6 Bilinearformen und quadratische Formen

Definition 6.1: *Seien $V_1, \ldots, V_r$ Vektorräume über dem Körper K. Eine Abbildung $f : V_1 \times \cdots \times V_r \to K$ heisst* **multilinear** *oder eine* **Multilinearform** *auf $V_1 \times \cdots \times V_r$, falls für alle $1 \leq i \leq r$, alle $v_i, v_i' \in V_i$ und alle $\lambda_i, \lambda_i' \in K$ gilt:*

$$f(v_1, \ldots, \lambda_i v_i + \lambda_i' v_i', \ldots, v_r) = \lambda_i f(v_1, \ldots, v_i, \ldots, v_r) + \lambda_i' f(v_1, \ldots, v_i', \ldots, v_r).$$

Mit $\mathrm{Mult}_K(V_1, \ldots, V_r)$ bezeichnen wir die Menge aller Multilinearformen von $V_1 \times \cdots \times V_r$ nach K.
Ist $r = 2$, so sagt man statt multilinear auch **bilinear** *und statt Multilinearform auch* **Bilinearform**. *Mit $\mathrm{Bil}_K(V_1, V_2)$ bezeichen wir die Menge aller Bilinearformen von $V_1 \times V_2$ nach K.* □

Beispiel 6.2: (i) Das kanonische innere Produkt auf $\mathbb{R}^n$, also

$$f\big((x_1, \ldots, x_n), (y_1, \ldots, y_n)\big) \;=\; \sum_{i=1}^{n} x_i y_i \;,$$

ist eine Bilinearform auf $\mathbb{R}^n \times \mathbb{R}^n$.
(ii) Die Abbildung

$$\begin{array}{cccc} f: & \mathbb{R}^2 \times \mathbb{R}_3 & \longrightarrow & \mathbb{R} \\ & \big((x_1, x_2), (y_1, y_2, y_3)^T\big) & \mapsto & (x_1, x_2) \cdot \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \end{array}$$

ist eine Bilinearform auf $\mathbb{R}^2 \times \mathbb{R}^3$.
(iii) Allgemein ist für eine Matrix $A \in \mathrm{Mat}_{m \times n}(K)$ die Abbildung

$$\begin{array}{cccc} f: & K^m \times K_n & \longrightarrow & K \\ & \big((x_1, \ldots, x_m), (y_1, \ldots, y_n)^T\big) & \mapsto & (x_1, \ldots, x_m) \cdot A \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \end{array}$$

eine Bilinearform auf $K^m \times K_n$.
(iv) Die Determinantenabbildung

$$\begin{array}{cccc} \det: & (K_n)^n & \longrightarrow & K \\ & (a_1, \ldots, a_n) & \mapsto & \det(a_1, \ldots, a_n) \end{array}$$

ist eine Multilinearform auf $(K_n)^n$, also $\det \in \mathrm{Mult}_K(K_n, \ldots, K_n)$. □

Beispiel 6.3: Sei $V = C(\mathbb{R})$ der Vektorraum der stetigen reellen Funktionen, sei $f(x,y)$ eine fixe stetige reelle Funktion in 2 Variablen, und seien $a, b \in \mathbb{R}$. Aus den Eigenschaften des (Riemann-) Integrals sieht man, dass die Abbildung $F : V \times V \to \mathbb{R}$ mit

$$F(g,h) \;=\; \int_a^b \int_a^b g(x) f(x,y) h(y) \, dx \, dy$$

eine Bilinearform ist. □

Beispiel 6.4: Eine Mulilinearform oder Bilinearform ist i.a. **nicht** linear. So ist etwa das Skalarprodukt (Bsp. 6.2(i)) nicht linear auf $\mathbb{R}^n \times \mathbb{R}^n$.
Umgekehrt ist i.a. eine lineare Abbildung nicht multilinear, wie man etwa an $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $f(x, y) = 2x + 3y$, sieht:

$$f(1 + 1, 1) = 7 \neq 10 = f(1, 1) + f(1, 1). \qquad \square$$

Satz 6.5: *Ist $f \in \mathrm{Mult}_K(V_1, \ldots, V_r)$ und enthält $v \in V_1 \times \cdots \times V_r$ eine Nullkomponente, dann ist $f(v) = 0$.*

Beweis: Sei $i \in \{1, \ldots, r\}$ so, dass $v_i = 0$. Dann gilt für ein beliebiges $x \in V_i$:

$$\begin{aligned} & f(v_1, \ldots, 0, \ldots, v_r) \\ = \; & f(v_1, \ldots, x - x, \ldots, v_r) \\ = \; & f(v_1, \ldots, x, \ldots, v_r) - f(v_1, \ldots, x, \ldots, v_r) \\ = \; & 0. \end{aligned} \qquad \square$$

Satz 6.6: *Seien $V_1, \ldots, V_r$ Vektorräume über dem Körper K. Dann ist $\mathrm{Mult}(V_1, \ldots, V_r)$ ein Vektorraum über K.*

Beweis: das ist offensichtlich aus der Definition. □

Satz 6.7: *Seien $f \in \mathrm{Mult}_K(V_1, \ldots, V_r)$ und $g \in \mathrm{Mult}_K(W_1, \ldots, W_s)$. Dann ist die Abbildung*

$$\begin{array}{rccc} f \otimes g : & \times_{i=1}^r V_i \;\times\; \times_{j=1}^s W_j & \longrightarrow & K \\ & (x_1, \ldots, x_r, y_1, \ldots, y_s) & \mapsto & f(x_1, \ldots, x_r) \cdot g(y_1, \ldots, y_s) \end{array}$$

ebenfalls multilinear, also $f \in \mathrm{Mult}_K(V_1, \ldots, V_r, W_1, \ldots, W_s)$.

