Lineare Algebra und Analytische Geometrie 2 für Lehramtsstudierende

Ein Skriptum zur Vorlesung
im Sommersemester 2012

Arne Dür und Franz Pauer

1. Auflage

Vorwort

Das vorliegende Skriptum soll den Hörerinnen und Hörern der Vorlesung „Lineare Algebra und Analytische Geometrie 2 für Lehramtsstudierende“ im Sommersemester 2012 das Mitschreiben und *Mitdenken* in der Vorlesung erleichtern. Das Skriptum enthält alle Definitionen und Sätze der Vorlesung, aber nur wenige Beispiele dazu. In der Vorlesung werden die Definitionen und Sätze motiviert, der Zusammenhang mit früheren Ergebnissen erläutert und Beispiele dazu besprochen. Der Inhalt des Skriptums „Lineare Algebra 1“ und der schriftlichen Unterlagen zur Vorlesung „Vertiefung Lineare Algebra 1“ von Franz Pauer wird als bekannt vorausgesetzt.

In den Kapiteln 1 und 4 werden die Themen Lineare Funktion, Determinante und Skalarprodukt aus dem Wintersemester weitergeführt und vertieft. Im Kapitel 2 werden Systeme linearer Ungleichungen und - als Anwendung davon - lineare Optimierung mit zwei Unbekannten besprochen. Im Kapitel 3 werden die Bewegungen starrer Körper in der Ebene und im Raum beschrieben, dieses Thema ist für die Mechanik (und Robotik) von großer Bedeutung. Im Kapitel 5 werden quadratische Funktionen und ihre Nullstellenmengen, die Quadriken, besprochen.

Dieses Skriptum umfasst (in leicht veränderter Form) Teile der folgenden Skripten:
Arne Dür und Franz Pauer: Lineare Algebra (5. Auflage), 2006.
Arne Dür und Franz Pauer: Analytische Geometrie (3. Auflage), 2005.
Franz Pauer: Lineare Optimierung (2. Auflage), 2003.

Innsbruck, Februar 2012

Inhaltsverzeichnis

Vorwort iii

Kapitel 1. Mehr über lineare Funktionen 1
§1. Der Graph einer linearen Funktion 1
§2. Basiswechsel 3
§3. Bild und Kern einer linearen Funktion 6
§4. Systeme linearer Gleichungen in koordinatenfreier Form 9
§5. Affine Funktionen 12

Kapitel 2. Systeme linearer Ungleichungen 15
§1. Wiederholung: Zwei Zugänge zur Geometrie der Ebene 15
§2. Lineare Ungleichungen und Halbräume 17
§3. Lineare Optimierung mit zwei Unbekannten 22

Kapitel 3. Bewegungen in euklidischen Räumen 28
§1. Isometrien 28
§2. Orthogonale Funktionen 31
§3. Spiegelungen 33
§4. Isometrien der Ebene 37
§5. Zeigerrechnung 42
§6. Isometrien des Raumes 45
§7. Symmetriegruppen 48

Kapitel 4. Multilineare Funktionen 50
§1. Multilineare Funktionen 50
§2. Alternierende Funktionen 53
§3. Entwicklung von Determinanten 57
§4. Die adjungierte Matrix 60
§5. Symmetrische Bilinearformen 61
§6. Positiv definite Matrizen 65

Kapitel 5. Quadratische Funktionen und Quadriken 68
§1. Linearformen 68
§2. Quadratische Formen 71
§3. Quadratische Funktionen 74
§4. Quadriken 77
§5. Quadriken in der Ebene 78

KAPITEL 1

Mehr über lineare Funktionen

In diesem Kapitel sei K ein Körper.

§1. Der Graph einer linearen Funktion

Satz 1: *Seien $V_1,\dots,V_\ell$ Vektorräume über K. Dann wird das kartesische Produkt*

$$V_1\times\cdots\times V_\ell=\{(x_1,\dots,x_\ell)\mid x_1\in V_1,\dots,x_\ell\in V_\ell\}$$

mit der komponentenweisen Addition

$$(x_1,\dots,x_\ell)+(y_1,\dots,y_\ell):=(x_1+y_1,\dots,x_\ell+y_\ell)$$

und der komponentenweisen Skalarmultiplikation

$$c(x_1,\dots,x_\ell):=(cx_1,\dots,cx_\ell)$$

mit $c\in K$ ein Vektorraum und heißt der Produktraum *von $V_1,\dots,V_\ell$.*
Für alle $j\in I$ ist die Projektion auf den j-ten Faktor

$$\mathrm{pr}_j:V_1\times\cdots\times V_\ell\to V_j\,,\ (x_1,\dots,x_\ell)\mapsto x_j\,,$$

K-linear. Wenn $(v_{11},\dots,v_{1n_1}),\dots,(v_{\ell 1},\dots,v_{\ell n_\ell})$ Basen von $V_1,\dots,V_\ell$ sind, dann ist

$$((v_{11},0,\dots,0),\dots,(v_{1n_1},0,\dots,0),\dots$$

$$\dots,((0,\dots,0,v_{\ell 1}),\dots,(0,\dots,0,v_{\ell n_\ell}))$$

eine Basis von $V_1\times\cdots\times V_\ell$, insbesondere gilt

$$\dim_K(V_1\times\cdots\times V_\ell)=\dim_K(V_1)+\cdots+\dim_K(V_\ell)\,.$$

Beweis: Es ist leicht zu zeigen, dass $V_1\times\cdots\times V_\ell$ mit der komponentenweisen Addition und Skalarmultiplikation ein Vektorraum ist und die Projektionen linear sind (Übung). Wir beweisen daher nur, dass $((v_{11},0,\dots,0),\dots,(0,\dots,0,v_{\ell n_\ell}))$ eine Basis von $V_1\times\cdots\times V_\ell$ ist. Wir schreiben $x_1\in V_1,\dots,x_\ell\in V_\ell$ als Linearkombinationen der Basen $(v_{11},\dots,v_{1n_1}),\dots,(v_{\ell 1},\dots,v_{\ell n_\ell})$:

$$x_1=\sum_{i=1}^{n_1}d_{1i}v_{1i},\ \dots,x_\ell=\sum_{i=1}^{n_\ell}d_{\ell i}v_{\ell i}\,.$$

Dann ist

$$\begin{aligned}(x_1,\dots,x_\ell) &= (x_1,0,\dots,0)+\dots+(0,\dots,0,x_\ell)\\ &= \sum_{i=1}^{n_1} d_{1i}(v_{1i},0,\dots,0)+\dots+\sum_{i=1}^{n_\ell} d_{\ell i}(0,\dots,0,v_{\ell i})\end{aligned}$$

und $((v_{11},0,\dots,0),\dots,(0,\dots,0,v_{\ell n_\ell}))$ ein Erzeugendensystem von $V_1\times\dots\times V_\ell$. Um die lineare Unabhängigkeit zu zeigen, seien $c_{11},\dots,c_{\ell n_\ell}\in K$ mit

$$\sum_{i=1}^{n_1} c_{1i}(v_{1i},0,\dots,0)+\dots+\sum_{i=1}^{n_\ell} c_{\ell i}(0,\dots,0,v_{\ell i}) = (0,\dots,0)\,.$$

Dann ist

$$(\sum_{i=1}^{n_1} c_{1i}v_{1i},\dots,\sum_{i=1}^{n_\ell} c_{\ell i}v_{\ell i}) = (0,\dots,0)\,,$$

also

$$\sum_{i=1}^{n_1} c_{1i}v_{1i}=0\ ,\ \dots\ ,\ \sum_{i=1}^{n_\ell} c_{\ell i}v_{\ell i}=0\,.$$

Da $(v_{11},\dots,v_{1n_\ell}),\dots,(v_{\ell 1},\dots,v_{\ell n_\ell})$ Basen von $V_1,\dots,V_\ell$ sind, folgt $c_{11}=\dots=c_{\ell n_\ell}=0$, was zu zeigen war.

Satz 2: *Es seien V und W Vektorräume über K und $(v_i)_{i\in I}$ eine (beliebige) Basis von V.*

Eine Funktion $f:V\to W$ ist genau dann linear, wenn der Graph von f ein Untervektorraum des Produktraums $V\times W$ ist.

In diesem Fall hat der Graph von f die Basis $((v_i,f(v_i)))_{i\in I}$. Insbesondere gilt

$$\dim_K(\mathrm{Graph}(f)) = \dim_K(V)\,.$$

Beweis: Nach Definition ist $\mathrm{Graph}(f)=\{(v,f(v))\mid v\in V\}\subset V\times W$. Seien $u,w\in V$ und $c\in K$. Wenn f linear ist, dann ist
$0_{V\times W}=(0_V,0_W)=(0_V,f(0_V))\in\mathrm{Graph}(f)$,
$(u,f(u))+(w,f(w))=(u+w,f(u)+f(w))=(u+w,f(u+w))\in\mathrm{Graph}(f)$
und
$c(w,f(w))=(cw,cf(w))=(cw,f(cw))\in\mathrm{Graph}(f)$,
also $\mathrm{Graph}(f)$ ein Untervektorraum von $V\times W$. Wenn umgekehrt $\mathrm{Graph}(f)$ ein Untervektorraum von $V\times W$ ist, dann sind
$(u,f(u))+(w,f(w))=(u+w,f(u)+f(w))\in\mathrm{Graph}(f)$ und
$c(w,f(w))=(cw,cf(w))\in\mathrm{Graph}(f)$, somit
$f(u+w)=f(u)+f(w)$ und $f(cw)=cf(w)$, also f linear.
Wenn f linear ist, dann ist auch die Funktion

$$F:V\to\mathrm{Graph}(f)\,,\,x\mapsto(x,f(x))\,,$$

linear und hat die Umkehrfunktion $\mathrm{Graph}(f) \to V$, $(x, f(x)) \mapsto x$. Daher ist F ein Isomorphismus und $(F(v_i))_{i\in I}$ eine Basis von $\mathrm{Graph}(f)$.

Beispiel 3 : Es sei k eine reelle Zahl und f die lineare Funktion $f : \mathbb{R} \longrightarrow \mathbb{R}$, $z \longmapsto kz$. Dann ist

$$\mathrm{Graph(f)} = \{(z, kz)|z \in \mathbb{R}\} = \{z(1, k)|z \in \mathbb{R}\} = \mathbb{R}(1, k) \subseteq \mathbb{R} \times \mathbb{R}$$

die Gerade durch $(0, 0)$ und $(1, k)$.

Beispiel 4 : Es seien a, b reelle Zahlen und g die lineare Funktion $g : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $(x, y) \longmapsto ax + by$. Dann ist

$$\mathrm{Graph(g)} = \{(x, y, ax+by)\,|\,x, y \in \mathbb{R}\} = \{x\cdot(1, 0, a) + y\cdot(0, 1, b)\,|\,x, y \in \mathbb{R}\} =$$

$$= \mathbb{R}(1, 0, a) + \mathbb{R}(0, 1, b) \subseteq \mathbb{R}^2 \times \mathbb{R}$$

die Ebene durch $(0, 0, 0)$, $(1, 0, a)$ und $(0, 1, b)$.

§2. Basiswechsel

In diesem Abschnitt sei V ein Vektorraum über K mit Dimension $n \in \mathbb{N}$, $\underline{v} := (v_1, \ldots, v_n)$ eine Basis von V, W ein Vektorraum über K mit Dimension $m \in \mathbb{N}$ und $\underline{w} := (w_1, \ldots, w_m)$ eine Basis von W. Wenn $f : V \longrightarrow W$ eine Funktion ist, schreiben wir kurz $f(\underline{v})$ anstatt $(f(v_1), \ldots, f(v_n))$.

