, \mathbb{R})$ sehr schöne Eigenschaften haben:

- sie sind diagonalisierbar,
- die einzige derartige Matrix, die nilpotent ist, ist die Nullmatrix,

und so weiter. Hier muss man betonen, dass dies nur für symmetrische **reelle** Matrizen gilt. Ist K ein anderer Körper, so haben symmetrische Matrizen $A \in M(n \times n, K)$ ganz andere Eigenschaften. So gibt es für $K = \mathbb{Z}/2\mathbb{Z}$ wie auch für $K = \mathbb{C}$ nilpotente syymetrische Matrizen $A \neq 0$, nämlich

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{für} \quad K = \mathbb{Z}/2\mathbb{Z}, \qquad \text{und} \quad \begin{bmatrix} 1 & i \\ i & -1 \end{bmatrix} \quad \text{für} \quad K = \mathbb{C}.$$

Einschub: Isotrope und anisotrope Vektoren.

Sei K ein Körper der Charakteristik $\mathrm{char}(K) \neq 2$. Sei $\langle -, - \rangle$ eine symmetrische Bilinearform auf dem K-Vektorraum V. Man nennt einen Vektor $v \in V$ *isotrop,* falls gilt $\langle v, v \rangle = 0$, andernfalls *anisotrop.* Ein Unterraum U, der nur aus isotropen Vektoren besteht, heißt *total-isotrop.*

5B. Hauptachsentransformation.

Wir betrachten nun den Fall eines euklid'schen Vektorraums V. Insbesondere ist also $K = \mathbb{R}$. Wir betrachten symmetrische Bilinearformen, die nicht notwendig positiv definit sind. Eine typische derartige Bilinearform, die in der Physik eine Rolle spielt ist die Lorentz-Form $x_1y_1 + x_2y_2 + x_3y_3 - cx_4y_4$ auf dem $\mathbb{R}^4$ (Raum-Zeit), dabei ist c die Lichtgeschwindigkeit. Darstellende Matrix (bezüglich der kanonischen Basis e_1, e_2, e_3, e_4) ist

$$\begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -c \end{bmatrix}$$

Hier gibt es viele isotrope Vektoren, zum Beispiel $e_1 - \frac{\sqrt{c}}{c} e_4$ (die Menge dieser isotropen Vektoren bildet den sogenannten "Lichtkegel").

Satz (Hauptachsentransformation). *Sei $(V, \langle -,- \rangle)$ ein euklid'scher Vektorraum. Sei $\langle -,- \rangle_A$ eine symmetrische Bilinearform auf V. Es gibt eine Orthonormalbasis von $(V, \langle -,- \rangle)$, die eine Orthogonalbasis für $\langle -,- \rangle_A$ ist.*

Man nennt eine derartige Orthonormalbasis für $\langle -,- \rangle_A$ (oder die von diesen Basisvektoren erzeugten Geraden) *Hauptachsen* der Bilinearform (oder auch Hauptachsen "der zugehörigen quadratischen Form").

Beweis: Die Aussage ist nur eine Umformulierung der Tatsache, dass es zu jeder symmetrischen Matrix A eine orthogonale Matrix P und eine Diagonalmatrix D mit

$$\boxed{P^{-1} A P = D}$$

gibt, hier allerdings wird diese Gleichung (wegen $P^{-1} = {}^tP$) in der Form

$$\boxed{{}^tP A P = D}$$

gelesen.

Zusatz. *Es gibt dann auch eine Orthogonalbasis von $(V, \langle -,- \rangle)$, die ebenfalls eine Orthogonalbasis für $\langle -,- \rangle_A$ ist, und für die zusätzlich gilt $\langle v_i, v_i \rangle_A \in \{1, -1, 0\}$.*

Beweis: Ist eine Orthonormalbasis $v_1, \dots, v_n$ für $(V, \langle -,- \rangle)$ gegeben, die eine Orthogonalbasis für $\langle -,- \rangle_A$ ist, und ist $d_i = \langle v_i, v_i \rangle$, so ersetzen wir im Fall $d_i \neq 0$ den Vektor v_i durch sein Vielfaches $\frac{\sqrt{|d_i|}}{d_i} v_i$. Die neue Basis ist für $\langle -,- \rangle$ nur noch eine Orthogonalbasis, andererseits ist nun $\langle v_i, v_i \rangle_A \in \{1, -1, 0\}$.

Bemerkung. Die hier gewählten Formulierungen betonen, dass es sich um **zwei** symmetrische Bilinearformen handelt (eine davon die kanonische), die gleichzeitig zu betrachten sind und die in Normalform gebracht werden sollen.

5C. Der Trägheitssatz von Sylvester.

Definition: Sei $(-,-)$ eine symmetrische Bilinearform auf dem reellen Vektorraum V. Man nennt $(-,-)$

- *positiv definit*, falls $(v,v) > 0$ für alle $0 \neq v \in V$ gilt,
- *positiv semidefinit*, falls $(v,v) \geq 0$ für alle $v \in V$ gilt,
- *negativ definit*, falls $(v,v) < 0$ für alle $0 \neq v \in V$ gilt,

- *negativ semidefinit*, falls $(v,v) \le 0$ für alle $v \in V$ gilt.

Trägheitssatz von Sylvester. Sei $K = \mathbb{R}$. *Sei V ein n-dimensionaler $\mathbb{R}$-Vektorraum. Sei $\langle -,- \rangle$ eine symmetrische Bilinearform auf V, sei $v_1, \ldots, v_n$ eine Orthogonalbasis. Sei p die Anzahl der v_i mit $\langle v_i, v_i \rangle > 0$, sei q die Anzahl der v_i mit $\langle v_i, v_i \rangle < 0$. Dann gilt: Die Zahlen p und q hängen nicht von der Wahl der Basis ab.* Man nennt das Tripel (p,q,r) mit $r = n-p-q$ die *Signatur* der Bilinearform.