Beweis: einfaches Ausrechnen. Etwa für die erste Komponente (analog für andere Komponenten):

$$\begin{aligned} & f \otimes g(x_1 + \lambda x_1', x_2, \ldots, x_r, y_1, \ldots, y_s) = \\ & f(x_1 + \lambda x_1', x_2, \ldots, x_r) \cdot g(y_1, \ldots, y_s) = \\ & \big[f(x_1, x_2, \ldots, x_r) + f(\lambda x_1', x_2, \ldots, x_r)\big] \cdot g(y_1, \ldots, y_s) = \\ & f(x_1, x_2, \ldots, x_r) \cdot g(y_1, \ldots, y_s) + f(\lambda x_1', x_2, \ldots, x_r) \cdot g(y_1, \ldots, y_s) = \\ & f \otimes g(x_1, \ldots, x_r, y_1, \ldots, y_s) + \lambda f \otimes g(x_1', \ldots, x_r, y_1, \ldots, y_s). \end{aligned}$$

Somit ist $f \otimes g$ multilinear auf $V_1 \times \cdots \times V_r \times W_1 \times \cdots \times W_s$. □

Definition 6.8: *Seien $f \in \mathrm{Mult}_K(V_1, \ldots, V_r)$ und $g \in \mathrm{Mult}_K(W_1, \ldots, W_s)$. Das in Satz 10.7 definierte Produkt $f \otimes g$ heisst das* **Tensorprodukt** *der Multilinearformen f und g.*
□

Beispiel 6.9: Für $1 \leq i \leq p$ ist die Projektion $\mathrm{pr}_i^{(p)} : K^p \to K$ linear, also $\mathrm{pr}_i^{(p)} \in \mathrm{Mult}_K(K^p)$. Seien $\mathrm{pr}_i^{(m)} \in \mathrm{Mult}_K(K^m)$ bzw. $\mathrm{pr}_j^{(n)} \in \mathrm{Mult}_K(K^n)$ die i-te bzw. j-te Projektion. Dann ist

$$\begin{array}{rcccl} \mathrm{pr}_i^{(m)} \otimes \mathrm{pr}_j^{(n)} : & K^{m+n} & \longrightarrow & K & \\ & (x_1, \ldots, x_m, y_1, \ldots, y_n) & \mapsto & x_i y_j & = (x_1, \ldots, x_m) E_{ij} (y_1, \ldots, y_n)^T \;, \end{array}$$

das Tensorprodukt dieser Projektionen, $\mathrm{pr}_i^{(m)} \otimes \mathrm{pr}_j^{(n)} \in \mathrm{Mult}_K(K, \ldots, K)$. □

Das Tensorprodukt kann auf offensichtliche Weise ausgedehnt werden auf mehrere Abbildungen $f_1 \otimes f_2 \otimes \cdots \otimes f_t$.

Satz 6.10: *Seien $V_1, \ldots, V_r$ endlichdimensionale Vektorräume über K. Seien $B_i = (b_1^{(i)}, \ldots, b_{n_i}^{(i)})$ geordnete Basen von V_i, und seien $B_i^* = (\beta_1^{(i)}, \ldots, \beta_{n_i}^{(i)})$ die dazu dualen (geordneten) Basen von V_i^* (vgl. 3.3.24). Dann ist*

$$B^* = \left(\beta_{j_1}^{(1)} \otimes \cdots \otimes \beta_{j_r}^{(r)}\right)_{1 \le j_i \le n_i}$$

eine (geordnete) Basis von $\mathrm{Mult}_K(V_1, \ldots, V_r)$.

Beweis: Für jedes $f \in \mathrm{Mult}_K(V_1, \ldots, V_r)$ gilt

$$\begin{aligned} & f\big(\textstyle\sum_{j_1=1}^{n_1} \lambda_{j_1}^{(1)} b_{j_1}^{(1)}, \ldots, \sum_{j_r=1}^{n_r} \lambda_{j_r}^{(r)} b_{j_r}^{(r)}\big) \\ = \;& \textstyle\sum_{j_1}^{n_1} \cdots \sum_{j_r}^{n_r} \lambda_{j_1}^{(1)} \cdots \lambda_{j_r}^{(r)} f(b_{j_1}^{(1)}, \ldots, b_{j_r}^{(r)}) \\ = \;& \textstyle\sum_{j_1}^{n_1} \cdots \sum_{j_r}^{n_r} f(b_{j_1}^{(1)}, \ldots, b_{j_r}^{(r)}) \beta_{j_1}^{(1)} \otimes \cdots \otimes \beta_{j_r}^{(r)} \big(\sum_{j_1=1}^{n_1} \lambda_{j_1}^{(1)} b_{j_1}^{(1)}, \ldots, \sum_{j_r=1}^{n_r} \lambda_{j_r}^{(r)} b_{j_r}^{(r)}\big)\,. \end{aligned}$$