Satz 5 : *Sei $f : V \longrightarrow W$ eine K-lineare Funktion mit Matrix $A := M(f, \underline{v}, \underline{w}) \in K^{m\times n}$, $u \in V$ und $c \in K^{n\times 1}$ die Koordinatenspalte von u bezüglich $\underline{v}$, also $u = \underline{v}c$. Dann ist*

$$f(\underline{v}) = \underline{w}A \qquad \textit{und} \qquad f(u) = \underline{w}Ac\,.$$

Beweis: Es ist

$$f(\underline{v}) = (f(v_1), \ldots, f(v_n)) = (\underline{w}A_{-1}, \ldots, \underline{w}A_{-n}) = \underline{w}A$$

und

$$f(u) = f(\underline{v}c) = f(\underline{v})c = \underline{w}Ac\,.$$

Satz 6 : *Sei $f : V \to W$ eine K-lineare Funktion mit Matrix*

$$A := M(f, \underline{v}, \underline{w}) \in K^{m\times n}\,.$$

Sei $\underline{v}'$ eine weitere Basis von V, $\underline{w}'$ eine weitere Basis von W, $T \in \mathrm{GL}_n(K)$ die Transformationsmatrix von $\underline{v}$ zu $\underline{v}'$ und $S \in \mathrm{GL}_m(K)$ die Transformationsmatrix von $\underline{w}$ zu $\underline{w}'$. Dann ist

$$M(f, \underline{v}', \underline{w}') = S^{-1}AT\,.$$

Im Spezialfall $V = W$, $\underline{v} = \underline{w}$ und $\underline{v}' = \underline{w}'$ ist $S = T$ und

$$M(f, \underline{v}') = T^{-1}AT\,.$$

Beweis: Nach Satz 5 ist

$\underline{w}'M(f, \underline{v}', \underline{w}') = f(\underline{v}') = f(\underline{v}T) = f(\underline{v})T = \underline{w}AT = (\underline{w}'S^{-1})AT = \underline{w}'S^{-1}AT\,.$

Da $\underline{w}'$ eine Basis ist, folgt daraus $M(f, \underline{v}', \underline{w}') = S^{-1}AT\,.$

Definition 7 :

(1) Zwei Matrizen $A, B \in K^{m\times n}$ heißen *äquivalent*, wenn es Matrizen $P \in \mathrm{GL}_m(K)$ und $Q \in \mathrm{GL}_n(K)$ gibt mit

$$B = PAQ\,.$$

(2) Zwei Matrizen $A, B \in K^{n\times n}$ heißen *ähnlich*, wenn es eine Matrix $T \in \mathrm{GL}_n(K)$ gibt mit

$$B = T^{-1}AT\,.$$

Satz 8 :

(1) *Sei $f : V \to W$ eine K-lineare Funktion mit Matrix*

$$A := M(f, \underline{v}, \underline{w}) \in K^{m\times n}\,.$$

Eine Matrix $B \in K^{m\times n}$ ist genau dann zu A äquivalent, wenn es eine Basis $\underline{v}'$ von V und eine Basis $\underline{w}'$ von W gibt mit

$$M(f, \underline{v}', \underline{w}') = B\,.$$

(2) *Sei $f : V \to V$ eine K-lineare Funktion mit Matrix*

$$A := M(f, \underline{v}) \in K^{n\times n}\,.$$

Eine Matrix $B \in K^{n\times n}$ ist genau dann zu A ähnlich, wenn es eine Basis $\underline{v}'$ von V gibt mit

$$M(f, \underline{v}') = B\,.$$

Beweis:

(1) Nach Satz 6 sind die Matrizen $M(f, \underline{v}, \underline{w})$ und $M(f, \underline{v}', \underline{w}')$ äquivalent. Wenn B zu A äquivalent ist, gibt es $P \in \mathrm{GL}_m(K)$ und $Q \in \mathrm{GL}_n(K)$ mit $B = PAQ$. Dann ist $\underline{v}' := \underline{v}Q$ eine Basis von V und $\underline{w}' := \underline{w}P^{-1}$ eine Basis von W. Nach Satz 6 ist $M(f, \underline{v}', \underline{w}') = PAQ = B$.

(2) analog.

Ähnliche Matrizen beschreiben (bezüglich verschiedener Basen) dieselbe lineare Funktion. Wird ein physikalisches Phänomen durch eine lineare Funktion von einem endlichdimensionalen Vektorraum V nach V beschrieben und diese (nach Wahl einer Basis von V) durch eine Matrix, dann haben nur jene Eigenschaften dieser Matrix „physikalische Bedeutung“, die sich beim Übergang zu einer ähnlichen Matrix nicht ändern. Beispiele für solche Eigenschaften von Matrizen sind die Determinante und die „Spur“.

Definition 9 : Sei $A \in K^{n\times n}$. Dann heißt

$$\operatorname{spur}(A) := \sum_{i=1}^{n} A_{ii}$$

die *Spur* von A.

Satz 10 :

(1) *Die Funktion* $\operatorname{spur} : K^{n\times n} \longrightarrow K$ *ist linear.*
(2) *Für* $A, B \in K^{n\times n}$ *gilt:* $\operatorname{spur}(AB) = \operatorname{spur}(BA)$.
(3) *Für* $A \in K^{n\times n}$ *und* $T \in GL_n(K)$ *gilt:* $\operatorname{spur}(T^{-1}AT) = \operatorname{spur}(A)$.
(4) *Sei* $f : V \to V$ *eine lineare Funktion und* A *die Matrix von* f *bezüglich* $\underline{v}$. *Dann ist*

$$\operatorname{spur}(f) := \operatorname{spur}(A)$$

unabhängig von der Wahl der Basis $\underline{v}$ *und heißt* Spur *von* f.

Beweis: (1) und (2) nachrechnen, (3) folgt aus (2), (4) aus (3).

Satz 11 :

(1) *Für* $A, B \in K^{n\times n}$ *gilt:* $\det(AB) = \det(BA)$.
(2) *Für* $A \in K^{n\times n}$ *und* $T \in GL_n(K)$ *gilt:* $\det(T^{-1}AT) = \det(A)$.
(3) *Sei* $f : V \to V$ *eine lineare Funktion und* A *die Matrix von* f *bezüglich* $\underline{v}$. *Dann ist*

$$\det(f) := \det(A)$$

unabhängig von der Wahl der Basis $\underline{v}$ *und heißt* Determinante *von* f.

Beweis:

(1) $\det(AB) = \det(A)\det(B) = \det(B)\det(A) = \det(BA)$,
(2) folgt aus (1),
(3) folgt aus (2).

§3. Bild und Kern einer linearen Funktion

In diesem Abschnitt seien V und W Vektorräume über K und $f : V \longrightarrow W$ eine lineare Funktion.

Definition 12 : Die Menge

$$\mathrm{Bild}(f) := \{f(v) \mid v \in V\} \subseteq W$$

heißt *Bild* von f und die Menge

$$\mathrm{Kern}(f) := \{v \in V \mid f(v) = 0_W\} \subseteq V$$

heißt *Kern* von f.

Satz 13 : $\mathrm{Bild}(f)$ *ist ein Untervektorraum von* W*,* $\mathrm{Kern}(f)$ *ist ein Untervektorraum von* V.
Die Dimension des Bildes von f *heißt* Rang von f *(Schreibweise* $\mathrm{rg}(f)$*).*

Beweis: Da f linear ist, ist $0_V \in \mathrm{Kern}(f)$. Für $u, v \in \mathrm{Kern}(f)$ und $c \in K$ folgt aus $f(u+v) = f(u) + f(v) = 0_W$ auch $u + v \in \mathrm{Kern}(f)$, sowie aus $f(cu) = cf(u) = 0_W$ auch $cu \in \mathrm{Kern}(f)$. Daher ist $\mathrm{Kern}(f)$ ein Untervektorraum von V. Analog zeigt man, dass $\mathrm{Bild}(f)$ ein Untervektorraum von W ist.

Satz 14 : *Sei* $A \in K^{m\times n}$ *und* $L(A,0) := \{x \in K^{n\times 1} \,|\, Ax = 0\}$ *der Lösungsraum des durch* A *definierten Systems homogener linearer Gleichungen. Fasst man die Matrix* A *als lineare Funktion*

$$A : K^{n\times 1} \to K^{m\times 1}\,,\, x \mapsto Ax\,,$$

auf, dann ist $\mathrm{Kern}(A) = L(A,0)$ *und* $\mathrm{Bild}(A) = {}_K\langle A_{-1}, \ldots, A_{-n}\rangle$.

Beweis: Es ist $\mathrm{Kern}(A) = \{x \in K^{n\times 1} \mid Ax = 0\} = L(A,0)$ und $\mathrm{Bild}(A) = = \{\,Ax \,|\, x \in K^{n\times 1}\,\} = \{\,\sum_{i=1}^{n} x_i A_{-i} \,|\, x_1, \ldots x_n \in K\,\} = {}_K\langle A_{-1}, \ldots, A_{-n}\rangle$.

Satz 15 : *Seien* V, W *endlich-dimensionale Vektorräume über* K*,* $f : V \to W$ *eine* K*-lineare Funktion und* $r := \mathrm{rg}(f)$*. Dann gibt es eine Basis* $(v_1, \ldots, v_n)$ *von* V *so, dass*

(1) $(f(v_1), \ldots, f(v_r))$ *eine Basis von* $\mathrm{Bild}(f)$ *und*
(2) $(v_{r+1}, \ldots, v_n)$ *eine Basis von* $\mathrm{Kern}(f)$

ist. Insbesondere gilt

$$\dim_K(V) = \dim_K(\mathrm{Bild}(f)) + \dim_K(\mathrm{Kern}(f))\,.$$

Ergänzt man die Basis $(f(v_1),\dots,f(v_r))$ *von* $\mathrm{Bild}(f)$ *zu einer Basis* $(w_1,\dots,w_m)$ *von* W*, dann ist*

$$D_r := \begin{pmatrix} 1 & 0 & \dots & & \\ 0 & \ddots & \ddots & & \\ \vdots & \ddots & 1 & & \\ & & & 0 & \\ & & & & \ddots \end{pmatrix} \in K^{m\times n}$$

(nur an den Stellen $(1,1),\dots,(r,r)$ *stehen Einsen und sonst Nullen) die Matrix von* f *bezüglich der Basen* $(v_1,\dots,v_n)$ *und* $(w_1,\dots,w_m)$*.*

Beweis: Sei $(w_1,\dots,w_r)$ eine Basis von $\mathrm{Bild}(f)$. Dann kann man Urbilder $v_1,\dots,v_r\in V$ von $w_1,\dots,w_r$ unter f wählen. Sei $(u_1,\dots,u_s)$ eine Basis von $\mathrm{Kern}(f)$. Dann ist

$$(v_1,\dots,v_r,u_1,\dots,u_s)$$

ein Erzeugendensystem von V, weil für $y\in V$ aus

$$f(y)=\sum_{i=1}^r a_iw_i=\sum_{i=1}^r a_if(v_i)=f(\sum_{i=1}^r a_iv_i)$$

folgt, dass $z:=y-\sum_{i=1}^r a_iv_i\in\mathrm{Kern}(f)$ ist. Daher ist $y=z+\sum_{i=1}^r a_iv_i$ eine Linearkombination von $(v_1,\dots,v_r,u_1,\dots,u_s)$.

Wir zeigen noch, dass $(v_1,\dots,v_r,u_1,\dots,u_s)$ linear unabhängig ist. Seien dazu $c_1,\dots,c_r,d_1,\dots,d_s\in K$ mit

$$\sum_{i=1}^r c_iv_i+\sum_{j=1}^s d_ju_j=0\,.$$

Dann ist $0=f(\sum_{i=1}^r c_iv_i+\sum_{j=1}^s d_ju_j)=\sum_{i=1}^r c_if(v_i)=\sum_{i=1}^r c_iw_i$.
Da $(w_1,\dots,w_r)$ linear unabhängig ist, sind alle c_i gleich 0. Dann ist $\sum_{j=1}^s d_ju_j=0$, und aus der linearen Unabhängigkeit von $u_1,\dots,u_s$ folgt $d_1=\dots=d_s=0$. Also ist $(v_1,\dots,v_r,u_1,\dots,u_s)$ die gesuchte Basis von V. Insbesondere ist $r+s=n$.