Beweis: Sei $v_1, \ldots, v_n$ eine Orthogonalbasis von V mit $\langle v_i, v_i \rangle > 0$ für $1 \le i \le p$, mit $\langle v_i, v_i \rangle < 0$ für $p+1 \le i \le p+q$ (und mit $\langle v_i, v_i \rangle = 0$ für $p+q < i \le n$). Sei V_+ das Erzeugnis von $v_1, \ldots, v_p$, sei V_- das Erzeugnis von $v_{p+1}, \ldots, v_{p+q}$ und sei V_0 das Erzeugnis von $v_{p+q+1}, \ldots, v_n$. Die Einschränkung von $\langle -,- \rangle$ auf V_+ ist positiv definit, ist dagegen $v \in V_- \oplus V_0$, so ist die Einschränkung auf $V_- \oplus V_0$ negativ semidefinit. Ist nun U ein beliebiger Unterraum, so dass die Einschränkung von $\langle -,- \rangle$ auf U positiv ist, so muss $U \cap (V_- \oplus V_0) = 0$ gelten, also ist $\dim U \le n-(q+r) = p$. Entsprechend sieht man: Ist U' ein beliebiger Unterraum, so dass die Einschränkung von $\langle -,- \rangle$ auf U' negativ definit ist, so muss $U' \cap (V_+ \oplus V_0) = 0$ gelten, also ist $\dim U' \le n-(p+r) = q$.

Ist also eine andere zweite Orthogonalbasis gegeben, mit Signatur (p', q', r'), so sehen wir $p' \le p, q' \le q$. Umgekehrt gilt aber auch $p \le p'$ und $q \le q'$.

Wir sehen also, dass die Zahlen p, q folgendermaßen charakterisiert werden können:

Sei $\langle -,- \rangle$ eine Bilinearform auf $\mathbb{R}^n$ mit Signatur (p,q,r). Dann gilt:

- *p ist die maximale Dimension eines Unterraums U von V, sodass die Einschränkung von $(-,-)$ auf U positiv definit ist.*
- *q ist die maximale Dimension eines Unterraums U von V, sodass die Einschränkung von $(-,-)$ auf U negativ definit ist.*

Zusatz.

- $r + \min(p,q)$ *ist die maximale Dimension eines total-isotropen Unterraums.*

Beweis des Zusatzes: Zuerst betrachten wir den Fall der Signatur $(1,1,0)$, sei also $\langle -,- \rangle$ durch die Matrix $A = \begin{bmatrix} a & 0 \\ 0 & -b \end{bmatrix}$ mit $a, b > 0$ gegeben. Dann ist offensichtlich $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ mit $x_1 = \sqrt{b}$ und $x_2 = \sqrt{a}$ ein isotroper Vektor. Ist also die Signatur (p,q,r), so erhält man auf diese Weise p orthogonale anisotrope Vektoren der Form $\lambda_i e_i + \lambda'_i e_{p+i} \neq 0$, mit $1 \le i \le \min(p,q)$. Zusammen mit den Vektoren $e_{p+q+1}, \ldots, e_n$ erhält man $r + \min(p,q)$ paarweise orthogonale isotrope Vektoren, sie spannen einen total-isotropen Unterraum der Dimension $r + \min(p,q)$ auf.

Sei umgekehrt U ein total-isotroper Unterraum, sei V_+, V_- wie im Beweis des Trägheitssatzes konstruiert (nach Wahl einer Orthogonalbasis). Es ist $U \cap V_+ = 0$, also $\dim U \le n-p$, und es ist auch $U \cap V_- = 0$, also $\dim U \le n-q$. Also ist $\dim U \le n - \max(p,q) = r + \min(p,q)$ (hier verwenden wir die offensichtliche Beziehung $\max(p,q) + \min(p,q) = p+q$).

Einschub: Polynome und ihre Nullstellenmengen.

Eine wichtige Frage der Algebra (wie auch der algebraischen Geometrie) ist die Untersuchung der Nullstellenmenge von Polynomen. Gegeben ist dabei ein Körper K und wir betrachten den Polynomring $K[T_1,\dots,T_n]$ in den Variablen $T_1,\dots,T_n$; es ist $K[T_1,\dots,T_n] = K[T_1]\cdots[T_n]$, oder, induktiv beschrieben, $K[T_1,\dots,T_n] = K[T_1,\dots,T_{n-1}][T_n]$ (man bildet also iterativ Polynomringe: zuerst $K[T_1]$, dann $K[T_1,T_2]$, usw). Polynome, also die Elemente von $K[T_1,\dots,T_n]$, lassen sich als Linearkombinationen von Monomen mit Koeffizienten in K schreiben, ein *Monom* ist von der Form $T_1^{d_1}T_1^{d_2}\cdots T_n^{d_n}$ mit natürlichen Zahlen $d_1,\dots,d_n$, der Grad dieses Monoms ist $\sum d_i$. Polynome sind also von der Form

$$P = P(T_1,\dots,T_n) = \sum_{d_1,\dots,d_n} c_{d_1,\dots,d_n}\, T_1^{d_1}T_1^{d_2}\cdots T_n^{d_n} \quad \text{mit} \quad c_{d_1,\dots,d_n} \in K$$

(summiert wird über Folgen der Form $(d_1,\dots,d_n) \in \mathbb{N}_0^n$, dabei dürfen nur endlich viele Summanden mit $c_{d_1,\dots,d_n} \neq 0$ auftreten). Der *Grad* von P ist das Maximum der Grade der Monome $T_1^{d_1}T_1^{d_2}\cdots T_n^{d_n}$ mit $c_{d_1,\dots,d_n} \neq 0$. Polynome vom Grad 1 heißen *linear,* solche vom Grad 2 heißen *quadratisch.*

Ist $P(T_1,\dots,T_n)$ ein Polynom, so bezeichnet man mit

$$V(P) = \{(x_1,\dots,x_n) \in K^n \mid P(x_1,\dots,x_n) = 0\}$$

die *Nullstellenmenge von P.* (Der Buchstabe V steht für das Wort *Varietät*, das in der algebraischen Geometrie üblicherweise vewendet wird.) Im Rahmen der linearen Algebra interessiert man sich natürlich für lineare Polynome und deren Nullstellenmengen: *die Nullstellenmenge eines linearen Polynoms ist eine Hyperebene.* Es zeigt sich, dass man mit den Methoden der linearen Algebra auch die quadratischen Polynome und ihre Nullstellenmengen untersuchen kann, dies soll jetzt geschehen. Die Nullstellenmenge eines quadratischen Polynoms nennt man eine *Quadrik.*

Ein Polynom, das eine Linearkombination von Monomen eines festen Grads d ist, heißt *homogen vom Grad d.* Ein homogenes quadratisches Polynom (man spricht auch von einer *quadratischen Form*) ist also auf folgende Weise gegeben:

$$P = P(T_1,\dots,T_n) = \sum\nolimits_{1\le i\le j\le n} \alpha_{ij}T_iT_j, \quad \text{mit} \quad a_{ij} \in K.$$

5D. Quadratische Formen.