Somit spannt B^* den Raum $\mathrm{Mult}_K(V_1, \ldots, V_r)$ auf.
Nun zur linearen Unabhängigkeit: ist eine Linearkombination von B^* die Nullform, also

$$f = \sum_{j_1=1}^{n_1} \cdots \sum_{j_r=1}^{n_r} \lambda_{j_1,\ldots,j_r} \beta_{j_1}^{(1)} \otimes \cdots \otimes \beta_{j_r}^{(r)} = 0,$$

dann gilt

$$0 = f(b_{j_1}^{(1)}, \ldots, b_{j_r}^{(r)}) = \lambda_{j_1,\ldots,j_r}.$$

Somit ist B^* eine Menge linear unabhängiger Multilinearformen. □

Korollar: *Sind $V_1, \ldots, V_r$ endlichdimensionale Vektorräume über K, dann gilt*

$$\dim(\mathrm{Mult}_K(V_1, \ldots, V_r)) = \dim(V_1) \cdots \dim(V_r).$$

Definition 6.11: *Sei V ein n-dimensionaler Vektorraum über dem Körper K, sei $B = (b_1, \ldots, b_n)$ eine geordnete Basis von V, und sei $f : V \times V \to K$ eine Bilinearform auf $V \times V$. Dann ist die* **Darstellungsmatrix** *von f bzgl. der Basis B die Matrix*

$$A = (a_{ij})_{1 \le i \le n, 1 \le j \le n}, \quad \text{wobei} \quad a_{ij} = f(b_i, b_j).$$ □

Satz 6.12: *Sei V ein n-dimensionaler Vektorraum über K und sei $B = (b_1, \ldots, b_n)$ eine geordnete Basis von V.*
(i) *Ist $A = (a_{ij})$ die Darstellungsmatrix der Bilinearform $f : V \times V \to K$ bzgl. der geordneten Basis B, dann gilt für beliebige $x = \sum_{i=1}^n x_i b_i$ und $y = \sum_{i=1}^n y_i b_i$:*

$$f(x, y) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i a_{ij} y_j = (x_1, \ldots, x_n) \cdot A \cdot (y_1, \ldots, y_n)^T.$$

(ii) *Umgekehrt erzeugt jede Matrix* $A \in \mathrm{Mat}_{n\times n}(K)$ *eine Bilinearform auf* $V \times V$ *mittels*

$$f(x_1b_1 + \cdots + x_nb_n, y_1b_1 + \ldots + y_nb_n) \;=\; (x_1, \ldots, x_n) \cdot A \cdot (y_1, \ldots, y_n)^T.$$

(iii) *Die Zuordnung der Darstellungsmatrix zur Bilinearform* f *ist ein Isomorphismus zwischen* $\mathrm{Bil}_K(V, V)$ *und* $\mathrm{Mat}_{n\times n}(K)$.

Beweis: ad (i): Wir rechnen einfach aus:

$$f(x, y) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_iy_if(b_i, b_j) = \sum_{i,j=1}^{n} x_ia_{ij}y_j = (x_1, \ldots, x_n) \cdot A \cdot (y_1, \ldots, y_n)^T.$$

ad (ii): dass die so definierte Abbildung f eine Bilinearform ist, folgt sofort aus den Rechengesetzen für Matrizen.
ad (iii): die Zuordnung ist offensichtlich eine Bijektion; zu gegebenem f ist die Matrix A eindeutig bestimmt; aus A andererseits ist eindeutig eine Bilinearform bestimmt. Ist A die Darstellungsmatrix von f, und B die Darstellungsmatrix von g, so ist

$$(f + \lambda g)(x, y) = f(x, y) + \lambda g(x, y) = xAy^T + \lambda xBy^T = x(A + \lambda B)y^T.$$

Also die Darstellungsmatrix einer Linearkombination von Bilinearformen ist dieselbe Linearkombination der zugehörigen Darstellungsmatrizen. □

Beispiel 6.13: (i) Die Darstellungsmatrix des Skalarprodukts in Beispiel 6.2(i) bzgl. der kanonischen Basis von K^n ist die Einheitsmatrix.
(ii) Wir betrachten die Bilinearform auf $\mathbb{R}^3 \times \mathbb{R}^3$ mit $f(x, y) = x_1(y_2 + y_3) + x_2y_3$. Bzgl. der kanonischen Basis hat f die Darstellungsmatrix

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

(iii) Die Matrix

$$A \;=\; \begin{pmatrix} 1 & 2 & 5 \\ -2 & 0 & 1 \\ 0 & -6 & 6 \end{pmatrix}$$

erzeugt die Bilinearform

$$f(x, y) = (x_1, x_2, x_3) \cdot A \cdot \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = x_1y_1 + 2x_1y_2 + 5x_1y_3 - 2x_2y_1 + x_2y_3 - 6x_3y_2 + 6x_3y_3$$

auf $\mathbb{R}^3 \times \mathbb{R}^3$. □

Satz 6.14: *Sei* V *ein Vektorraum der Dimension* n *über* K. *Seien* $B = (b_1, \ldots, b_n)$ *und* $C = (c_1, \ldots, c_n)$ *geordnete Basen von* V, *sei* P *die Basistransformationsmatrix von* C *nach* B, *und sei* $f : V \times V \to K$ *eine Bilinearform mit Darstellungsmatrix* A *bzgl.* B. *Dann ist die Darstellungsmatrix von* f *bzgl.* C *gegeben durch* P^TAP.

Beweis: Sei $P = (p_{ij})$; vgl. Def. 3.3.27 und 3.3.32. Wir haben also

$$c_j = \sum_{i=1}^{n} p_{ij} b_i \quad \text{für } j = 1, \ldots, n.$$

Wegen der Bilinearität gilt daher

$$\begin{aligned} f(c_i, c_j) &= f\left(\sum_{r=1}^n p_{ri} b_r, \sum_{s=1}^n p_{sj} b_s\right) \\ &= \sum_{r=1}^n \sum_{s=1}^n p_{ri} p_{sj} f(b_r, b_s) \\ &= \sum_{r=1}^n \sum_{s=1}^n p_{ri} a_{rs} p_{sj} \\ &= [P^T AP]_{ij}. \end{aligned}$$