Satz 16: *Seien* $n\in\mathbb{N}$ *bzw.* $m\in\mathbb{N}$ *die Dimensionen von* V *bzw.* W*,* $\underline{v}$ *eine Basis von* V*,* $\underline{w}$ *eine Basis von* W *und* $A:=M(f,\underline{v},\underline{w})\in K^{m\times n}$ *die Matrix von* f*. Dann ist*

$$\mathrm{rg}(f)=\mathrm{rg}(A)$$

und

$$\dim_K(\mathrm{Kern}(f))=\dim_K(\mathrm{Kern}(A))\,.$$

Beweis: Die Koordinaten-Funktionen

$$h:V\to K^{n\times 1}\,,\,\underline{v}c\mapsto c\,,$$

und

$$k : W \to K^{m\times 1}\,, \underline{w}d \mapsto d\,,$$

sind Isomorphismen. Fassen wir A als lineare Funktion

$$K^{n\times 1} \to K^{m\times 1}\,, x \mapsto Ax\,,$$

auf, dann ist $f = k^{-1} \circ A \circ h$. Weil k und h Isomorphismen sind, prüft man leicht nach, dass $k(\mathrm{Bild}(f)) = \mathrm{Bild}(A)$ und $h(\mathrm{Kern}(f)) = \mathrm{Kern}(A)$ ist. Daher ist

$$\mathrm{rg}(f) = \dim_K(\mathrm{Bild}(f)) = \dim_K(k(\mathrm{Bild}(f))) = \dim_K(\mathrm{Bild}(A)) = \mathrm{rg}(A)$$

und

$$\dim_K(Kern(f)) = \dim_K(h(Kern(f))) = \dim_K(\mathrm{Kern}(A))\,.$$

Satz 17 :

(1) *Jede Matrix $A \in K^{m\times n}$ vom Rang r ist zur Matrix*

$$D_r := \begin{pmatrix} 1 & 0 & \dots & & & \\ 0 & \ddots & \ddots & & & \\ \vdots & \ddots & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \end{pmatrix} \in K^{m\times n}$$

äquivalent, wo nur an den Stellen $(1,1),\dots,(r,r)$ Einsen stehen und sonst Nullen.

(2) *Zwei Matrizen $A,B \in K^{m\times n}$ sind genau dann äquivalent, wenn sie den gleichen Rang besitzen.*

Beweis:

(1) folgt aus Satz 8, Satz 15 und Satz 16.
(2) folgt aus (1), Satz 8 und Satz 16.

Zur Berechnung von invertierbaren Matrizen P,Q mit $PAQ = D_r$:
Die Matrix A kann durch elementare Zeilenumformungen in eine Matrix PA in Stufenform umgeformt werden. Analog kann PA durch elementare Spaltenumformungen (Multiplikation mit Elementarmatrizen von rechts) auf die Form $PAQ = D_r$ gebracht werden. Die Matrix $Q \in \mathrm{GL}_n(K)$ kann berechnet werden, indem man die elementaren Spaltenumformungen auf die Matrix

$$\begin{pmatrix} PA \\ I_n \end{pmatrix}$$

anwendet und als Ergebnis

$$\begin{pmatrix} PA \\ I_n \end{pmatrix} Q = \begin{pmatrix} PAQ \\ Q \end{pmatrix}$$

erhält.

Satz 18 : *Für $A \in K^{m\times n}$ gilt*

$$\operatorname{rg}(A^T) = \operatorname{rg}(A)\,.$$

Insbesondere haben der Zeilenraum und der Spaltenraum von A die gleiche Dimension („Zeilenrang= Spaltenrang“).

Beweis: Nach Satz 17,(2) gibt es Matrizen $P \in \mathrm{GL}_m(K)$ und $Q \in \mathrm{GL}_n(K)$ so, dass $PAQ = D_r$ ist, wobei $r := \operatorname{rg}(A)$ ist. Dann ist

$$D_r = (D_r)^T = (PAQ)^T = Q^T A^T P^T$$

mit $P^T \in \mathrm{GL}_m(K)$ und $Q^T \in \mathrm{GL}_n(K)$, also A^T äquivalent zu D_r. Nach Satz 17 folgt, dass auch A^T den Rang r hat.
Der Zeilenraum von A ist der Spaltenraum von A^T.

Es sei $A \in K^{m\times n}$ und $L(A,0)$ die Lösungsmenge des durch A definierten homogenen Systems linearer Gleichungen. In der Vorlesung „Einführung in die Mathematik 1“ wurde gezeigt, dass $L(A,0)$ ein Untervektorraum von $K^{n\times 1}$ ist, dessen Dimension gleich $n - \operatorname{rg}(A)$ ist. Der Rang von A ist nach Definition die Dimension des Spaltenraums von A, also die maximale Anzahl linear unabhängiger Spalten von A. Nach Satz 18 ist $\operatorname{rg}(A) = \operatorname{rg}(A^T)$, also auch gleich der maximalen Anzahl linear unabhängiger Zeilen von A. Die Spalten von A entsprechen den „Unbekannten“ des durch A gegebenen Systems linearer Gleichungen, die Zeilen von A den „Gleichungen“. Daher kann die Dimension von $L(A,0)$ als „Anzahl der Unbekannten minus Anzahl der linear unabhängigen Gleichungen“ beschrieben werden.

§4. Systeme linearer Gleichungen in koordinatenfreier Form

Definition 19 : Ein *System linearer Gleichungen in koordinatenfreier Form* ist eine Aufgabe:

- Gegeben sind eine lineare Funktion $f : V \to W$ und ein Vektor $y \in W$.
- Gesucht ist eine „gute Beschreibung“ der Menge

  $$L(f,y) := f^{-1}(\{y\}) = \{x \in V \mid f(x) = y\}$$

  aller Vektoren $x \in V$, für die $f(x) = y$ ist.

Die Menge $L(f,y)$ heißt *Lösungsmenge* des durch f und y gegebenen Systems linearer Gleichungen. Ihre Elemente heißen *Lösungen* dieses Systems.

Das durch f und y gegebene System linearer Gleichungen heißt *homogen*, wenn $y = 0_W$ ist, ansonsten *inhomogen*. Die Lösungsmenge eines homogenen Systems linearer Gleichungen ist

$$L(f,0) = \operatorname{Kern}(f)\,.$$

Satz 20: *Sei $f : V \to W$ K-linear, $y \in W$ und $z \in L(f,y)$ (also ist $L(f,y)$ insbesondere nicht leer). Dann ist*

$$L(f,y) = z + \operatorname{Kern}(f)$$

ein affiner Unterraum von V mit Aufpunkt z und parallelem Untervektorraum $\operatorname{Kern}(f)$.
Das durch f und y gegebene System „lösen“ bedeutet daher: finde (irgend)ein Urbild z von y unter f und (irgend)eine Basis von $\operatorname{Kern}(f)$.
Falls V endlichdimensional ist, gilt weiters

$$\dim_K(L(f,y)) = \dim_K(V) - \operatorname{rg}(f)\,.$$

Beweis: Sei $v \in \operatorname{Kern}(f)$. Dann ist $f(z+v) = f(z) + f(v) = y + 0 = y$, also $z + v \in L(f,y)$.

Sei $x \in L(f,y)$. Dann ist $f(x-z) = f(x) - f(z) = y - y = 0$, also $x - z \in \operatorname{Kern}(f)$ und $x = z + (x-z) \in \{z + v \mid v \in \operatorname{Kern}(f)\}$.
Nach Satz 15 ist $\dim_K(\operatorname{Kern}(f)) = \dim_K(V) - \operatorname{rg}(f)$.

Beispiel 21: Fasst man eine Matrix $A \in K^{m\times n}$ als eine lineare Funktion

$$f : K^{n\times 1} \to K^{m\times 1}\,,\, x \mapsto Ax\,,$$

auf, dann ist $L(f,y) = L(A,y)$.

Beispiel 22: Sei $\mathcal{C}(\mathbb{R},\mathbb{R}) := \{f \mid f : \mathbb{R} \to \mathbb{R} \text{ stetig}\}$,
$\mathcal{C}^1(\mathbb{R},\mathbb{R}) := \{f \mid f : \mathbb{R} \to \mathbb{R} \text{ stetig differenzierbar}\}$ und

$$D : \mathcal{C}^1(\mathbb{R},\mathbb{R}) \to \mathcal{C}(\mathbb{R},\mathbb{R})\,,\, f \mapsto f'\,,$$

wobei f' die Ableitung der Funktion f bezeichnet. Dann sind $\mathcal{C}(\mathbb{R},\mathbb{R})$ und $\mathcal{C}^1(\mathbb{R},\mathbb{R})$ mit der punktweisen Addition und Skalarmultiplikation Vektorräume über $\mathbb{R}$, und die Funktion D ist $\mathbb{R}$-linear. Der Unterraum $\operatorname{Kern}(D)$ besteht aus allen konstanten Funktionen. Eine Funktion $f \in \mathcal{C}^1(\mathbb{R},\mathbb{R})$ heißt *Stammfunktion* von $g \in \mathcal{C}(\mathbb{R},\mathbb{R})$, wenn $Df = g$ ist. Die Menge aller Stammfunktionen von g ist

$$L(D,g) = f + \operatorname{Kern}(D).$$

Satz 23: *Seien V,W Vektorräume über K der Dimensionen n,m mit Basen $\underline{v},\underline{w}$, sei $f : V \to W$ K-linear mit Matrix*

$$A := M(f,\underline{v},\underline{w}) \in K^{m\times n}$$

und $y = \underline{w}b \in W$. Dann bildet der Koordinaten-Isomorphismus

$$V \to K^{n\times 1}\,,\, \underline{v}c \mapsto c\,,$$

$L(f,y)$ auf $L(A,b)$ ab und $\operatorname{Kern}(f)$ *auf* $L(A,0)$.

Beweis: Nach Satz 5 ist $\underline{v}c \in L(f,y)$ genau dann wenn $\underline{w}(Ac) = \underline{w}b$, also $c \in L(A,b)$ ist.

Nach Satz 23 kann für $f: V \to W$ und $y \in W$ das System linearer Gleichungen (f,y) wie folgt gelöst werden:

(1) Wähle Basen $\underline{v}, \underline{w}$ von V, W.
(2) Berechne die Matrix $A := M(f, \underline{v}, \underline{w})$ und die Koordinatenspalte b von y bezüglich $\underline{w}$.
(3) Berechne die Lösungsmenge $L(A,b)$.
Wenn $L(A,b)$ leer ist, dann ist auch $L(f,y)$ leer.
Wenn $z \in L(A,b)$ und $(u_1, \ldots, u_s)$ eine Basis von $L(A,0)$ ist, dann ist $\underline{v}z \in L(f,y)$ und $(\underline{v}u_1, \ldots, \underline{v}u_s)$ eine Basis von Kern(f).

Im Schulunterricht entsprechen Systeme linearer Gleichungen in koordinatenfreier Form gewissen „Textaufgaben", und die Umwandlung in die Form $Ax = b$ nennt man „den Ansatz finden".