Wir setzen jetzt voraus, dass K ein Körper mit $\operatorname{char}(K) \neq 2$ *ist.* Sei eine quadratische Form (also ein homogenes quadratisches Polynom)

$$P = P(T_1,\dots,T_n) = \sum\nolimits_{1\le i\le j\le n} \alpha_{ij}T_iT_j, \quad \text{mit} \quad a_{ij} \in K.$$

gegeben. Wir setzen

$$a_{ij} = \begin{cases} \frac{1}{2}\alpha_{ij} & \text{falls} \quad i < j, \\ \alpha_{ii} & \text{falls} \quad i = j, \\ \frac{1}{2}\alpha_{ji} & \text{falls} \quad i > j. \end{cases}$$

Wir erhalten auf diese Weise eine symmetrische Matrix $A = (a_{ij})_{ij}$, und es gilt

$$P(x_1, \ldots, x_n) = {}^t x A x \quad \text{für} \quad x = {}^t[x_1 \ \ldots \ x_n].$$

Erinnerung: Die symmetrische Matrix A liefert eine symmetrische Bilinearform $\langle -, - \rangle_A$ mit $\langle x, y \rangle_A = {}^t x A y$. Also gilt: *Ist P ein quadratisches Polynom in n Variablen mit Koeffizienten in K, so gibt es eine symmetrische $(n \times n)$-Matrix A, so dass die Funktion $x = (x_1, \ldots, x_n) \mapsto P(x_1, \ldots, x_n)$ durch $x \mapsto \langle x, x \rangle_A = {}^t x A x$ gegeben ist.* Man nennt die Bilinearform $\langle -, - \rangle_A$ die *Polarisierung* der quadratischen Form P. Kennt man P, so erhält die Polarisierung $\langle -, -, \rangle$ von P durch

$$\langle v, w \rangle = \frac{1}{2}(P(v + w) - P(v) - P(w)$$

(hier sieht man, worum man $\operatorname{char} K \neq 2$ voraussetzen muss).

Insgesamt gilt: *Ist K ein Körper der Charakteristik* $\operatorname{char}(K) \neq 2$*, so entsprechen die Bilinearformen auf K^n bijektiv den quadratischen Formen auf K^n*, unter folgenden Zuordnungen: Ist $\langle -, - \rangle$ eine Bilinearform auf K^n, so bilde $P(x) = \langle x, x \rangle$; ist umgekehrt eine quadratische Form P gegeben, so bildet man die Polarisierung von P.

Beispiel 1. Sei $n = 2$ und sei

$$P(T_1, T_2) = 5T_1^2 - 8T_1T_2 + 5T_2^2, \quad \text{also} \quad A = \begin{bmatrix} 5 & -4 \\ -4 & 5 \end{bmatrix}.$$

Das charakteristische Polynom ist

$$\chi_A(T) = (T - 5)^2 - 16 = T^2 - 10T + 9 = (T - 1)(T - 9),$$

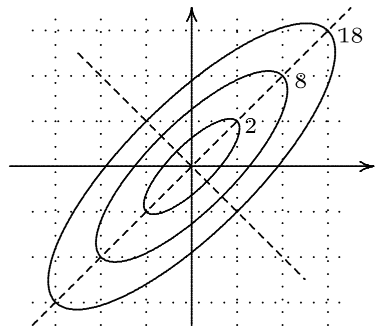
die Eigenwerte sind also 1 und 9. Ein Eigenvektor zum Eigenwert 1 ist der Vektor ${}^t[1, 1]$, ein Eigenvektor zum Eigenwert 9 ist ${}^t[-1, 1]$. Wir erhalten also folgende Orthonormalbasis

$$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Setzen wir $S = \frac{1}{\sqrt{2}} \left[\begin{smallmatrix} 1 & -1 \\ 1 & 1 \end{smallmatrix}\right]$, so ist dies eine orthogonale Matrix (und zwar eine Drehung) und es gilt

$${}^t SAS = \begin{bmatrix} 1 & 0 \\ 0 & 9 \end{bmatrix}.$$

Natürlich liefert eine quadratische Form $P(T_1, T_2)$ eine Funktion $P\colon \mathbb{R}^2 \to \mathbb{R}$, sie ist durch $(x_1, x_2) \mapsto P(x_1, x_2)$ definiert und wird üblicherweise ebenfalls mit P bezeichnet. Hier sind einige "Höhenlinien" der Funktion P, und zwar die Mengen $V(P(T_1, T_2) - c)$ mit $c = 2, 8, 18$, gezeichnet. Diese Nullstellenmengen sind jeweils Ellipsen, ihre Hauptachsen sind gerade die Eigengeraden der Matrix A, sie sind gestrichelt eingezeichnet.