Somit ist $P^T AP$ die Darstellungsmatrix von f bzgl. C. □

Beispiel 6.15: Auf $\mathbb{R}^3$ betrachten wir die Bilinearform f gegeben durch

$$2x_1y_1 - 3x_2y_3 + x_3y_3.$$

Bzgl. der kanonischen Basis hat f die Darstellungsmatrix

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & -3 \\ 0 & 0 & 1 \end{pmatrix}.$$

Um die Darstellungsmatrix von f bzgl. der geordneten Basis

$$C = (\,(1,1,1), (-2,1,1), (2,1,0)\,)$$

zu bestimmen, verwenden wir die Transformationsmatrix von C in die Standardbasis:

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

Die Darstellungmatrix bzgl. C ist dann

$$P^T AP = \begin{pmatrix} 0 & -6 & 4 \\ -6 & 6 & -8 \\ 1 & -11 & 8 \end{pmatrix}.$$ □

In Definition 2.2.3 haben wir eine (nicht notwendig quadratische) Matrix B zeilenäquivalent zur (nicht notwendig quadratischen) Matrix A genannt, wenn $B = P \cdot A$ für eine reguläre Matrix P. Ebenso haben wir eine Matrix B spaltenäquivalent zu A genannt, wenn $B = A \cdot Q$ für eine reguläre Matrix Q. Daraus ergibt sich auf natürliche Weise der Begriff der "Äquivalenz": A und B heissen **äquivalent**, geschrieben $A \sim B$, g.d.w.

$$B = P \cdot A \cdot Q$$

für reguläre Matrizen P und Q. Äquivalente Matrizen haben denselben Rang und dieselbe Hermite Normalform.

In Def. 3.3.39 haben wir zwei $n \times n$ Matrizen A und B **ähnlich** genannt, wenn es eine reguläre Matrix P gibt, sodass

$$B = P^{-1} \cdot A \cdot P \ .$$

Ähnliche Matrizen stellen dieselbe lineare Funktion dar, bzgl. verschiedener Basen.

Der Satz 6.14 begründet die folgende Definition von "Kongruenz".

Definition 6.16: *Seien $A, B \in \mathrm{Mat}_{n\times n}(K)$. A und B heissen* **kongruent**, *geschrieben $A \simeq B$, gdw. es eine reguläre Matrix P gibt, sodass $B = P^T AP$.* □

Satz 6.17: *Die Relation "$\simeq$" (Kongruenz) ist eine Äquivalenzrelation auf $\mathrm{Mat}_{n\times n}(K)$, welche feiner ist als die Relation "$\sim$" (Äquivalenz) auf $\mathrm{Mat}_{n\times n}(K)$.*

Beweis: Offensichtlich gilt $A \simeq A$.
$A \simeq B$ gdw $B = P^T AP$ gdw $BP^{-1} = P^T A$ gdw $(P^{-1})^T BP^{-1} = (P^T)^{-1}BP^{-1} = A$ gdw $B \simeq A$.
Aus $A \simeq B \simeq C$ ergibt sich

$$C = Q^t BQ = Q^T P^T APQ = (PQ)^T APQ,$$

also $A \simeq C$.
Somit ist "$\simeq$" eine Äquivalenzrelation.
Mit P ist natürlich auch P^T regulär, also ist "$\simeq$" eine Verfeinerung von "$\sim$". □

Definition 6.18: *Eine Bilinearform $f : V \times V \to K$ heisst* **symmetrisch** *gdw. $f(x, y) = f(y, x)$ für alle $x, y \in V$ gilt.*
f heisst **schief-symmetrisch** *gdw. $f(x, y) = -f(y, x)$ für alle $x, y \in V$ gilt.* □

Satz 6.19: *Jede Bilinearform $f : V \times V \to K$ kann auf eindeutige Weise zerlegt werden in die Summe $f = g + h$ einer symmetrischen Bilinearform g und einer schief-symmetrischen Bilinearform h. (Dabei müssen wir annehmen, dass $1/2$ in K existiert, also $1 + 1 \neq 0$)*

Beweis: Sei

$$\begin{array}{rcl} g(x,y) & := & \frac{1}{2}\big(f(x,y) + f(y,x)\big), \\ h(x,y) & := & \frac{1}{2}\big(f(x,y) - f(y,x)\big). \end{array}$$

Dann ist offensichtlich g symmetrisch, h schief-symmetrisch, und es gilt $f = g + h$.
Nun zur Eindeutigkeit: lässt sich f auch schreiben als $f = g' + h'$, wobei g' symmetrisch und h' schief-symmetrisch ist, dann gilt

$$g - g' = h' - h \ .$$

Dabei ist die linke Seite symmetrisch, und die rechte Seite schief-symmetrisch. Das ist nur möglich für $g - g' = 0 = h' - h$, also $g = g'$ und $h = h'$. □

Ist V endlichdimensional, dann ergibt sich Satz 6.19 aus dem Satz 2.1.25 über die Zerlegung einer quadratischen Matrix in die Summe einer symmetrischen und einer schief-symmetrischen Matrix.