Beispiel 24: Wir suchen alle Polynomfunktionen $p \in \mathbb{R}[x]$ mit

$$p(-1) = 2, p(1) = 1, p(2) = 1 \text{ und } \operatorname{gr}(p) < 5.$$

Sei $V := \{q \in \mathbb{R}[x] \mid \operatorname{gr}(q) < 5\}$, $W := \mathbb{R}^3$,

$$f: V \longrightarrow W, \; q \longmapsto (q(-1), q(1), q(2)),$$

und $y := (2,1,1) \in W$. Die Funktion f ist linear.
Wir wählen die Basis $\underline{v} := (1, x, x^2, x^3, x^4)$ von V und die Standardbasis $\underline{w} := (e_1, e_2, e_3)$ von $W = \mathbb{R}^3$. Dann ist

$$A := M(f, \underline{v}, \underline{w}) := \begin{pmatrix} 1 & -1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 & 16 \end{pmatrix} \quad \text{und} \quad b := \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}.$$

Man berechnet mit dem Gauß-Verfahren

$$L(A,b) = \begin{pmatrix} \frac{4}{3} \\ -\frac{1}{2} \\ \frac{1}{6} \\ 0 \\ 0 \end{pmatrix} + \mathbb{R} \begin{pmatrix} -2 \\ 1 \\ 2 \\ -1 \\ 0 \end{pmatrix} + \mathbb{R} \begin{pmatrix} -4 \\ 0 \\ 5 \\ 0 \\ -1 \end{pmatrix}.$$

Daher ist

$$L(f,y) = \{\frac{4}{3} - \frac{1}{2}x + \frac{1}{6}x^2 + $$
$$+ c(-2 + x + 2x^2 - x^3) + d(-4 + 5x^2 - x^4) \mid c, d \in \mathbb{R}\}.$$

Satz 25: *Sei V ein endlich-dimensionaler Vektorraum über K und Z ein affiner Unterraum von V. Dann ist Z die Lösungsmenge eines Systems linearer Gleichungen, d.h. es gibt eine lineare Funktion*

$f : V \to W$ und einen Vektor $y \in W$ mit

$$Z = L(f, y)\,.$$

(Dann ist Z durch f und y „in impliziter Form“ *gegeben).*
Wenn der affine Unterraum Z durch einen Aufpunkt p und eine Basis $(u_1, \dots, u_k)$ des parallelen Untervektorraums gegeben ist, dann kann ein solches System linearer Gleichungen auf die folgende Weise berechnet werden:

Ergänze $(u_1, \dots, u_k)$ zu einer Basis $(u_1, \dots, u_k, u_{k+1}, \dots, u_n)$ von V. Setze

$$f : V \longrightarrow K^{n-k}\ ,\ \sum_{i=1}^{n} c_i u_i \longmapsto (c_{k+1}, c_{k+2}, \dots, c_n)$$

und $y := f(p)$.

Beweis: Seien f und y wie im Satz definiert. Dann ist $\mathrm{Kern}(f) = {}_K\!< u_1, \dots, u_k >$ und $p \in L(f, y)$. Nach Satz 20 ist $Z = L(f, y)$.

§5. Affine Funktionen

V und W seien Vektorräume über einem Körper K.

Definition 26 : Eine Funktion a: $V \longrightarrow W$ heißt *affin*, wenn es eine lineare Funktion $f : V \longrightarrow W$ und einen Vektor $w \in W$ gibt mit

$$a = t_w \circ f\,,$$

wobei t_w die Translation um w in W ist, d.h. es ist

$$a(x) = f(x) + w$$

für alle $x \in V$. Insbesondere ist $w = a(0)$ und $f = t_{(-w)} \circ a$, also sind t_w und f eindeutig durch a bestimmt und heißen der *Translationsanteil* bzw. der *lineare Anteil* von a.

Beispiel 27 : Lineare Funktionen und Translationen sind affine Funktionen.

Beispiel 28 : Jede lineare Funktion von $K^{n\times 1}$ nach $K^{m\times 1}$ ist von der Form $x \mapsto Ax$ mit $A \in K^{m\times n}$. Daher ist jede affine Funktion von $K^{n\times 1}$ nach $K^{m\times 1}$ von der Form $x \mapsto Ax + b$ mit $A \in K^{m\times n}$ und $b \in K^{m\times 1}$, d.h. von der Form

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} A_{11}x_1 + A_{12}x_2 + \cdots + A_{1n}x_n + b_1 \\ A_{21}x_1 + A_{22}x_2 + \cdots + A_{2n}x_n + b_2 \\ \vdots \\ A_{m1}x_1 + A_{m2}x_2 + \cdots + A_{mn}x_n + b_m \end{pmatrix}.$$

Satz 29: *Eine Funktion zwischen zwei Vektorräumen V und W ist genau dann affin, wenn ihr Graph ein affiner Unterraum von $V \times W$ ist.*

Beweis: Der Graph einer affinen Funktion $a : V \longrightarrow W$ ist

$$Graph(a) = \{(v, a(v)) \,|\, v \in V\} = \{(v, f(v) + w) \,|\, v \in V\} =$$

$$= \{(0, w) + (v, f(v)) \,|\, v \in V\} = (0, w) + Graph(f) \subseteq V \times W,$$

dabei ist f der lineare Anteil von a und $w := a(0)$. Weil f linear ist, ist $Graph(f)$ ein Untervektorraum von $V \times W$ und somit $(0, w) + Graph(f)$ ein affiner Unterraum.

Ist der Graph einer Funktion $g : V \longrightarrow W$ ein affiner Unterraum $(0, x) + U$, wobei U ein Untervektorraum von $V \times W$ ist und $x \in W$, dann ist U der Graph einer linearen Funktion f und $g = t_x \circ f$ ist eine affine Funktion.

Satz 30: *Sei $a : V \to W$ eine affine Funktion und M, N affine Unterräume von V. Dann ist $a(M)$ ein affiner Unterraum von W. Wenn M und N parallel sind, dann auch ihre Bilder $a(M)$ und $a(N)$.*

Beweis: Sei $w \in W$ und $f : V \to W$ eine lineare Funktion so, dass $a = t_w \circ f$ ist. Sei $p \in M$ und U der zu M parallele Untervektorraum von V. Dann ist auch $f(U)$ ein Untervektorraum von W, und

$$a(M) = f(M) + w = (f(p) + f(U)) + w = (w + f(p)) + f(U)$$

ist ein affiner Unterraum von W.

Satz 31: *Die Hintereinanderausführung affiner Funktionen sowie die Umkehrfunktion einer bijektiven affinen Funktion sind wieder affin.*

Beweis: Seien $a : V \to W$ und $b : Y \to Z$ affine Funktionen mit
Bild$(a) \subset Y$. Seien $f : V \to W$, $g : Y \to Z$ linear und $w \in W$, $z \in Z$ mit $a = t_w \circ f$ und $b = t_z \circ g$. Dann ist für alle $x \in V$

$$\begin{aligned} b(a(x)) &= g(a(x)) + z = g(f(x) + w) + z = g(f(x)) + g(w) + z \\ &= (t_{g(w)+z}(g(f(x)))\,, \end{aligned}$$

also ist $b \circ a$ wieder affin. Wenn a bijektiv ist, dann ist auch f bijektiv und die Umkehrfunktion $t_{-f^{-1}(w)} \circ f^{-1}$ ebenfalls affin.

Satz 32: *Sei $a : V \to W$ eine Funktion. Die folgenden Aussagen sind äquivalent:*

(1) *Die Funktion a ist affin.*

(2) *Für jede affine Linearkombination $\sum_{i\in I} c_i v_i$ einer endlichen Familie in V ist*

$$a\left(\sum_{i\in I} c_i v_i\right) = \sum_{i\in I} c_i a(v_i)\,.$$

(„Das Bild einer affinen Linearkombination ist die affine Linearkombination der Bilder“.)

Beweis:

(1) $\Rightarrow$ (2): Sei $w \in W$ und $f : V \to W$ eine lineare Funktion so, dass $a = t_w \circ f$ ist. Dann ist

$$a\left(\sum_{i\in I} c_i v_i\right) = w + \sum_{i\in I} c_i f(v_i) = \left(\sum_{i\in I} c_i\right) w + \sum_{i\in I} c_i f(v_i) =$$
$$= \sum_{i\in I} c_i (w + f(v_i)) = \sum_{i\in I} c_i a(v_i)\,.$$

(2) $\Rightarrow$ (1): Sei $w := a(0)$ und $f := t_{(-w)} \circ a$. Es ist zu zeigen, dass f linear ist.

Seien $c_1, c_2 \in K$ und $x_1, x_2 \in V$. Dann ist

$$f(c_1x_1 + c_2x_2) = a(c_1x_1 + c_2x_2) - w =$$
$$= a(c_1x_1 + c_2x_2 + (1 - c_1 - c_2)0) - w =$$
$$= c_1 a(x_1) + c_2 a(x_2) + (1 - c_1 - c_2) \cdot a(0) - w =$$
$$= c_1(a(x_1) - w) + c_2(a(x_2) - w) = c_1 f(x_1) + c_2 f(x_2)\,.$$

Definition 33: Es sei $(v_i)_{i\in I}$ eine endliche Familie in V.
Der *Schwerpunkt* von $(v_i)_{i\in I}$ ist

$$\frac{1}{\#(I)} \sum_{i\in I} v_i\,.$$

Der Schwerpunkt von (v_1, v_2) heißt *Mittelpunkt der Strecke* zwischen v_1 und v_2.

Satz 34: *Es seien W ein Vektorraum, $a : V \to W$ eine affine Funktion und $(v_i)_{i\in I}$ eine endliche Familie in V. Dann gilt:*

(1) *Das Bild der konvexen Hülle von $(v_i)_{i\in I}$ bezüglich a ist die konvexe Hülle der Familie $(a(v_i))_{i\in I}$ in W.*
(2) *Das Bild des Schwerpunkts von $(v_i)_{i\in I}$ ist der Schwerpunkt von $(a(v_i))_{i\in I}$.*

Beweis: Folgt aus Satz 32.

Beispiel 35: Es seien P ein Polytop in V und $a : V \to \mathbb{R}$ eine affine Funktion. Dann ist $a(P)$ ein abgeschlossenes Intervall in $\mathbb{R}$.

KAPITEL 2

Systeme linearer Ungleichungen

§1. Wiederholung: Zwei Zugänge zur Geometrie der Ebene

§1.1. Die Ebene (nach Wahl eines Nullpunktes) als Vektorraum. Wir nehmen an, wir haben ein „beliebig großes" Zeichenblatt E und die folgenden Zeichengeräte:

- einen „beliebig fein gespitzten" Bleistift,
- ein „beliebig langes" Lineal und
- ein Dreieck.

Wir betrachten das Zeichenblatt als Menge von „Punkten" und wählen einen davon aus. Diesen ausgewählten Punkt nennen wir *Nullpunkt* und bezeichnen ihn mit $0 \in E$.

Wir nehmen an, dass mit Lineal und Bleistift durch je zwei Punkte eine „Gerade" gezeichnet werden kann und dass mit Lineal, Dreieck und Bleistift jede Gerade in jeden Punkt „parallelverschoben" werden kann.

Je zwei Punkten $A, B \in E$ können wir wie folgt einen dritten Punkt, den wir mit $A+B$ bezeichnen, zuordnen:

- Falls $0, A$ und B nicht auf einer Geraden liegen:
  Zeichne eine Gerade durch 0 und B und verschiebe sie in den Punkt A. Zeichne eine Gerade durch 0 und A und verschiebe sie in den Punkt B.
  Dann sei $A+B$ der „Schnittpunkt" dieser zwei Geraden. q
- Falls $0, A$ und B auf einer Geraden liegen:
  Wähle einen Punkt $H \in E$, der nicht auf dieser Geraden liegt.
  Konstruiere wie oben die Punkte $A+H$ und $(A+H)+B$.
  Verschiebe die Gerade durch 0 und H in den Punkt $(A+H)+B$.
  Dann sei $A+B$ der Schnittpunkt dieser Geraden mit der Geraden durch 0 und A.

Wir nehmen nun zusätzlich an, dass das Lineal mit einer Skala versehen ist, aus der jede reelle Zahl abgelesen werden kann. Dann kann jeder reellen Zahl c und jedem Punkt $A \in E$ ein weiterer Punkt, den wir mit $c \cdot A$ bezeichnen, wie folgt zugeordnet werden:

- Zeichne mit dem Lineal eine Gerade durch 0, die den Punkt A nicht enthält.
- Lege das Lineal so, dass die Zahl 0 über dem Punkt 0 liegt und zeichne dann die Punkte P bzw. Q, über denen die Zahlen 1 bzw. c liegen, auf dieser Geraden ein.
- Verschiebe die Gerade durch A und P in den Punkt Q.

- Dann sei $c \cdot A$ der Schnittpunkt dieser Geraden mit der Geraden durch 0 und A.