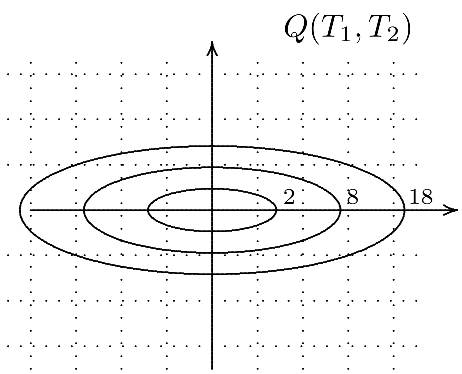
Wir sehen: Durch die Drehung S_1 können wir erreichen, dass die Koordinatenachsen die Hauptachsen sind, wir erhalten zur transformierten Matrix $B = {}^tS_1AS_1 = \begin{bmatrix} 1 & 0 \\ 0 & 9 \end{bmatrix}$ die quadratische Form

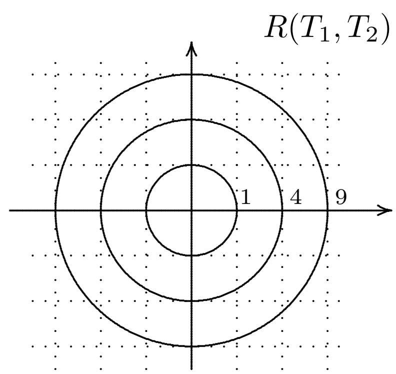
$$Q(T_1, T_2) = T_1^2 + 9T_2^2.$$

Die entsprechende Funktion $Q\colon \mathbb{R}^2 \to \mathbb{R}$ wird duch das folgende linke Bild beschrieben; durch eine zusätzliche Skalierung der Koordinatenachsen erhalten wir aus $Q(T_1, T_2)$ das quadratische Polynom

$$R(T_1, T_2) = T_1^2 + T_2^2,$$

dabei verwenden wir die Matrix $S_2 = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{3} \end{bmatrix}$ und bilden $C = {}^tS_2BS_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Das folgende recht Bild zeigt Nullstellenmengen $V(R(T_1, T_2) - c)$ für $c > 0$, es sind jeweils Kreise:



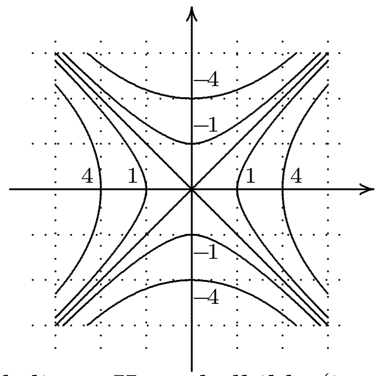


Die Matrix A, die wir hier betrachtet haben, hat die Signatur $(2, 0, 0)$. Betrachten wir statt $P = P(T_1, T_2)$ das Polynom $-P$, so ist die Signatur der zugehörigen Matrix $(0, -2, 0)$, das Bild der Nullstellenmengen $V(\dots)$ (also der "Höhenlinien") ändert sich bis auf die Beschriftung nicht: die Höhenlinien von P zum Wert c sind für $-P$ diejenigen zum Wert $-c$.

Beispiel 2. Wesentlich verschiedene Höhenlinien erhalten wir, wenn es sich um eine Matrix A mit Signatur $(1,1,0)$ handelt, dann nämlich sind die Höhenlinien für $c \neq 0$ Hyperbeln, und für $c = 0$ erhält man ein sich schneidendes Geradenpaar. Das typische derartige Beispiel erhält man durch

$$P(T_1, T_2) = T_1^2 - T_2^2,$$

hier das entsprechende Höhenlinienbild:



Ein wesentlicher Unterschied dieses Hyperbelbilds (im Gegensatz zum Ellipsenbild) ist das Vorhandensein von Null verschiedener isotroper Vektoren: Die Signatur von A ist $(1,1,0)$, also ist die maximale Dimension eines total-isotropen Unterraums 1, die Nullstellenmenge $V(P)$ ist ein Paar sich schneidender Geraden. Die Faktorisierung

$$P(T_1, T_1) = T_1^2 - T_2^2 = (T_1 - T_2)(T_1 + T_2)$$

zeigt $V(P) = V(T_1 - T_2) \cup V(T_1 + T_2)$. Die Hauptachsen für $P(T_1, T_2) = T_1^2 - T_2^2$ sind die Koordinatenachsen (denn dies sind ja die Eigengeraden der Matrix $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$), — nicht etwa die beiden total-isotropen Geraden!

Es bleiben die Fälle mit Signatur $(1,0,1), (0,1,1)$ und $(0,0,2)$; in den ersten beiden Fällen sind die "Höhenlinien" jeweils Geraden (oder besser: Paare paralleler Geraden), im letzten Fall ist A die Nullmatrix und P ist das Nullpolynom.

5E. Quadriken

Erinnerung: Eine Quadrik ist die Nullstellenmenge eines (nicht notwendig homogenen) quadratischen Polynoms. Ein quadratisches Polynom ist von der Form

$$P = P(T_1, \dots, T_n) = \sum\nolimits_{1 \le i \le j \le n} \alpha_{ij} T_i T_j + \sum\nolimits_{1 \le j \le n} \alpha_{0j} T_j + \alpha_{00},$$

mit Koeffizienten α_{ij} in einem Körper K. Auch hier wollen wir mit einer Matrix arbeiten. Ist $\operatorname{char}(K) \neq 2$, so ordnen wir P die erweiterte Koeffizientenmatrix A' mit

$$a_{ij} = \begin{cases} \frac{1}{2}\alpha_{ij} & \text{falls} \quad i < j \\ \alpha_{ii} & \text{falls} \quad i = j \\ \frac{1}{2}\alpha_{ji} & \text{falls} \quad i > j \end{cases}$$

für $0 \leq i, j \leq n$ zu. Neben der Matrix A' betrachten wir auch die Matrix $A = (a_{ij})_{1 \leq i,j \leq n}$. Dem Vektor $x = {}^t[x_1 \ \cdots \ x_n]$ ordnen wir den erweiterten Vektor $x' = {}^t[1 \ x_1 \ \cdots \ x_n]$ zu. Wir erhalten dann

$$P(x) = {}^t(x')A'x'.$$

Wir sehen also, dass sich auch nicht-homogene quadratische Polynome mit Hilfe von symmetrischen Matrizen beschreiben lassen.