Beispiel 6.20: Die Bilinearform

$$f((x_1, x_2), (y_1, y_2)) = x_1y_1 + x_1y_2 + 2x_2y_1 + x_2y_2$$

auf $\mathbb{R}^2 \times \mathbb{R}^2$ lässt sich schreiben als $f = g + h$, wobei g die symmetrische Bilinearform

$$\begin{aligned}
g((x_1,x_2),(y_1,y_2)) &= \tfrac{1}{2}\big(f((x_1,x_2),(y_1,y_2) + f((y_1,y_2),(x_1,x_2))\big) \\
&= \tfrac{1}{2}\big(x_1y_1 + x_1y_2 + 2x_2y_1 + x_2y_2y_1x_1 + y_1x_2 + 2y_2x_1 + y_2x_2\big) \\
&= \tfrac{1}{2}\big(2x_1y_1 + 3x_1y_2 + 3x_2y_1 + 2x_2y_2\big)
\end{aligned}$$

ist, und h die schief-symmetrische Bilinearform

$$\begin{aligned}
h((x_1,x_2),(y_1,y_2)) &= \tfrac{1}{2}\big(f((x_1,x_2),(y_1,y_2) - f((y_1,y_2),(x_1,x_2))\big) \\
&= \tfrac{1}{2}\big(x_1y_1 + x_1y_2 + 2x_2y_1 + x_2y_2 - y_1x_1 - y_1x_2 - 2y_2x_1 - y_2x_2\big) \\
&= \tfrac{1}{2}\big(-x_1y_2 + x_2y_1\big)
\end{aligned}$$

ist. □

Lemma 6.21: *Die Darstellungsmatrix einer symmetrischen Bilinearform ist symmetrisch, und umgekehrt erzeugt eine symmetrische Matrix eine symmetrische Bilinearform.*

Beweis: das sieht man aus der Betrachtung einer Basis: $f(b_i, b_j) = f(b_j, b_i)$. □

Definition 6.22: *Sei V ein Vektorraum über K, und sei $f : V \times V \to K$ eine symmetrische Bilinearform. Dann ist die Abbildung*

$$q : \begin{array}{ccc} V & \longrightarrow & K \\ x & \mapsto & f(x,x) \end{array}$$

eine **quadratische Form** *auf V. Diese Abbildung q heisst auch* **die zu f gehörige quadratische Form**, *geschrieben q_f.* □

Beispiel 6.23: Die Abbildung

$$q : \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \mapsto & x^2 - xy + y^2 \end{array}$$

ist eine quadratische Form auf $\mathbb{R}^2$. Das sieht man auch aus der Darstellung

$$q(x,y) = (x,y) \cdot \begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}.$$

Die Darstellungsmatrix bzgl. der kanonischen Basis ist also symmetrisch. □

Satz 6.24: *Sei V ein Vektorraum über $\mathbb{R}$ und sei $f : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform auf V (wie in Satz 6.19 nehmen wir an, dass $1/2$ in K existiert, also $1+1 \neq 0$). Dann gilt folgendes:*

(i) $q_f(\lambda x) = \lambda^2 q_f(x)$;

(ii) $f(x,y) = \frac{1}{2}\big(q_f(x+y) - q_f(x) - q_f(y)\big)$;

(iii) $f(x,y) = \frac{1}{4}\big(q_f(x+y) - q_f(x-y)\big)$.

Beweis: ad (i): $q_f(\lambda x) = f(\lambda x, \lambda x) = \lambda^2 f(x,x) = \lambda^2 q_f(x)$.
ad (ii): weil f symmetrisch ist, haben wir

$$q_f(x+y) = f(x+y, x+y) = f(x,x) + f(x,y) + f(y,x) + f(y,y) = q_f(x) + 2f(x,y) + q_f(y).$$

Daraus folgt unmittelbar die Behauptung.
ad (iii): wegen (i) haben wir $q_f(-x) = q_f(x)$, und daher wegen (ii)

$$q_f(x-y) = q_f(x) - 2f(x,y) + q_f(y).$$

Subtrahiert man diese Beziehung von (ii), so ergibt sich (iii). □

Satz 6.25: *Jede reelle quadratische Form $q : V \to \mathbb{R}$ (V ein Vektorraum über $\mathbb{R}$) gehört zu einer eindeutig bestimmten symmetrischen Bilinearform.*

Beweis: Sei $q : V \to \mathbb{R}$ eine reelle quadratische Form. Wir nehmen an, dass $q_f = q = q_g$ für zwei symmetrische Bilinearformen $f, g : V \times V \to \mathbb{R}$. Dann ergibt sich aus Satz 6.24 sofort $f = g$. □

Dieser Satz gibt uns die Berechtigung für die folgende Definition.

Definition 6.26: *Sei $q : V \to \mathbb{R}$ eine reelle quadratische Form auf dem endlichdimensionalen reellen Vektorraum V. Die* **Darstellungsmatrix** *von q ist die Darstellungsmatrix der zugehörigen symmetrischen Bilinearform.* □

Beispiel 6.27: Die Abbildung $q : \mathbb{R}^2 \to \mathbb{R}$ gegeben durch

$$q(x,y) \;=\; 4x^2 + 6xy + 9y^2 \;=\; (x,y) \cdot \begin{pmatrix} 4 & 3 \\ 3 & 9 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}$$

ist eine quadratische Form. Die Darstellungsmatrix von q ist

$$\begin{pmatrix} 4 & 3 \\ 3 & 9 \end{pmatrix}$$

und die zugehörige Bilinearform ist

$$f\big((x,y),(x',y')\big) \;=\; 4xx' + 3(xy' + x'y) + 9yy'.$$ □

Wegen Satz 6.14 repräsentieren symmetrische Matrizen A und B dieselbe quadratische Form bzgl. verschiedener Basen genau dann, wenn sie kongruent sind. Im folgenden wollen wir nun eine besonders günstige Darstellung einer quadratischen Form unter Kongruenztransformationen finden.

Natürlich haben kongruente Matrizen denselben Rang, da durch Multiplikation mit einer regulären Matrix (welche ja ein Produkt von Elementarmatrizen ist) der Rang nicht verändert wird.