Wir nehmen an, dass die so definierten Rechenoperationen $+$ („Addition von Punkten“) und $\cdot$ („Skalarmultiplikation von reellen Zahlen mit Punkten“) die Rechenregeln eines Vektorraums erfüllen. Dann ist die Zeichenebene E ein Vektorraum und ihre Punkte sind Vektoren.

Es sei A ein von 0 verschiedener Punkt. Die Gerade durch 0 und A ist (nach Definition der Skalarmultiplikation) die Menge aller skalaren Vielfachen von A. Wenn die Punkte $0, A$ und B nicht auf einer Geraden liegen, dann ist das Punktepaar (A,B) eine $\mathbb{R}$-Basis des Vektorraums E. Wenn a bzw. b die Koordinaten eines Punktes P bezüglich dieser Basis sind, dann erhält man $a \cdot A$ (bzw. $b \cdot B$) als Schnittpunkt der Geraden $\mathbb{R} \cdot A$ mit der Geraden, die man erhält, indem man die Gerade durch 0 und B in den Punkt P verschiebt (bzw. als Schnittpunkt der Geraden $\mathbb{R} \cdot B$ mit der Geraden, die man erhält, indem man die Gerade durch 0 und A in den Punkt P verschiebt). Der Punkt P wird eindeutig durch das Zahlenpaar (a,b) beschrieben.

Die Gerade durch zwei Punkte P und Q ist dann die Menge

$$P + \mathbb{R}(Q-P) = \{P + c(Q-P) \mid c \in \mathbb{R}\}.$$

Die Wahl eines Nullpunktes 0 in der Ebene und von zwei Punkten A, B so, dass $0, A$ und B nicht auf einer Geraden liegen, nennt man auch *Wahl eines Koordinatensystems*. Man kann diese Wahl auch dadurch treffen, dass man ein Paar von Geraden, die genau einen Punkt gemeinsam haben, und auf jeder Geraden einen Punkt, der nicht der Schnittpunkt ist, wählt. Der Schnittpunkt ist dann der Nullpunkt und das Paar der auf den Geraden gewählten Punkte ist die Basis.

§1.2. Die Ebene als affiner Raum.

Definition 36 : Sei V ein Vektorraum über einem Körper K, A eine Menge und

$$V \times A \to A\,,\ (v,a) \mapsto v \cdot a\,,$$

eine Operation der Gruppe $(V,+)$ auf A. (Also: Für alle $a \in A$, $v,w \in V$ ist $0 \cdot a = a$ und $(v+w) \cdot a = v \cdot (w \cdot a)$.
A zusammen mit dieser Operation ist ein *affiner Raum über V*, wenn es für alle Elemente $a,b \in A$ genau einen Vektor $v \in V$ gibt mit $v \cdot a = b$.
Die Elemente von A heißen dann *Punkte*, die Elemente von V *Vektoren* des affinen Raums.

Satz 37 : *Sei A ein affiner Raum über V und $a \in A$. Die Funktion*

$$V \longrightarrow A\,,\ v \longmapsto v \cdot a\,,$$

ist bijektiv. (Nach Wahl eines „Nullpunktes“ kann ein affiner Raum als Vektorraum betrachtet werden).

Beweis: Folgt aus der Definition.

Sie E die Zeichenebene $T(E)$ der Vektorraum der Translationen von E. Dann ist E mit

$$T(E) \times E \longrightarrow E\,,\ (t,x) \longmapsto t(x)\,,$$

ein affiner Raum über $T(E)$.
Möchte man in der Zeichenebene keinen „Nullpunkt" wählen, kann man sie als affinen Raum betrachten. Dann muss man zwischen Punkten ($\in E$) und Vektoren ($\in T(E)$) unterscheiden. Punkte können dann nicht addiert werden, aber Vektoren können addiert werden und auf Punkten „wirken".

Sind P und Q Punkte von E und $P \neq Q$, dann gibt es genau eine Translation in $T(E)$, die P auf Q abbildet. Sie wird häufig mit $\vec{PQ}$ bezeichnet. Die Menge

$$\{t\vec{PQ}\,|\,t \in \mathbb{R}\} \subseteq T(E)$$

ist die Gerade durch $0_{T(E)} = id_E$ und $\vec{PQ}$ in $T(E)$. Die Gerade durch P und Q in E ist dann als

$$\{(t\vec{PQ})(P)\,|\,t \in \mathbb{R}\} \subseteq E$$

definiert. Wegen $(\vec{PQ})(P) = Q$ und $(0 \cdot \vec{PQ})(P) = id_E(P) = P$ sind P und Q Punkte dieser Geraden. Die Translation $\vec{PQ}$ wird als „Richtungsvektor" dieser Geraden bezeichnet.

§2. Lineare Ungleichungen und Halbräume

In diesem Abschnitt seien n eine positive ganze Zahl, V ein n-dimensionaler reeller Vektorraum und $V^* := \mathrm{Lin}_{\mathbb{R}}(V, \mathbb{R})$ der Vektorraum aller linearen Funktionen von V nach $\mathbb{R}$. Mit $\mathbb{R}_{\geq 0}$ bezeichnen wir die Menge aller reellen Zahlen, die nicht negativ sind.

Definition 38 : Eine *lineare Ungleichung* in V ist durch eine lineare Funktion $0 \neq f \in V^*$ und eine Zahl $b \in \mathbb{R}$ gegeben. Gesucht sind alle Vektoren $v \in V$ mit

$$f(v) \leq b.$$

Die Menge $L(f, \leq b) := \{v \in V \mid f(v) \leq b\}$ heißt *Lösungsmenge* der durch f und b gegebenen linearen Ungleichung, die Elemente von $L(f, \leq b)$ sind *Lösungen* dieser Ungleichung.

Die durch $0 \neq f \in V^*$ und $b \in \mathbb{R}$ gegebene lineare Ungleichung ist *homogen*, wenn $b = 0$ ist.

Sei $L(f, \geq b) := \{v \in V \mid f(v) \geq b\}$. Dann ist

$$L(f, \geq b) = L(-f, \leq -b)$$

die Lösungsmenge der durch $-f$ und $-b$ gegebenen linearen Ungleichung.

Beispiel 39 : Sei $V := \mathbb{R}^{n\times 1}$, $a_1, \ldots, a_n, b \in \mathbb{R}$ und

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}, \, (x_1, \ldots, x_n) \longmapsto \sum_{i=0}^{n} a_i x_i \, .$$

Dann ist $L((a_0, a_1, \ldots, a_n), \leq b) := L(f, \leq b) = \{x \in \mathbb{R}^n \mid \sum_{i=0}^{n} a_i x_i \leq b\}$.

Satz 40 : *Sei $0 \neq f \in V^*$ und $b \in \mathbb{R}$. Dann kann $L(f, \leq b)$ wie folgt durch endlich viele Daten beschrieben werden:*

Berechne eine Basis $(v_1, \ldots, v_{n-1})$ von Kern(f) *und v_n so, dass $f(v_n) = 1$ ist.*

Dann ist

$$L(f, \leq b) = \Big\{ \sum_{i=1}^{n} c_i v_i \mid c_1, \ldots, c_n \in \mathbb{R}, \; c_n \leq b \Big\}.$$

Beweis: Das $n-$Tupel $(v_1, \ldots, v_n)$ ist eine Basis von V.

$\subseteq$: Sei $v := \sum_{i=1}^{n} c_i v_i \in V$ und $f(v) \leq b$. Dann ist

$$b \geq f(v) = \sum_{i=1}^{n} c_i f(v_i) = c_n \, .$$

$\supseteq$: Seien $c_1, \ldots, c_n \in \mathbb{R}$ mit $c_n \leq b$. Wegen

$$f\big(\sum_{i=1}^{n} c_i v_i \big) = c_n f(v_n) \leq b$$

ist $\sum_{i=1}^{n} c_i v_i \in L(f, \leq b)$.

Beispiel 41 : Sei V der 3-dimensionale Vektorraum aller Polynomfunktionen von $\mathbb{R}$ nach $\mathbb{R}$ vom Grad 0, 1 oder 2, $b := 1$ und f die lineare Funktion

$$f : V \to \mathbb{R}, \quad h \mapsto \int_0^1 h(t) \mathrm{d}t.$$

Dann bilden

$$v_1 : \mathbb{R} \to \mathbb{R}, \quad t \mapsto t - \frac{1}{2}, \quad \text{und}$$
$$v_2 : \mathbb{R} \to \mathbb{R}, \quad t \mapsto t^2 - \frac{1}{3}$$

eine $\mathbb{R}$-Basis von Kern(f). Für

$$v_3 : \mathbb{R} \to \mathbb{R}, \quad t \mapsto 1$$

ist $f(v_3) = 1$. Also ist

$$L(f, \leq 1) = \{c_1 v_1 + c_2 v_2 + c_3 v_3 \mid c_1, c_2, c_3 \in \mathbb{R}, c_3 \leq 1\} =$$
$$= \Big\{ g \in V \mid g(t) = c_2 t^2 + c_1 t + (c_3 - \frac{1}{3} c_2 - \frac{1}{2} c_1), \; c_1, c_2, c_3 \in \mathbb{R}, \; c_3 \leq 1 \Big\}.$$

Beispiel 42 : Sei $V := \mathbb{R}^4$, $b := 2$ und

$$f : V \to \mathbb{R}, \quad (a_1, a_2, a_3, a_4) \mapsto a_1 + a_2 + a_3 + a_4.$$

Dann ist $\big((1,0,0,-1),(1,0,-1,0),(1,-1,0,0)\big)$ eine Basis von $\mathrm{Kern}(f)$ und $f\big((0,0,0,1)\big) = 1$. Also ist

$$L(f, \leq 2) = \{(c_1 + c_2 + c_3, -c_3, -c_2, -c_1 + c_4) \mid c_1, c_2, c_3, c_4 \in \mathbb{R},\ c_4 \leq 2\}.$$

Definition 43 : Für $v, w \in V$ sei

$$[v, w] := \{sv + tw \mid s, t \in \mathbb{R},\ s, t \geq 0,\ s + t = 1\}$$

die Strecke zwischen v und w.

Definition 44 : Es seien H eine Hyperebene in V (d. h.: ein affiner Unterraum von V der Dimension $n-1$) und $w \in V \setminus H$. Dann ist

$$\big\{v \in V \mid [v, w] \cap H \subseteq \{v\}\big\}$$

der durch H und w gegebene *Halbraum*. Die Hyperebene H ist der *Rand* dieses Halbraums. Wenn $n = 2$ ist, nennt man einen Halbraum *Halbebene*. Deren Rand ist eine Gerade.

Satz 45 : *Jeder Halbraum in V ist die Lösungsmenge einer linearen Ungleichung.*

Die Lösungsmenge einer linearen Ungleichung ist ein Halbraum.

Eine lineare Ungleichung ist genau dann homogen, wenn 0 *ein Element des Randes ihrer Lösungsmenge ist.*

Beweis: Es seien H eine Hyperebene in V und $w \in V \setminus H$. Wähle eine lineare Funktion $f : V \to \mathbb{R}$ so, dass der Untervektorraum $\mathrm{Kern}(f)$ zu H parallel ist. Dann gibt es eine Zahl $b \in \mathbb{R}$ so, dass $f(H) = \{b\}$ ist. Wenn $f(w) > b$ ist, ersetze f durch $-f$ und b durch $-b$. Dann ist $f(w) < b$ und $L(f, \leq b)$ ist der durch H und w gegebene Halbraum.

Sei nun durch $0 \neq g \in V^*$ und $c \in \mathbb{R}$ eine lineare Ungleichung gegeben. Dann ist $L(g, \leq c)$ der durch $g^{-1}(c)$ und ein Element von $g^{-1}(c-1)$ gegebene Halbraum.

Die Aussage über homogene lineare Ungleichungen prüft man nun leicht nach.