Es sei daran erinnert (Übungsaufgabe 9-2), dass sich Isometrien des $\mathbb{R}^n$ ebenfalls sehr gut durch Matrizen beschreiben lassen, wenn man auch hier erweiterte Vektoren heranzieht: Eine Isometrie ϕ des $\mathbb{R}^n$ ist von der Form $\phi(x) = Ax+a$, für alle $x \in \mathbb{R}^n$, dabei ist $A \in \mathcal{O}(n)$ und $a \in \mathbb{R}^n$, also $\phi = t_a \circ f_A$. Wir ordnen diesem Paar A, a die $(n+1) \times (n+1)$-Matrix $\left[\begin{smallmatrix} 1 & 0 \\ a & A \end{smallmatrix}\right]$ zu. Natürlich gilt

$$\begin{bmatrix} 1 & 0 \\ a & A \end{bmatrix} \begin{bmatrix} 1 \\ x \end{bmatrix} = \begin{bmatrix} 1 \\ Ax+a \end{bmatrix},$$

dies zeigt, dass der Isometrie $t_a \circ f_A$ die Matrix $\left[\begin{smallmatrix} 1 & 0 \\ a & A \end{smallmatrix}\right]$ zugeordnet ist. Man kann dies auch so formulieren: Wir betrachten die Zuordnung $\eta \colon \mathcal{B}(n) \to GL(n+1, \mathbb{R})$

$$\eta(t_a \circ f_A) = \begin{bmatrix} 1 & 0 \\ a & A \end{bmatrix} \quad \text{für} \quad a \in \mathbb{R}^n \quad \text{und} \quad A \in \mathcal{O}(n)$$

dies ist ein injektiver Gruppen-Homomorphismus; ist $\phi \in \mathcal{B}(n)$ und $x \in \mathbb{R}^n$, so gilt:

$$\eta(\phi) \begin{bmatrix} 1 \\ x \end{bmatrix} = \begin{bmatrix} 1 \\ \phi(x) \end{bmatrix}.$$

Unser Ziel ist es, Normalformen der Quadriken im $\mathbb{R}^n$ zu beschreiben: Jedes reelle Polynom P in n Variablen liefert eine Abbildung $\mathbb{R}^n \to \mathbb{R}$, nämlich

$${}^t[x_1, \ldots, x_n] \mapsto P(x_1, \ldots, x_n),$$

diese Abbildung bezeichnet man meist ebenso mit P. Durch jeden Basiswechsel und jede Verschiebung des Koordinatensystems im $\mathbb{R}^n$ werden die Koeffizienten von P verändert; wir fragen, inwieweit wir durch eine Verschiebung des Koordinatensystems und einen orthogonalen Basiswechsel die Form von P vereinfachen können. Gesucht ist also eine Isometrie ϕ des $\mathbb{R}^n$, so dass die Hintereinanderschaltung

$$\mathbb{R}^n \xrightarrow{\phi} \mathbb{R}^n \xrightarrow{P} \mathbb{R}$$

eine möglichst einfache Form hat. Da wir hier vor allem an der Nullstellenmenge $V(P)$ interessiert sind, werden wir auch erlauben, dass P durch ein skalares

Vielfaches γP mit $\gamma \in \mathbb{R} \setminus \{0\}$ ersetzt wird, denn für derartige Vielfache gilt $V(\gamma P) = V(P)$.

Satz (Isometrische Klassifikation der Quadriken in $\mathbb{R}^n$). *Sei P ein quadratisches Polynom in n Variablen mit reellen Koeffizienten. Dann gibt es eine Isometrie φ des $\mathbb{R}^n$, und* **positive** *reelle Zahlen r_i mit $1 \leq i \leq m$, so dass $P \circ \phi$ ein skalares Vielfaches eines der folgenden Polynome ist:*

(**a**) $$\sum_{i=1}^{p} r_i T_i^2 - \sum_{i=p+1}^{m} r_i T_i^2,$$

(**b**) $$\sum_{i=1}^{p} r_i T_i^2 - \sum_{i=p+1}^{m} r_i T_i^2 - 1,$$

(**c**) $$\sum_{i=1}^{p} r_i T_i^2 - \sum_{i=p+1}^{m} r_i T_i^2 - T_{m+1}.$$

dabei ist jeweils $0 \leq p \leq m \leq n$, im Fall (c) *ist $m < n$; in den Fällen* (a) *und* (c) *kann man zusätzlich $p \geq \frac{1}{2}m$ voraussetzen.*

Zusatz. Es ist einfach festzustellen, welcher der Fälle **(a)**,**(b)** oder **(c)** vorliegt: Sei $P = P(T_1, \ldots, T_n) = P_2 + L$, dabei sei P_2 ein homogenes quadratisches Polynom, während L ein Polynom mit Grad höchstens 1 ist. Die (erweiterte) Koeffizientenmatrix von P sei A', die Koeffizientenmatrix von P_2 sei A. Sei $m = \operatorname{rang}(A)$, $m' = \operatorname{rang}(A')$. Dann gilt:

$m' = m$	Fall **(a)**,
$m' = m + 1$	Fall **(b)**,
$m' = m + 2$	Fall **(c)**.

Beweis: Sei (p, q, r) die Signatur von A. Es ist $m = p + q$, dies ist der Rang der Matrix A.

Schritt 1 (Hauptachsentransformation). Wir wählen in $\mathbb{R}^n$ eine Orthonormalbasis $v_1, \ldots, v_n$ bezüglich des kanonischen inneren Produkts, die aus Eigenvektoren von A besteht. Seien $d_1, \ldots, d_n$ die zugehörigen Eigenwerte. Wir werden immer voraussetzen, dass $d_i = 0$ für $m < i \leq n$ gilt. Falls notwendig, können wir annehmen, dass die Vektoren $v_1, \ldots, v_n$ so angeordnet sind, dass $d_i > 0$ für $1 \leq i \leq p$ und $d_i < 0$ für $p < i \leq m (= p + q)$ gilt (oder aber auch, dass zuerst die Eigenvektoren mit negativen, danach die mit positiven Eigenwerten kommen). Es gibt also eine orthogonale Matrix S mit ${}^tSAS = D = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{bmatrix}$ gilt. Statt A' betrachten wir nun die transformierte Matrix $B' = {}^tG_1 A' G_1$, dabei ist G_1 die direkte Summe der (1×1)-Matrix $[1]$ und der Matrix S:

$$B' = {}^tG_1 A' G_1 = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & {}^tS & \\ 0 & & & \end{array}\right] \cdot \left[\begin{array}{c|ccc} a_{00} & a_{01} & \cdots & a_{0n} \\ \hline a_{10} & & & \\ \vdots & & A & \\ a_{n0} & & & \end{array}\right] \cdot \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & S & \\ 0 & & & \end{array}\right] = \left[\begin{array}{c|ccc} b_0 & b_1 & \cdots & b_n \\ \hline b_1 & d_1 & & \\ \vdots & & \ddots & \\ b_n & & & d_n \end{array}\right]$$