Satz 6.28: *Sei $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$, und A symmetrisch. Dann ist A kongruent zu einer eindeutig bestimmten Matrix der Form*

$$\tilde{A} \;=\; \begin{pmatrix} I_r & & \\ & -I_s & \\ & & 0 \end{pmatrix}.$$

Beweis: Eine zu A kongruente Matrix hat die Form $B = P^T A P$, für eine reguläre Matrix P. P lässt sich schreiben als Produkt von Elementarmatrizen, insbesondere Spaltenumformungen. P^T ist also das entsprechende Produkt von Zeilenumformungen. Durch gleiche

Zeilen- und Spaltenumformungen lässt sich eine symmetrische Matrix diagonalisieren und die Diagonalelemente auch normieren zu ± 1. Es lässt sich aber nicht das Vorzeichen in der Diagonale ändern. □

Aus Satz 6.28 ergibt sich unmittelbar der folgende Satz von Sylvester [1].

Satz 6.29: (Sylvester) *Sei V ein Vektorraum der Dimension n über $\mathbb{R}$ und sei $q : V \to \mathbb{R}$ eine quadratische Form auf V. Dann gibt es eine geordnete Basis $(b_1, \ldots, b_n)$ von V, und natürliche Zahlen r, s mit $0 \leq r + s \leq n$, sodass für jedes $x = \sum_{i=1}^n x_i b_i$ gilt*

$$q(x) \;=\; x_1^2 + \cdots + x_r^2 - x_{r+1}^2 - \cdots - x_{r+s}^2.$$

Diese Zahlen r und s hängen nicht von einer solchen Basis ab.

Definition 6.30: *Sei q eine reelle quadratische Form über dem endlichdimensionalen Vektorraum V, wie in Satz 6.29. Dann heisst die Darstellung von q wie in Satz 6.29 (bzgl. einer geeigneten Basis) die* **kanonische Form** *von q. Weiters heisst $r + s$ der* **Rang** *der quadratischen Form q, r heisst der* **Index** *von q und $r - s$ heisst die* **Signatur** *von q.* □

Beispiel 6.31: Wir betrachten die quadratische Form

$$\begin{array}{rlcl} q: & \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ & (x,y,z) & \mapsto & x^2 - 2xy + 4yz - 2y^2 + 4z^2 \;. \end{array}$$

Die Darstellungsmatrix von q (bzgl. der kanonischen Basis) ist die symmetrische Matrix

$$A \;=\; \begin{pmatrix} 1 & -1 & 0 \\ -1 & -2 & 2 \\ 0 & 2 & 4 \end{pmatrix} \;.$$

Um die kanonische Form von q zu bestimmen, könnte man wie in Satz 10.28 eine geeignete Basis bestimmen. Man kann aber auch die Methode der "Vervollständigung der Quadrate" anwenden, wie wir an diesem Beispiel demonstrieren:

$$q(x,y,z) = (x-y)^2 + 4yz - 3y^2 + 4z^2 = (x-y)^2 + (y+2z)^2 - 4y^2.$$

Also für

$$\alpha = x - y, \quad \beta = y + 2z, \quad \gamma = 2y \qquad (1)$$

bzw.

$$x = \alpha + \frac{1}{2}\gamma, \quad y = \frac{1}{2}\gamma, \quad z = \frac{1}{2}\beta - \frac{1}{4}\gamma \qquad (2)$$

lässt sich q schreiben als

$$q(\alpha, \beta, \gamma) = \alpha^2 + \beta^2 - \gamma^2.$$

Aus (1) lesen wir die Transformationsmatrix P ab:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \;=\; P \cdot \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} \;=\; \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{4} \end{pmatrix} \cdot \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} \;.$$

Bzgl. der Basis

$$\big(\, (1,-1,0), (0,1,2), (0,2,0) \,\big)$$

[1] James Joseph Sylvester, 1814–1897

hat die quadratische Form q die Darstellung

$$\tilde{A} \;=\; P^T \cdot A \cdot P \;=\; \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \; .$$

Die kanonische Form von q ist also

$$q(x_1, x_2, x_3) = x_1^2 + x_2^2 - x_3^2,$$

der Rang von q ist 3, der Index ist 2 und die Signatur ist 1. □

Definition 6.32: *Sei V ein Vektorraum über $\mathbb{R}$ und $q : V \to \mathbb{R}$ eine quadratische Form auf V.*

(i) *q heisst* **positiv definit** *gdw. $\forall x \in V \setminus \{0\} : q(x) > 0$;*

(ii) *q heisst* **positiv semidefinit** *gdw. $\forall x \in V \setminus \{0\} : q(x) \geq 0$;*

(iii) *q heisst* **negativ definit** *gdw. $\forall x \in V \setminus \{0\} : q(x) < 0$;*

(iv) *q heisst* **negativ semidefinit** *gdw. $\forall x \in V \setminus \{0\} : q(x) \leq 0$.*

In allen anderen Fällen heisst q **indefinit**.
Ebenso nennt man eine symmetrische Matrix $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ positiv defninit, positiv semidefinit, negativ definit, negativ semidefinit gdw.

$$\forall x \in V \setminus \{0\} : xAx^T >, \geq, <, \leq 0$$

gilt. □

Satz 6.33: *Sei $q : V \to \mathbb{R}$ eine reelle quadratische Form auf einem n-dimensionalen Vektorraum. Sei A die Darstellungmatrix von q. Sei $\tilde{A}$ die zu A kongruente Matrix von der Form wie in Satz 6.28, also mit r mal der 1 und s mal der -1 in der Diagonale. Dann ist q (bzw. A)*

(i) *positiv definit gdw. $r = n$ und $s = 0$;*

(ii) *positiv semidefinit gdw. $r \leq n$ und $s = 0$;*

(iii) *negativ definit gdw. $r = 0$ und $s = n$;*

(iv) *negativ semidefinit gdw. $r = 0$ und $s \leq n$.*

Beweis: Sei

$$q(x) = x_1^2 + \ldots + x_r^2 - x_{r+1}^2 - \ldots - x_{r+s}^2, \qquad r + s \leq n$$

die kanonische Form von q. Daraus liest man die Behauptungen unmittelbar ab.
Ausserdem gilt

$$\forall x \in V \setminus \{0\} : xAx^T > 0 \quad \Longleftrightarrow \quad \forall x \in V \setminus \{0\} : xP^TAPx^T > 0 \; ,$$

wenn P eine reguläre Matrix ist. Eine analoge Beziehung gilt für $\geq, <, \leq$. □

Ohne Beweis geben wir noch folgendes Kriterium für die Definitheit an.