Definition 46 : Sei k eine positive ganze Zahl. Ein *System von k linearen Ungleichungen* ist durch lineare Funktionen $f_1, \ldots, f_k \in V^*$ und reelle Zahlen $b_1, \ldots, b_k \in \mathbb{R}$ gegeben. Gesucht sind alle Vektoren $v \in V$ mit

$$f_1(v) \leq b_1, \ldots, f_k(v) \leq b_k.$$

Die *Lösungsmenge* des Systems ist

$$L(f_1,\dots,f_k,\leq b_1,\dots,\leq b_k) := \{v \in V \mid f_1(v) \leq b_1,\dots,f_k(v) \leq b_k\}.$$

Es ist

$$L(f_1,\dots,f_k,\leq b_1,\dots,\leq b_k) = \bigcap_{i=1}^{k} L(f_i,\leq b_i),$$

also ist $L(f_1,\dots,f_k,\leq b_1,\dots,b_k)$ der Durchschnitt von k Halbräumen.

Definition 47 : Ein *Polyeder in V* ist der Durchschnitt von endlich vielen Halbräumen in V.

Nach Satz 45 ist jedes Polyeder in V die Lösungsmenge eines Systems von endlich vielen linearen Ungleichungen.

Wenn $\underline{v} = (v_1,\dots,v_n)$ eine Basis von V ist, dann werden die linearen Funktionen $f_1,\dots,f_k \in V^*$ eindeutig durch die n-Tupel $\big(f_i(v_1),\dots,f_i(v_n)\big) \in \mathbb{R}^n$ beschrieben, $1 \leq i \leq k$.

Sei $A \in \mathbb{R}^{k\times n}$ definiert durch $A_{ij} := f_i(v_j)$. Für $b \in \mathbb{R}^{k\times 1}$ sei

$$L(A,\leq b) := \{x \in \mathbb{R}^{n\times 1} \mid Ax \leq b\}.$$

Dann ist $L(f_1,\dots,f_k,\leq b_1,\dots,b_k) = \{\underline{v}x \mid x \in L(A,\leq b)\}$.

Beispiel 48 : Die Aufgabe “Finde alle $x \in \mathbb{R}^4$ mit

$$\begin{aligned} 2x_1 + x_2 - x_3 + x_4 &\leq 2, \\ x_1 - 2x_2 + 3x_4 &\geq 1, \\ \text{und } x_1 - x_2 + x_3 - x_4 &= 1 \text{ ”} \end{aligned}$$

ist System linearer Ungleichungen, weil die Menge dieser $x \in \mathbb{R}^4$ gleich $L(A,\leq (2,-1,1,-1)^T)$ ist, wobei

$$A = \begin{pmatrix} 2 & 1 & -1 & 1 \\ -1 & 2 & 0 & -3 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \end{pmatrix} \in \mathbb{R}^{4\times 4}.$$

(“Jedes System von linearen Gleichungen und Ungleichungen kann in ein System der Form $Ax \leq b$ umgeschrieben werden”).

Definition 49 : Eine nichtleere Teilmenge L von V heißt *Kegel (in V)*, wenn jede nichtnegative Linearkombination von Elementen in L wieder in L liegt, d. h.: Für alle $\ell \in \mathbb{N}$, $v_0,\dots,v_\ell \in L$, $c_0,\dots,c_\ell \in \mathbb{R}_{\geq 0}$ ist $c_0v_0 + \cdots + c_\ell v_\ell \in L$.

Für $\ell \in \mathbb{N}$ und $v_0, \ldots, v_\ell \in V$ heißt die Menge

$$\mathcal{K}(v_0, \ldots, v_\ell) := \big\{ \sum_{i=0}^{\ell} c_i v_i \mid c_0, \ldots, c_\ell \in \mathbb{R}_{\geq 0} \big\}$$

der von $v_0, \ldots, v_\ell$ erzeugte Kegel. Ein Kegel L ist *endlich erzeugt*, wenn es $\ell \in \mathbb{N}$ und $v_0, \ldots, v_\ell \in V$ gibt, so dass $L = \mathcal{K}(v_0, \ldots, v_\ell)$.

Beispiel 50 : Sei $\dim_{\mathbb{R}}(V) = 1$ und $0 \neq v \in V$. Dann gibt es in V genau vier Kegel und zwar $\{0\}$, $\mathcal{K}(v)$, $\mathcal{K}(-v)$ und V.

Beispiel 51 : Sei $\dim_{\mathbb{R}}(V) = 2$ und L ein Kegel in V mit $\{0\} \neq L \neq V$. Dann gibt es eine Basis (v_1, v_2) von V so, dass $L = \mathcal{K}(v_1)$ oder $L = \mathcal{K}(v_1, -v_1)$ oder $L = \mathcal{K}(v_1, v_2)$ oder $L = \mathcal{K}(v_1, v_2, -v_1)$ ist.

Beispiel 52 : Die Menge

$$\{z \in \mathbb{R}^3 \mid z_1^2 + z_2^2 \leq z_3^2,\, z_3 \geq 0\}$$

ist ein Kegel in $\mathbb{R}^3$, aber nicht endlich erzeugt.

Definition 53 : Eine Teilmenge M von V ist *konvex*, wenn für je zwei Elemente $v, w \in M$ auch die Strecke $[v, w]$ in M enthalten ist.

Beispiel 54 : Polyeder sind konvex.

Beispiel 55 : Die konvexe Hülle einer Menge N ist konvex.
Denn: Seien $\sum_{i=0}^{\ell} c_i v_i$ und $\sum_{j=0}^{m} d_j w_j$ konvexe Linearkombinationen von Elementen in N und $s, t \in \mathbb{R}_{\geq 0}$ mit $s + t = 1$. Wegen

$$\sum_{i=0}^{\ell} s c_i + \sum_{j=0}^{m} t d_j = s\big(\sum_{i=0}^{\ell} c_i\big) + t\big(\sum_{j=0}^{m} d_j\big) = s + t = 1$$

ist $s(\sum_{i=0}^{\ell} c_i v_i) + t(\sum_{j=0}^{m} d_j w_j) \in \operatorname{conv}(N)$.

Satz 56 : *Sind $v_0, \ldots, v_\ell$ Elemente einer konvexen Menge M in V, dann enthält M auch alle konvexen Linearkombinationen von $v_0, \ldots, v_\ell$.*

Insbesondere: Die konvexe Hülle $\operatorname{conv}(N)$ einer Teilmenge N von V ist die (bezüglich Inklusion) kleinste konvexe Teilmenge von V, die N enthält.

Beweis: Induktion über ℓ.

$\ell = 1$: $\operatorname{conv}(\{v_0, v_1\}) = [v_0, v_1] \subseteq M$.

$\ell > 1$: Sei $\sum_{i=0}^{\ell} c_i v_i$ eine konvexe Linearkombination von $v_0, \ldots, v_\ell$ und $s := \sum_{i=1}^{\ell-1} c_i > 0$. Dann ist $w := \sum_{i=1}^{\ell-1} \frac{c_i}{s} v_i$ eine konvexe Linearkombination von $v_0, \ldots, v_{\ell-1}$, nach Induktionsannahme also ein Element von M. Da M konvex und $c_\ell = 1 - s$ ist folgt

$$\sum_{i=0}^{\ell} c_i v_i = sw + (1-s) v_\ell \in M.$$

Definition 57 : Für Teilmengen $A \subseteq V$, $B \subseteq V$ sei

$$A + B := \{v + w \mid v \in A,\ w \in B\}$$

die *Summe von A und B*.

§3. Lineare Optimierung mit zwei Unbekannten

Beispiel 58 : Eine Firma stellt aus 3 Rohstoffen A, B und C zwei Produkte P und Q her.
Von A bzw. B bzw. C sind 28 bzw. 32 bzw. 40 Tonnen verfügbar.
Für je ein Stück von P bzw. Q werden gebraucht:

	A	B	C
P	1kg	2kg	5kg
Q	4kg	4kg	2kg

Der Gewinn für ein Stück von P bzw. Q beträgt 4 bzw. 5 Euro.

Wie viele Stück von P und von Q soll die Firma produzieren, um einen möglichst großen Gewinn zu erzielen?

Um diese Aufgabe mathematisch zu modellieren, legen wir zuerst fest, was wir suchen. Gesucht sind zwei Zahlen: die Stückzahlen von Produkt P und von Produkt Q. Dann lesen wir aus dem Text ab, welche Bedingungen diese Zahlen erfüllen sollen. Weiters müssen wir Annahmen über die „Gewinnfunktion“ treffen. Also:

- *Gesucht* ist ein Zahlenpaar $(p, q) \in \mathbb{R}^2$, dabei gibt p bzw. q an, wieviele Stück von P bzw. Q hergestellt werden sollen.

- Dieses Zahlenpaar (p,q) muss die folgenden *Bedingungen* erfüllen :
  - $p+4q \leq 28000,\ 2p+4q \leq 32000,\ 5p+2q \leq 40000$, weiters $p \geq 0$ und $q \geq 0$, weil p und q als Stückzahlen nicht negativ sein können.
    Die Menge der Zahlenpaare (p,q), die diese Bedingungen erfüllen, nennen wir den *zulässigen Bereich* dieser Optimierungsaufgabe.
  - Der Gewinn soll möglichst groß sein, d.h.: wenn wir mit G die Funktion von $\mathbb{R}^2$ nach $\mathbb{R}$ bezeichnen, die jedem Zahlenpaar (x,y) den Gewinn (in Euro) bei Produktion von x Stück von P und y Stück von Q zuordnet, dann soll für alle (x,y) aus dem zulässigen Bereich
    $G(x,y) \leq G(p,q)$ sein.
- Wenn wir annehmen, dass die Funktion G linear ist, dann ist $G(x,y) = 4x+5y$ für alle Zahlenpaare (x,y), also soll das Zahlenpaar (p,q) im zulässigen Bereich so gewählt werden, dass $4p+5q$ größtmöglich ist.

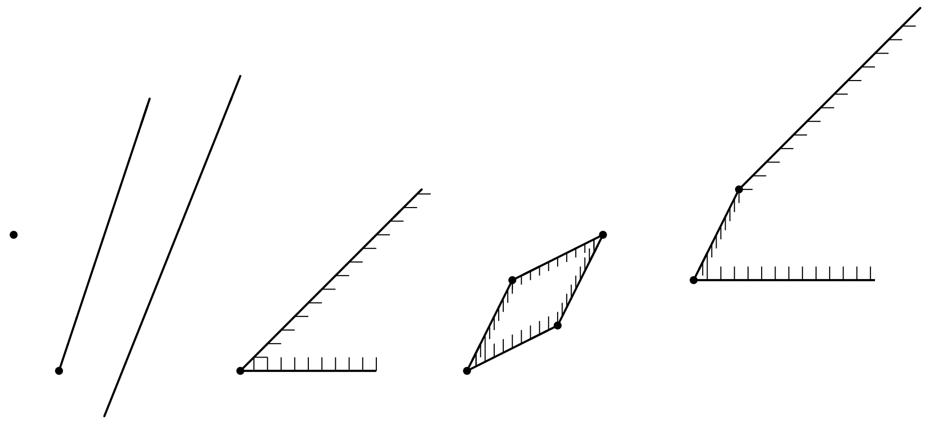
Ist die Annahme, dass die Gewinnfunktion linear ist, sinnvoll? Diese Annahme bedeutet, dass sich der Gewinn verdoppelt, verdreifacht, ... , wenn man die Produktion verdoppelt, verdreifacht, ... Jede Firma weiß, dass das in der Regel nicht der Fall ist. Wenn der Markt mit einem Produkt „überschwemmt“ wird, verfällt der Preis. Wenn die Gewinnfunktion linear ist, dann muss der Gewinn bei gleichzeitiger Produktion von p Stück von P und q Stück von Q die Summe der Gewinne nach der Produktion von p bzw. q Stück von P bzw. Q sein. Auch das ist nicht immer so. Wenn die Kunden zum Beispiel das Produkt P besser als das Produkt Q empfinden, dann kann die Produktion von P den Absatz von Q beeinträchtigen.