Schritt 2. Sei $u = {}^t[u_1, \ldots, u_n]$ mit

$$u_i = \begin{cases} -\frac{b_i}{d_i} & \text{für} \quad 1 \le i \le p+q \\ 0 & \text{für} \quad p+q < i \le n \end{cases}$$

und betrachte die Matrix

$$G_2 = \left[\begin{array}{c:ccc} 1 & 0 & \cdots & 0 \\ \hdashline & 1 & & \\ u & & \ddots & \\ & & & 1 \end{array}\right]$$

Bilden wir nun $C' = {}^tG_2 B' G_2$, so erhalten wir eine Matrix der Form

$$C' = {}^tG_2 B' G_2 = \left[\begin{array}{c:ccc:ccc} c & 0 & \cdots & 0 & b_{m+1} & \cdots & b_n \\ \hdashline 0 & d_1 & & & & & \\ \vdots & & \ddots & & & 0 & \\ 0 & & & d_m & & & \\ \hdashline b_{m+1} & & & & & & \\ \vdots & & 0 & & & 0 & \\ b_n & & & & & & \end{array}\right]$$

Schritt 3. Wir unterscheiden nun 3 Fälle: Wir erinnern daran, dass wir den Rang der Matrix A' (und damit auch den der Matrizen B' und C') mit m' bezeichnet haben.

Fall (a): $c = 0$ und $b_{m+1} = \cdots = b_n = 0$. *In diesem Fall ist* $m' = m$. Hier haben wir es mit dem homogenen quadratischen Polynom

$$\sum_{i=1}^{m} d_i T_i^2$$

zu tun, dabei können wir voraussetzen $d_i > 0$ für $1 \le i \le p$ und $d_i < 0$ für $p < i \le m$. Zusätzlich können wir hier auch annehmen, dass $p \ge \frac{1}{2}m$ gilt (sonst multiplizieren wir das Polynom mit -1 und vertauschen die T_i).

Fall (b): $c \ne 0$ und $b_{m+1} = \cdots = b_n = 0$. *In diesem Fall gilt* $m' = m+1$. Wir multiplizieren wir das Polynom mit $-\frac{1}{c}$ und erhalten

$$\sum_{i=1}^{m} d_i T_i^2 - 1,$$

dabei können wir wieder annehmen, dass $d_i > 0$ für $1 \le i \le p$ und $d_i < 0$ für $p < i \le m$ gilt.

Fall (c): Es gibt ein i mit $b_i \ne 0$. *In diesem Fall gilt* $m' = m+2$. Wir betrachten den Vektor $w = {}^t[b_{m+1}, \ldots, b_n] \in \mathbb{R}^{n-m}$ und normieren ihn: $w_1 = \frac{1}{||w||} w$. Sei $w_1, \ldots, w_{n-m}$ eine Orthonormalbasis von $\mathbb{R}^{n-m}$ (bekanntlich erhalten wir eine solche Orthonormalbasis, indem wir eine beliebige Basis, die w_1 als erstes Element enthält, nach dem Gram-Schmidt-Verfahren orthonormalisieren). Sei nun $R \in \mathrm{GL}(n-m,\mathbb{R})$ mit Spalten $w_1, \ldots, w_{n-m}$, dies ist eine orthogonale Matrix. Sei $v = -\frac{c}{2\langle v,v\rangle} w$. Wir bilden die Matrix

$$G_3 = \left[\begin{array}{c|ccc|ccc}
1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
\hline
0 & 1 & & & & & \\
\vdots & & \ddots & & & 0 & \\
0 & & & 1 & & & \\
\hline
v & & 0 & & & R & \\
\end{array}\right]$$

Bilden wir nun $D' = {}^tG_3C'G_3$, so erhalten wir die Matrix

$$D' = {}^tG_3C'G_3 = \left[\begin{array}{c|ccc|cccc}
0 & 0 & \cdots & 0 & ||w|| & 0 & \cdots & 0 \\
\hline
0 & d_1 & & & & & & \\
\vdots & & \ddots & & & 0 & & \\
0 & & & d_m & & & & \\
\hline
||w|| & & & & & & & \\
0 & & 0 & & & 0 & & \\
\vdots & & & & & & & \\
0 & & & & & & & \\
\end{array}\right]$$

Das zugehörige quadratische Polynom $\sum_{i=1}^m d_i T_i^2 + 2||w||T_{m+1}$ können wir durch Multiplikation mit $-\frac{1}{2||w||}$ in die Form

$$\sum_{i=1}^m d_i' T_i^2 - T_{m+1}$$

bringen. Dabei können wir voraussetzen, dass $d_i' > 0$ für $1 \leq i \leq p$ und $d_i' < 0$ für $p < i \leq m$ gilt. Zusätzlich können wir hier wieder annehmen, dass $p \geq \frac{1}{2}m$ gilt (andernfalls ersetze T_{m+1} durch $-T_{m+1}$).

Es sollte betont werden, dass im Schritt 3 die drei Fälle (a), (b), (c) gerade den drei Möglichkeiten für $m' - m = 0, 1, 2$ entsprechen. Insbesondere sehen wir, dass in den Fällen $m' - m = 0$ und 1 nur die beiden Schritte 1 und 2 von Bedeutung sind.

Zum Abschluss des Beweises sollten wir uns vergewissern, dass die betrachteten Matrizen G_1, G_2, G_3 wirklich Isometrien des $\mathbb{R}^n$ beschreiben. Die Matrix G_1 beschreibt eine orthogonale Änderung des Koordinatensystems, die Matrix G_2 eine Translation, die dritte Matrix G_3 ist Hintereinanderanderschaltung einer orthogonalen Änderung des Koordinatensystems und einer Translation.

Folgerung. *Sei P reelles quadratisches Polynom. Genau dann ist $V(P)$ leer, wenn A' positiv semidefinit oder negativ semidefinit ist und $m' = m+1$ gilt.*

Beweis: Gehe die einzelnen Normalformen der isometrischen Klassifikation durch.