Satz 6.35: *Sei q eine reelle quadratische Form auf dem n-dimensionalen Vektorraum V und $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ die zu q gehörige symmetrische Darstellungsmatrix. Dann ist q (bzw. A)*

(i) *positiv definit gdw. alle Hauptminoren von A positiv sind;*

(ii) *negativ definit gdw. die Hauptminoren von A abwechselnd negativ und positiv sind, also $M_1 < 0, M_2 > 0, M_3 < 0$, etc.*

Beispiel 6.36: Wir betrachten die quadratische Form aus Beispiel 10.31, also

$$\begin{array}{rlcl} q: & \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ & (x,y,z) & \mapsto & x^2 - 2xy + 4yz - 2y^2 + 4z^2 \ , \end{array}$$

mit der Darstellungsmatrix

$$A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & -2 & 2 \\ 0 & 2 & 4 \end{pmatrix} .$$

Die kanonische Form von q ist

$$q(x,y,z) = (x-y)^2 + (y+2z)^2 - 4y^2,$$

und die zu A kongruente kanonische Matrix ist

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

Die quadratische Form q ist also weder positv noch negativ definit. Insbesondere gilt

$$q(1,1,1) = 5, \quad q(1,3,1) = -7 \ .$$

Ändern wir nun die quadratische Form zu

$$p(x) := q(x) + 5y^2 = x^2 - 2xy + 4yz + 3y^2 + 4z^2 \ ,$$

so hat die neue quadratische Form p die Darstellungsmatrix

$$\begin{pmatrix} 1 & -1 & 0 \\ -1 & 3 & 2 \\ 0 & 2 & 4 \end{pmatrix} .$$

Die Hauptminoren sind $M_1 = 1, M_2 = 2, M_3 = 4$, die quadratische Form p ist also positiv definit. □

Satz 6.37: *Sei $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ eine symmetrische Matrix, sei $\lambda \in \mathbb{C}$ ein Eigenwert von A, und sei $z = x + iy \in \mathbb{C}_n$ ($x, y \in \mathbb{R}_n$) ein Eigenvektor zum Eigenwert λ. Dann ist λ reell, und sowohl x als auch y sind (falls verschieden vom Nullvektor) reelle Eigenvektoren von A zum Eigenwert λ.*

Beweis: Mit $\overline{z}$ bezeichnen wir den zu z komplex konjugierten Vektor.
Es gilt also $A \cdot z = \lambda \cdot z$, und weiters

$$\overline{z}^T A z = \overline{z}^T \lambda z = \lambda \overline{z}^T z \qquad (1)$$

Ausserdem

$$\overline{Az} = \overline{A}\overline{z} = A\overline{z} \quad \text{und} \quad \overline{Az} = \overline{\lambda z} = \overline{\lambda}\overline{z} \quad (2)$$

Somit erhalten wir

$$\begin{aligned} \lambda\overline{z}^T z \ &=_{(1)} \overline{z}^T Az = (\overline{z}^T Az)^{TT} = (z^T A\overline{z})^T = \quad \text{(weil } 1 \times 1\text{-Matrix)} \\ &= z^T A\overline{z} =_{(2)} z^T \overline{Az} =_{(2)} z^T \overline{\lambda}\overline{z} = \overline{\lambda} z^T \overline{z} = \overline{\lambda}\overline{z}^T z \ . \end{aligned}$$

Da laut Voraussetzung $\overline{z}^T z = \langle z|z\rangle \neq 0$, haben wir $\lambda = \overline{\lambda}$, also $\lambda \in \mathbb{R}$.
Aus

$$Ax + iAy = Az = \lambda z = \lambda x + i\lambda y$$

und weil sowohl A, als auch x und y reell sind, erhalten wir

$$Ax = \lambda x \quad \text{und} \quad Ay = \lambda y.$$

Sowohl x als auch y sind also (falls verschieden vom Nullvektor) reelle Eigenvektoren von A zum Eigenwert λ. □

Wir zeigen nun noch, wie man mittels Eigenwerten und -vektoren auf der Einheitskugel in $\mathbb{R}_n$ das Maximum einer quadratischen Form bestimmen kann. Die Einheitskugel in $\mathbb{R}_n$ besteht aus allen Vektoren mit (Euklidscher) Norm 1:

$$S_n \ := \ \{x \in \mathbb{R}_n \mid \ \| x \| = \sqrt{x_1^2 + \ldots + x_n^2} = 1\}.$$

Die Einheitskugel S_n ist beschränkt und abgeschlossen, also kompakt. Eine stetige Funktion besitzt also auf S_n ein Maximum.