Die Annahme, dass die Gewinnfunktion linear ist, ist jedoch unter gewissen Umständen sinnvoll. Zum Beispiel, wenn die Abnahme der Produkte von vorneherein gesichert ist. Oder, wenn nur kleine Veränderungen der Produktion geplant sind. Die Annahme, dass der Gewinn um 2 Prozent steigt, wenn die Produktion um 2 Prozent erhöht wird, scheint realistisch zu sein.

Definition 59 : Ein *lineares Programm* auf $\mathbb{R}^2$ ist die folgende Aufgabe:

- *Gegeben* sind endlich viele lineare Ungleichungen mit zwei Unbekannten und eine lineare Funktion $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$.
  Die Funktion f nennt man *Zielfunktion*, den Durchschnitt der Lösungsmengen aller linearen Ungleichungen nennt man den *zulässigen Bereich*.
- *Gesucht* ist ein *optimaler Punkt*, das ist ein Zahlenpaar im zulässigen Bereich mit größtmöglichem (oder kleinstmöglichem) Funktionswert bezüglich f.

Nach §2 ist der zulässige Bereich Durchschnitt von Halbebenen. Er kann leer, ein Punkt, eine Halbgerade, eine Gerade, eine Halbebene, ein spitzer Kegel (in der Ebene), ein konvexes Vieleck oder die Summe eines spitzen Kegels (in der Ebene) und eines konvexen Vielecks sein.



Ein optimaler Punkt kann *graphisch* wie folgt bestimmt werden:

- Für jede lineare Ungleichung $ax+by \leq c$ zeichne die Gerade

$$L((a,b),c) = \{\,(x,y)\,|\,ax+by = c\,\}$$

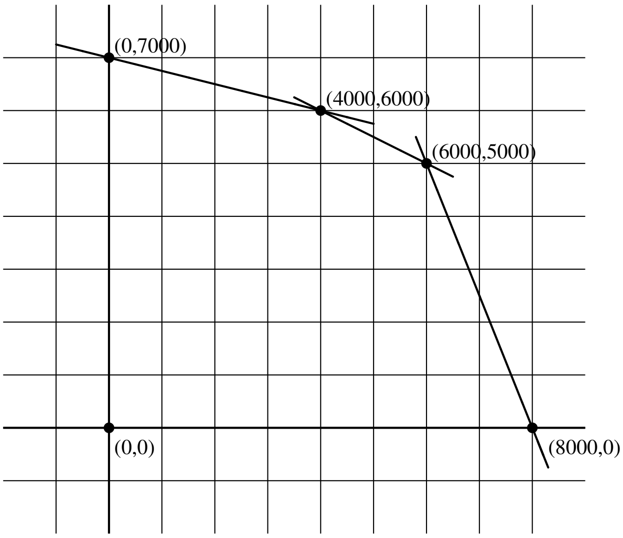
  ein.
  Wähle dann irgendeinen Punkt (u,v), der nicht auf dieser Geraden liegt, und berechne $au+bv$. Wenn diese Zahl kleiner als c ist, dann liegt dieser Punkt auf der Halbebene $L((a,b),\leq c)$. Sonst ist $L((a,b),\leq c)$ die andere Halbebene.

- Bestimme durch Zeichnung den Durchschnitt dieser Halbebenen, den zulässigen Bereich.
  Falls er leer ist, gibt es keinen optimalen Punkt. Wir nehmen daher im Folgenden an, dass der zulässige Bereich nicht leer ist.

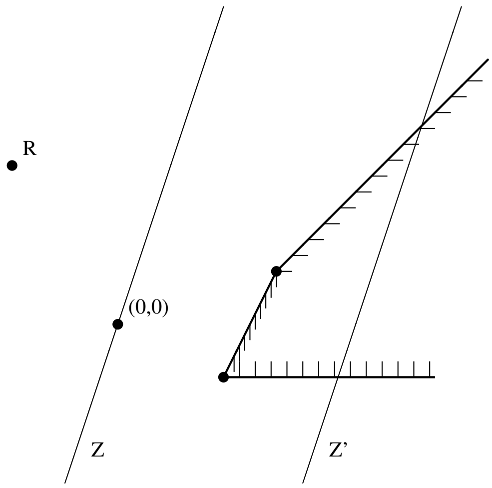
Beispiel: Als Lösungsmenge des Systems

$$\begin{aligned} x+4y &\leq 28000 \\ 2x+4y &\leq 32000 \\ 5x+2y &\leq 40000 \\ -x &\leq 0 \\ -y &\leq 0 \end{aligned}$$

linearer Ungleichungen erhalten wir den folgenden umrandeten Bereich:

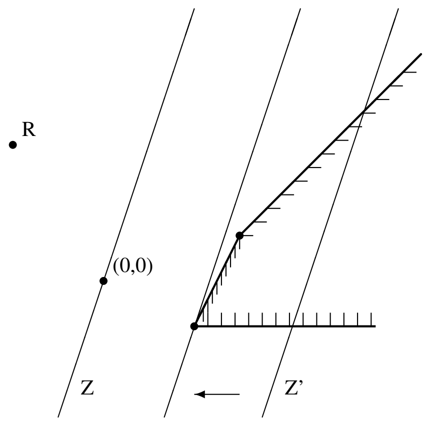


- Zeichne das Urbild von 0 bezüglich der Zielfunktion. Dieses ist eine Gerade in $\mathbb{R}^2$ durch $(0,0)$, wir nennen sie Z.
- Wir verschieben Z zunächst so, dass die verschobene Gerade Z' einen Punkt des zulässigen Bereichs enthält.
- Wähle einen Punkt R, dessen Funktionswert bezüglich der Zielfunktion positiv ist (o.E.d.A. nehmen wir an, dass die Zielfunktion nicht die Nullfunktion ist).

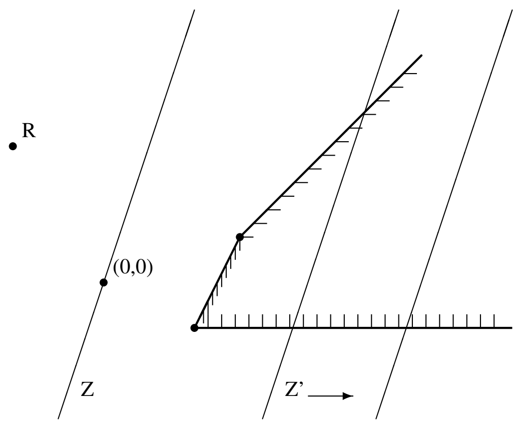


- Wenn wir ein *Maximum* suchen, dann verschieben wir Z' so lange in die Richtung von $(0,0)$ nach R, wie noch ein zulässiger Punkt auf der verschobenen Geraden liegt. Die Funktionswerte bezüglich der Zielfunktion der Punkte auf den Geraden werden dabei immer größer.

– Falls das beliebig weit möglich ist, gibt es keinen optimalen Punkt.
– Sobald das nicht mehr möglich ist, ist jeder Punkt des zulässigen Bereichs, der dann noch auf dieser Geraden liegt, ein optimaler Punkt.



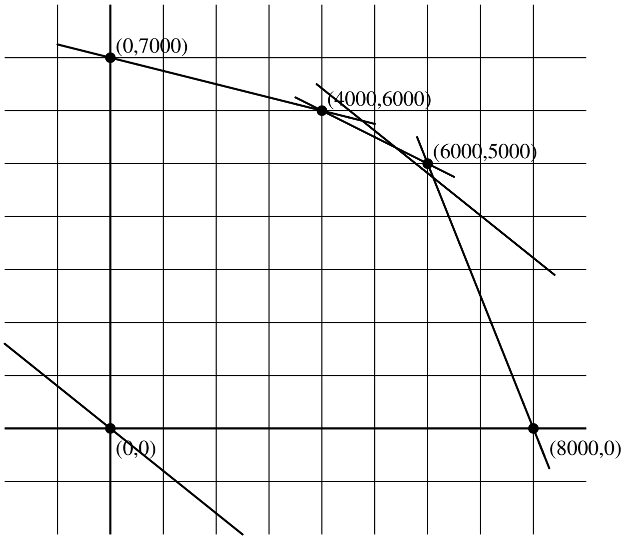
- Wenn wir ein *Minimum* suchen, dann verschieben wir Z' in die andere Richtung und gehen analog vor.



Beispiel 60 : Wir lösen nun das lineare Programm aus Beispiel 58

$$x+4y \le 28000\,,\ 2x+4y \le 32000\,,\ 5x+2y \le 40000\,,\ -x \le 0\,,\ -y \le 0\,,$$
$$4x+5y \text{ maximal}$$

Das Urbild von 0 bezüglich der Zielfunktion ist die Gerade $Z := \mathbb{R}(-5,4)$. Diese Gerade enthält bereits einen Punkt des zulässigen Bereichs, nämlich $(0,0)$. Werten wir die Zielfunktion zum Beispiel in $(1,1)$ aus, erhalten wir die positive Zahl 9. Also verschieben wir die Gerade Z solange in der Richtung von $(0,0)$ nach $(1,1)$, wie die verschobene Gerade noch Punkte des zulässigen Bereichs enthält. Schließlich liegt nur noch der Punkt $(6000,5000)$ auf der verschobenen Geraden. Also ist der Gewinn maximal, wenn 6000 Stück von Produkt P und 5000 Stück von Produkt Q hergestellt werden.



Aus den Überlegungen oben folgt unmittelbar: Wenn es einen optimalen Punkt gibt, dann ist mindestens einer der Randpunkte des zulässigen Bereichs ein optimaler Punkt. Das gilt auch für lineare Programme auf $\mathbb{R}^n$ mit $n \geq 3$ und ist die Grundlage für den *Simplexalgorithmus* zu ihrer Lösung

KAPITEL 3

Bewegungen in euklidischen Räumen

Es seien V mit $\langle -,- \rangle$ ein euklidischer Raum und $n \in \mathbb{N}$ seine Dimension. In diesem Kapitel werden „Bewegungen starrer Körper“ (zum Beispiel eines Roboterarms) mathematisch beschrieben. Dabei verstehen wir unter „Bewegung“ eines Körpers in V nicht einen „in der Zeit ablaufenden Vorgang“, sondern eine Funktion $f : V \longrightarrow V$, die dem „Ort“ $v \in V$ eines „Massenpunktes“ zur „Zeit 0“ seinen „Ort“ $f(v)$ zur „Zeit 1“ zuordnet. Ein Körper ist „starr“, wenn nach jeder Bewegung die Abstände je zweier seiner „Massenpunkte“ gleichgeblieben sind.

§1. Isometrien

Definition 61: Eine Funktion $f : V \to V$ heißt *Isometrie*, wenn für alle v, $w \in V$

$$\|f(v) - f(w)\| = \|v - w\|$$

ist. („Isometrien erhalten Abstände.“)
Eine Funktion $f : V \to V$ heißt *orthogonal*, wenn für alle v, $w \in V$

$$\langle f(v), f(w) \rangle = \langle v, w \rangle$$

ist. („Orthogonale Funktionen erhalten das Skalarprodukt.“)

Satz 62: *Jede Translation in V ist eine Isometrie. Die Identität Id_V ist die einzige orthogonale Translation.*

Beweis: Sei t eine Translation und $u := t(0)$. Dann ist

$$\|t(v) - t(w)\| = \|(v+u) - (w+u)\| = \|v - w\|.$$

Wenn t orthogonal ist, dann ist

$$\langle u, u \rangle = \langle t(0), t(0) \rangle = \langle 0, 0 \rangle = 0,$$

also $u = 0$ und $t = \mathrm{Id}_V$.