Sei V ein Vektorraum. Unter einer *affinen Abbildung* $\phi\colon V \to V$ versteht man die Hintereinanderschaltung eines Endomorphismus und einer Translation (dies ist für jeden Grundkörper K definiert), es ist also $\phi(v) = g(v)+u$ für alle $v \in V$, dabei ist $g\colon V \to V$ eine lineare Abbildung und $a \in V$ ein fest gewählter Vektor. Eine derartige affine Abbildung ist genau dann invertierbar, wenn die zugehörige lineare Abbildung invertierbar ist. Sei nun ϕ eine affine Abbildung des Vektorraums K^n in sich, etwa $\phi(x) = Ax + a$, mit $A \in \mathrm{GL}(n,K)$ und $a \in K^n$. Wie im Fall von Isometrien des $\mathbb{R}^n$ ordnen wir diesem Paar A, a die $(n+1)\times(n+1)$-Matrix $\begin{bmatrix} 1 & 0 \\ a & A \end{bmatrix}$ zu und arbeiten mit “erweiterten Vektoren”: Dem Vektor $x = {}^t[x_1 \ \cdots \ x_n]$ ordnen wir den erweiterten Vektor $x' = {}^t[1 \ x_1 \ \cdots \ x_n]$ zu. Es gilt auch hier

$$\begin{bmatrix} 1 & 0 \\ a & A \end{bmatrix} \begin{bmatrix} 1 \\ x \end{bmatrix} = \begin{bmatrix} 1 \\ Ax+a \end{bmatrix},$$

dies zeigt, wie man jeder affinen Abbildung eine darstellende Matrix zuordnen kann.

Satz (Affine Klassifikation der Quadriken in $\mathbb{R}^n$). *Sei P ein quadratisches Polynom in n Variablen mit reellen Koeffizienten. Dann gibt es eine invertierbare affine Abbildung φ des $\mathbb{R}^n$, so dass $P\circ\phi$ ein skalares Vielfaches eines der folgenden Polynome ist:*

(a) $$\sum_{i=1}^{p} T_i^2 - \sum_{i=p+1}^{m} T_i^2,$$

(b) $$\sum_{i=1}^{p} T_i^2 - \sum_{i=p+1}^{m} T_i^2 - 1,$$

(c) $$\sum_{i=1}^{p} T_i^2 - \sum_{i=p+1}^{m} T_i^2 - T_{m+1}.$$

dabei ist jeweils $0 \le p \le m \le n$, im Fall (c) *ist $m < n$; in den Fällen* (a) *und* (c) *kann man zusätzlich $p \ge \frac{1}{2}m$ voraussetzen.*

Beweis: wir erhalten dies Ergebnis unmittelbar aus der isometrischen Klassifkation, wenn wir bedenken, dass wir durch Skalierung der Koordinatenachsen die positiven Zahlen r_i durch 1 ersetzen können.

5F. Kegelschnitte

Wir betrachten nun Quadriken in der reellen Ebene $\mathbb{R}^2$, und zwar Ellipsen, Parabeln und Hyperbeln. Man spricht hier von “Kegelschnitten”, denn sie entstehen, wie jetzt gezeigt werden soll, als Schnittmengen des Kegels $M = V(T_1^2 + T_2^2 - T_3^2)$ mit einer Ebene H, die nicht durch den Ursprung geht; die Wahl einer geeigneten derartigen Ebene liefert alle möglichen Ellipsen, Parabeln und Hyperbeln (schneidet man diesen Kegel $V(T_1^2+T_2^2-T_3^2)$ mit einer Ebene, die durch den Ursprung geht, so erhält man zwei sich schneidende Geraden, oder eine Doppelgerade oder einen einzelnen Punkt). Um aber die Teilmenge $M \cap H$ als Quadrik in der Ebene H beschreiben zu können, müssen wir H mit $\mathbb{R}^2$ identifizieren, wir betrachten also eine injektive Abbildung $\phi\colon \mathbb{R}^2 \to \mathbb{R}^3$, deren Bild H ist – natürlich nicht eine beliebige Abbildung, sondern eine, die zumindest abstandserhaltend ist (wie wir gleich sehen werden, ist diese Bedingung ausreichend). Es empfiehlt sich, folgende Begriffe zu verwenden:

Ist K ein beliebiger Körper und sind V, W K-Vektorräume, so betrachten wir *affine Abbildungen* $\phi\colon V \to W$. Eine derartige affine Abbildung ϕ ist durch $\phi(v) = g(v)+w$ für alle $v \in V$ gegeben, dabei ist $g\colon V \to W$ eine lineare Abbildung und $w \in W$ ein fest gewählter Vektor. Ist $V = K^n$ und $W = K^m$, so können wir diesem Paar A, w die $(m+1)\times(n+1)$-Matrix $\begin{bmatrix} 1 & 0 \\ w & A \end{bmatrix}$ zugeordnen. Es sei $A = (a_{ij})_{ij}$.

Ist $P(T_1,\dots,T_m)$ ein Polynom in m Variablen mit Koeffizienten in K, so erhält man ein Polynom $P\circ\phi$ in n Variablen durch

$$P\circ\phi(T_1,\dots,T_n) = P\Big(\sum_{j=1}^n a_{1j}T^j + w_1, \dots, \sum_{j=1}^n a_{mj} + w_m\Big).$$

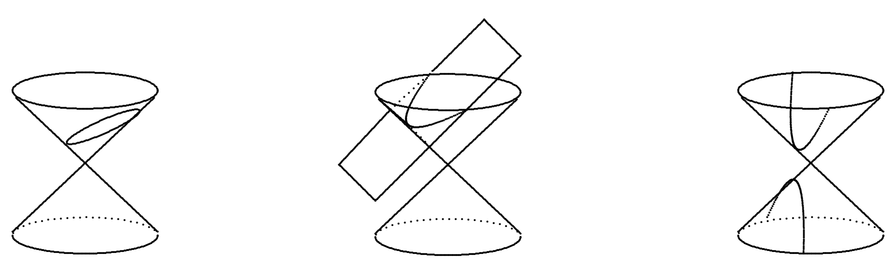
Ist d der Grad von P, so ist der Grad von $P\circ\phi$ höchstens d.