Satz 6.38: *Sei $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ eine symmetrische Matrix, sei $q(x) = x^T Ax$ die zugehörige quadratische Form auf $\mathbb{R}_n$ (Spaltenvektoren). Sei $s \in S_n$ so, dass $q(s)$ das Maximum von q auf S_n ist, also*

$$q(s) = \max\{q(t) \mid t \in S_n\} \ .$$

Dann ist s ein Eigenvektor von A; d.h. es gibt ein $\lambda \in \mathbb{R}$ mit $As = \lambda s$.
Beweis: Sei also $s \in S_n$ so, dass $q(s)$ das Maximum von q auf S_n ist. Sei $W = \{s\}^\perp$, also der Orthogonalraum zu s. Die Dimension von W ist $n - 1$. Sei $w \in W$ mit $\| w \| = 1$. Wir betrachten die Kurve

$$C(t) = (\cos t)s + (\sin t)w.$$

Die Tangenten an die Einheitskugel S_n bei s sind die Einheitsvektoren $w \in W$. Die Kurve $C(t)$ liegt auf der Einheitskugel, denn $\| C(t) \| = 1$:

$$\begin{aligned} \| C(t) \| \ &= \sqrt{\langle C(t)|C(t)\rangle} \\ &= \sqrt{\langle (\cos t)s + (\sin t)w|(\cos t)s + (\sin t)w\rangle} \\ &= \sqrt{\cos^2 t\langle s|s\rangle + (\cos t)(\sin t)\langle s|w\rangle + (\sin t)(\cos t)\langle w|s\rangle + \sin^2 t\langle w|w\rangle} \\ &= \sqrt{\cos^2 t\langle s|s\rangle + \sin^2 t\langle w|w\rangle} \\ &= 1 \end{aligned}$$

Ausserdem ist $C(0) = s$, die Kurve $C(t)$ geht also durch den Punkt s. Die Ableitung von $C(t)$ ist

$$C'(t) = (-\sin t)s + (\cos t)w,$$

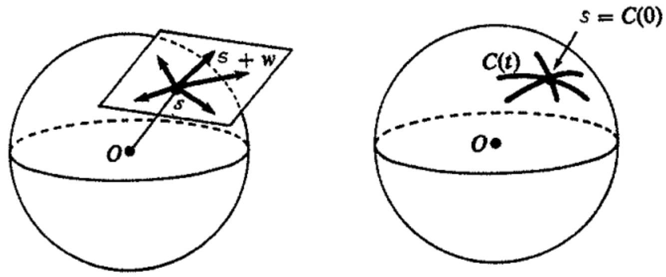

Figure 1: Kurve $C(t) = (\cos t)s + (\sin t)w$

also $C'(0) = w$. Am Punkt s geht die Kurve also in Richtung w.
Nun betrachten wir die Funktion

$$g(t) = q(C(t)) = C(t)^T \cdot A \cdot C(t).$$

Die Ableitung von g ist wegen der Symmetrie von A

$$g'(t) = C'(t)^T \cdot A \cdot C(t) + C(t)^T \cdot A \cdot C'(t) = 2\,C'(t)^T \cdot A \cdot C(t).$$

Ausserdem

$$g'(t) = q'(C(t)) \cdot C'(t).$$

Da $q(s)$ ein Maximum ist, und $g(0) = q(s)$, haben wir $g'(0) = q'(s) \cdot C'(0) = 0$.
Daraus erhalten wir

$$0 = g'(0) = 2\,C'(0)^T \cdot A \cdot C(0) = 2\,w^T As.$$

Also ist As orthogonal zu allen $w \in W$. Nun ist aber $W^\perp = \text{span}\{s\}$. Es gibt also ein $\lambda \in \mathbb{R}$ mit $As = \lambda s$. □

Korollar: *Das Maximum von q auf der Einheitskugel S_n ist der grösste Eigenwert von A.*

Beweis: Für jeden Eigenwert λ und jeden zugehörigen Eigenvektor s auf der Einheitskugel ($\| s \| = 1$) gilt

$$q(s) = s^T As = s^T \lambda s = \lambda s^T s = \lambda.$$

Der Wert von q bei einem Eigenvektor auf der Einheitskugel ist also genau der zugehörige Eigenwert.
Satz 6.38 besagt, dass das Maximum von q auf der Einheitskugel bei einem Eigenvektor angenommen wird. Daher ist das Maximum von q auf der Einheitskugel gleich dem grössten Eigenwert. □

Beispiel 6.39: Wir betrachten die quadratische Form $q(x, y) = 2x^2 - 3xy + y^2$ auf $\mathbb{R}_2$.
Die Darstellungmatrix von q ist

$$A = \begin{pmatrix} 2 & -\frac{3}{2} \\ -\frac{3}{2} & 1 \end{pmatrix}.$$

Wir suchen das Maximum von q auf dem Einheitskreis S_2.
Das charakteristische Polynom von A ist $x^2 - 3x - 1/4$, die Eigenwerte sind also

$$\lambda_{1,2} \;=\; \frac{3 \pm \sqrt{10}}{2} \;.$$

Die zugehörigen Eigenvektoren sind

$$v(\lambda_i) \;=\; \begin{pmatrix} 1 \\ \frac{2}{3}(2 - \lambda_i) \end{pmatrix} \;.$$

Die Eigenvektoren auf dem Einheitskreis sind

$$s(\lambda_i) \;=\; \frac{v(\lambda_i)}{\| \, v(\lambda_i) \, \|} \;.$$

Das Maximum von q auf dem Einheitskreis ist der grösste Eigenwert, also

$$\max\{q(t) \mid t \in S_2\} \;=\; \frac{3 + \sqrt{10}}{2} \;=\; \lambda_1 \;,$$

und dieses Maximum wird angenommen am Punkt $s(\lambda_1)$.
Ebenso wird das Minimum $\lambda_2 = \frac{3-\sqrt{10}}{2}$ angenommen am Punkt $s(\lambda_2)$. $\square$