Nun zurück zum Fall $K = \mathbb{R}$. Wie im Fall $n = m$ zeigt man:

Lemma. *Sei $g\colon \mathbb{R}^n \to \mathbb{R}^m$ eine lineare Abbildung und $w \in \mathbb{R}^m$. Die affine Abbbildung $\phi(x) = g(x)+w$ für $x \in \mathbb{R}^n$ ist genau dann abstandserhaltend, wenn die Vektoren $g(e_1),\dots,g(e_n)$ orthonormiert sind.*

Satz. *Zu jeder Ellipse, Parabel und Hyperbel N im $\mathbb{R}^2$ gibt es eine abstandserhaltende affine Abbildung $\phi\colon \mathbb{R}^2 \to \mathbb{R}^3$ mit $N = \phi^{-1}(V(T_1^2+T_2^2-T_3^2))$.*

Eine abstandserhaltende affine Abbildung $\phi\colon \mathbb{R}^2 \to \mathbb{R}^3$ liefert als Bild eine Ebene H im $\mathbb{R}^3$, und wir können $\mathbb{R}^2$ vermöge ϕ mit dieser Ebene H identifizieren. Dann ist $\phi^{-1}(V(T_1^2+T_2^2-T_3^2)) = H \cap V(T_1^2+T_2^2-T_3^2)$. Der Satz besagt also, dass alle Ellipsen, Parabeln und Hyperbeln im $\mathbb{R}^2$ "Kegelschnitte" sind. Als weitere ("entartete") Kegelschnitte erhält man auch zwei sich schneidende Geraden, eine Doppelgerade und einen einzelnen Punkt; in diesen "entarteten" Fällen wählt man Ebenen, die den Ursprung enthalten.



Ellipse Parabel Hyperbel

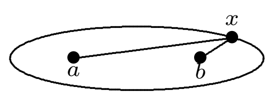
Beweis: Wir wählen Ebenen der Form

$$H = \{ \begin{bmatrix} y-r \\ x \\ \lambda y+r \end{bmatrix} \mid x, y \in \mathbb{R} \}$$

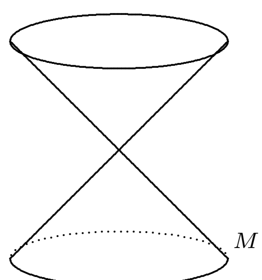
mit $\lambda \geq 0$ und $r > 0$. Wir erhalten als Schnitt $H \cap V(T_1^2 + T_2^2 - T_3^2)$ für $0 \leq \lambda < 1$ eine Ellipse, für $\lambda = 1$ eine Parabel, und für $\lambda > 1$ eine Hyperbel, und zwar erhalten wir alle möglichen Normalformen.

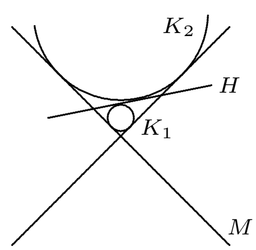
Bemerkung 1. Analog zeigt man: *Schneidet man einen beliebigen elliptischen Zylinder, zum Beispiel den Zylinder $V(T_1^2+T_2^2-1)$ im $\mathbb{R}^3$ mit einer Ebene, so erhält man eine Ellipse, ein Geradenpaar, eine (Doppel-)Gerade oder die leere Menge.*

Bemerkung 2. Dass man beim Schneiden eines Kreis-Zylinders Ellipsen erhält, erscheint nicht zu überraschend, man sieht dies beim Metzger! Dagegen ist das Auftreten der Ellipsen beim Schneiden eines Kreis-Kegels gar nicht so einleuchtend. Es gibt dafür einen elementar-geometrischen Beweis, der von Dandelin stammt: man kann Ellipsen auf folgende Weise konstruieren: gegeben sind zwei Punkte a, b in der Ebene $\mathbb{R}^2$ und eine Zahl $r > d(a,b)$, man betrachte die Menge $E(a,b;r)$ aller Punkte $x \in \mathbb{R}^2$ mit $d(x,a) + d(x,b) = r$ (die Summe der Abstände, die x von a und von b hat, soll also gleich r sein); *es ist $E(a,b;r)$ eine Ellipse und jede Ellipse erhält man auf diese Weise;* man nennt a, b die Brennpunkte der Ellipse $E(a,b;r)$ (so konstruieren Gärtner Ellipsen).



Die Ebene H schneide den Kegel $M = V(T_1^2 + T_2^2 - T_3^2)$ in einer Ellipse. Wie erhält man deren Brennpunkte? Man sucht die beiden Kugeln K_1, K_2, die den Kegel M und die Ebene H berühren (man nennt sie die *Dandelin'sche Kugeln*); *die Berührpunkte der Kugeln $K_1 \cap H$ und $K_2 \cap H$ sind gerade die Brennpunkte der Ellipse $M \cap H$.*





Beweis: Sei $a_1 = K_1 \cap H$, $a_2 = K_2 \cap H$. Für einen beliebigen Punkt $x \in M \cap H$ ist zu zeigen, dass die Summe der Abstände $d(x,a_1)+d(x,a_2)$ konstant ist. Betrachte die Kreise $Z_1 = K_1 \cap M$ und $Z_2 = K_2 \cap M$, sie liegen offensichtlich in parallelen Ebenen. Sei L die Urspungsgerade durch x, sie schneide den Kreis Z_1 im Punkt x_1 und den Kreis Z_2 im Punkt x_2. Der Abstand $d(x_1,x_2)$ ist sicher von x unabhängig, denn dieser Abstand hängt nur von den Kreisen Z_1, Z_2 ab. Es ist aber $d(x,x_1) = d(x,a_1)$, denn die Gerade durch x und x_1, wie auch die durch x und a_1 sind Tangenten an die Kugel K_1. Entsprechend ist $d(x,x_2) = d(x,a_2)$. Insgesamt sehen wir

$$d(x,a_1) + d(x,a_2) = d(x,x_1) + d(x,x_2) = d(x_1,x_2).$$