Satz 63: *Für eine Funktion $f : V \to V$ sind die folgenden Aussagen äquivalent:*

(1) *f ist orthogonal.*
(2) *f ist eine Isometrie und linear.*
(3) *f ist eine Isometrie und $f(0) = 0$.*

Beweis:

(1) $\Rightarrow$ (2) : Für alle v, $w \in V$ ist

$$\|f(v)-f(w)\|^2 = \langle f(v)-f(w), f(v)-f(w)\rangle =$$
$$= \langle f(v), f(v)\rangle - 2\langle f(v), f(w)\rangle + \langle f(w), f(w)\rangle \underset{(1)}{=}$$
$$= \langle v, v\rangle - 2\langle v, w\rangle + \langle w, w\rangle = \langle v-w, v-w\rangle = \|v-w\|^2,$$

also ist f eine Isometrie.
Wir zeigen noch, dass f linear ist, das heißt: für alle v, $w \in V$ und für alle c, $d \in \mathbb{R}$ ist $f(cv+dw)-cf(v)-df(w)=0$.
Dazu genügt es zu zeigen, dass

$$\|f(cv+dw)-cf(v)-df(w)\|^2 = 0 \text{ ist:}$$

$$\|f(cv+dw)-cf(v)-df(w)\|^2 =$$
$$\langle f(cv+dw)-cf(v)-df(w), f(cv+dw)-cf(v)-df(w)\rangle =$$
$$= \langle f(cv+dw), f(cv+dw)\rangle + c^2\langle f(v), f(v)\rangle +$$
$$+d^2\langle f(w), f(w)\rangle - 2c\langle f(cv+dw), f(v)\rangle -$$
$$-2d\langle f(cv+dw), f(w)\rangle + 2cd\langle f(v), f(w)\rangle \underset{(1)}{=}$$
$$\underset{(1)}{=} \langle cv+dw, cv+dw\rangle + c^2\langle v, v\rangle + d^2\langle w, w\rangle -$$
$$-2c\langle cv+dw, v\rangle - 2d\langle cv+dw, w\rangle + 2cd\langle v, w\rangle =$$
$$= \langle (cv+dw)-cv-dw, (cv+dw)-cv-dw\rangle = \langle 0, 0\rangle = 0$$

(2) $\Rightarrow$ (3) : trivial

(3) $\Rightarrow$ (1) : Für alle $u \in V$ ist

$$\langle f(u), f(u)\rangle = \|f(u)\|^2 = \|f(u)-0\|^2 =$$
$$\|f(u)-f(0)\|^2 = \|u-0\|^2 = \|u\|^2 = \langle u, u\rangle.$$

Daher gilt für alle v, $w \in V$:

$$\langle f(v), f(w)\rangle = \frac{1}{2}\left(\langle f(v), f(v)\rangle +\right.$$
$$\left.+\langle f(w), f(w)\rangle - \langle f(v)-f(w), f(v)-f(w)\rangle\right) =$$
$$= \frac{1}{2}\left(\langle v, v\rangle + \langle w, w\rangle - \|f(v)-f(w)\|^2\right) \underset{(3)}{=}$$
$$\underset{(3)}{=} \frac{1}{2}\left(\|v\|^2 + \|w\|^2 - \|v-w\|^2\right) = \langle v, w\rangle.$$

Satz 64 :

(1) *Die Zusammensetzung von Isometrien ist eine Isometrie.*

(2) *Jede affine Funktion mit orthogonalem linearen Anteil ist eine Isometrie.*

(3) *Jede Isometrie $f : V \to V$ ist eine affine Funktion, ihr Translationsanteil ist die Translation um $f(0)$ und ihr linearer Anteil ist orthogonal. Kurz: $f = t_{f(0)} \circ g$ mit $g : V \to V$ orthogonal (und damit linear).*

(4) *Die Menge aller Isometrien von V ist mit der Zusammensetzung von Funktionen eine Gruppe und heißt* Isometriegruppe von V. *Inbesondere ist jede Isometrie $f = t_u \circ g$ bijektiv und es ist*

$$f^{-1} = t_{-g^{-1}(u)} \circ g^{-1} .$$

Beweis:

(1) Es seien f, h Isometrien von V und v, $w \in V$. Dann ist

$$\begin{aligned}\|(f \circ h)(v) - (f \circ h)(w)\| &= \|f(h(v)) - f(h(w))\| = \\ &= \|h(v) - h(w)\| = \|v - w\| ,\end{aligned}$$

also $f \circ h$ eine Isometrie.

(2) Nach Satz 62 und Satz 63 sind Translationen und orthogonale Funktionen Isometrien, also auch ihre Zusammensetzung.

(3) Sei $g := t_{-f(0)} \circ f$. Dann ist $f = t_{f(0)} \circ g$. Nach (1) ist g eine Isometrie, weiters ist $g(0) = -f(0) + f(0) = 0$. Nun folgt aus Satz 63, dass g orthogonal und linear ist.

(4) Es sei $g : V \to V$ der lineare Anteil einer Isometrie. Nach (3) ist g orthogonal.
Ist $v \in V$ so, dass $g(v) = 0$ ist, dann ist $\langle v, v\rangle = \langle g(v), g(v)\rangle = \langle 0, 0\rangle = 0$, daher $v = 0$. Also ist g injektiv. Weil Definitions- und Bildbereich von g dieselbe Dimension haben, ist g auch bijektiv.
Für alle $v, w \in V$ ist

$$\left\langle g^{-1}(v), g^{-1}(w)\right\rangle = \left\langle g(g^{-1}(v)), g(g^{-1}(w))\right\rangle = \langle v, w\rangle ,$$

daher ist g^{-1} orthogonal. Nun prüft man leicht nach, dass die Isometrie $t_{-g^{-1}(u)} \circ g^{-1}$ die Umkehrfunktion von $t_u \circ g$ ist.

Satz 65 :

(1) *Die Menge $T(V)$ aller Translationen von V ist eine Untergruppe der Isometriegruppe und heißt* Translationsgruppe von V.

(2) *Die Menge $\mathcal{O}(V)$ aller orthogonalen Funktionen von V ist eine Untergruppe sowohl der Isometriegruppe als auch der Gruppe $\mathrm{GL}_{\mathbb{R}}(V)$ und heißt* orthogonale Gruppe von V.

(3) *Die Translationsgruppe $T(V)$ ist kommutativ, die orthogonale Gruppe $\mathcal{O}(V)$ ist für $n \geq 2$ nicht kommutativ.*

Beweis: Übung.

§2. Orthogonale Funktionen

Es sei $f : V \to V$ eine lineare Funktion.

Satz 66 : *Es seien $\underline{v}$ eine ON-Basis von V und $A \in \mathbb{R}^{n\times n}$ die Matrix von f bezüglich $\underline{v}$. Dann sind die folgenden Aussagen äquivalent:*

(1) *Die Funktion f ist orthogonal.*
(2) *Die Familie $f(\underline{v}) := (f(v_1), ..., f(v_n))$ ist eine ON-Basis von V.*
(3) *Die Matrix A ist orthogonal.*

Insbesondere gilt: Wenn f orthogonal ist, dann ist A^T die Matrix von f^{-1} bezüglich $\underline{v}$.

Beweis:

(1) $\Rightarrow$ (2) : Wenn f orthogonal ist, dann ist

$$\left\langle f(v_i), f\left(v_j\right)\right\rangle = \left\langle v_i, v_j\right\rangle = \delta_{ij},\ 1 \le i, j \le n.$$

Daher ist die Familie $f(\underline{v})$ orthonormal, also auch linear unabhängig. Weil V n-dimensional ist, folgt daraus die Behauptung.

(2) $\Rightarrow$ (3) : Wenn $f(\underline{v}) = \underline{v}A$ eine ON-Basis ist, dann ist A orthogonal.

(3) $\Rightarrow$ (1) : Wenn A orthogonal ist, dann ist $\underline{v}A$ eine ON-Basis. Für $y, z \in \mathbb{R}^{n\times 1}$ ist daher

$$\langle f(\underline{v}y), f(\underline{v}z)\rangle = \langle (\underline{v}A)y, (\underline{v}A)z\rangle = \langle y, z\rangle = \langle \underline{v}y, \underline{v}z\rangle .$$

Somit ist f orthogonal.

Satz 67 : *Wenn f orthogonal ist, dann ist*

$$\det(f) \in \{1, -1\} \textit{ und } \mathrm{Sp}_{\mathbb{R}}(f) \subseteq \{1, -1\}.$$

Beweis: Sei $A \in \mathbb{R}^{n\times n}$ die Matrix von f bezüglich einer ON-Basis. Nach Satz 66 ist $\det(f) = \det(A) \in \{1, -1\}$.
Sei v ein Eigenvektor von f zum Eigenwert $c \in \mathbb{R}$, dann ist
$0 \neq \langle v, v\rangle = \langle f(v), f(v)\rangle = \langle cv, cv\rangle = c^2 \langle v, v\rangle$, also $c^2 = 1$.

Definition 68 : Ein Untervektorraum W von V ist genau dann *unter f stabil*, wenn $f(W) \subseteq W$ ist.

Satz 69 : *Es seien f orthogonal und W ein unter f stabiler Untervektorraum von V. Dann ist auch $W^\perp$ stabil unter f.*

Beweis: Sei $v \in W^{\perp}$. Wir zeigen, dass auch $f(v) \in W^{\perp}$ ist.
Für alle $w \in W$ ist $f(w) \in W$ und, weil f linear und bijektiv ist, auch $f^{-1}(w) \in W$. Daher ist

$$\langle f(v), w\rangle = \left\langle f^{-1}(f(v)), f^{-1}(w)\right\rangle = \left\langle v, f^{-1}(w)\right\rangle = 0.$$

Beispiel 70 : Sei $A \in \mathbb{R}^{n\times n}$ eine Matrix in oberer Blockdreiecksform

$$\begin{pmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{pmatrix}.$$

Die Blockgröße von B_{11} sei k.
Betrachten wir A als lineare Funktion

$$A : \mathbb{R}^{n\times 1} \to \mathbb{R}^{n\times 1} \quad , \quad x \longmapsto Ax\ ,$$

dann ist der von $e_1, ..., e_k$ erzeugte Untervektorraum stabil unter A.
Wenn A orthogonal ist, ist auch

$${}_{\mathbb{R}}\langle e_1, ..., e_k\rangle^{\perp} = {}_{\mathbb{R}}\langle e_{k+1}, ..., e_n\rangle$$

stabil unter A, also $B_{12} = 0$. Daher hat jede orthogonale Matrix in Blockdreiecksform sogar Blockdiagonalgestalt.

Definition 71 : Es sei $M \in \mathbb{C}^{m\times n}$ eine komplexe Matrix. Dann heißt

$$\mathrm{Re}(M) := \left(\mathrm{Re}\left(M_{ij}\right)\right)_{i,j} \in \mathbb{R}^{m\times n} \text{ bzw. } \mathrm{Im}(M) := \left(\mathrm{Im}\left(M_{ij}\right)\right)_{i,j} \in \mathbb{R}^{m\times n}$$

der *Realteil* bzw. *Imaginärteil* von M.

Hilfssatz 72 :

(1) *Sei $A \in \mathbb{R}^{n\times n}$. Dann gibt es n-Spalten $0 \neq x, y \in \mathbb{R}^{n\times 1}$ so, dass $Ax, Ay \in {}_{\mathbb{R}}\langle x, y\rangle$ ist.*
(2) *Ist $h : V \to V$ eine $\mathbb{R}$-lineare Funktion, dann gibt es einen ein- oder zweidimensionalen Untervektorraum von V, der unter h stabil ist.*

Beweis:

(1) Wegen $A \in \mathbb{R}^{n\times n} \subseteq \mathbb{C}^{n\times n}$ gibt es einen Eigenvektor $z \in \mathbb{C}^{n\times 1}$ von A. Sei $c \in \mathbb{C}$ der Eigenwert von z. Dann sind
$\mathrm{Re}(z), \mathrm{Im}(z) \in \mathbb{R}^{n\times 1}$ und $A(\mathrm{Re}(z)) = A\left(\frac{1}{2}(z+\overline{z})\right) =$
$= \frac{1}{2}(Az + A\overline{z})$ =(weil A reell ist) $\frac{1}{2}\left(Az + \overline{Az}\right) = \mathrm{Re}(Az) =$
$= \mathrm{Re}(cz) = \mathrm{Re}(c)\mathrm{Re}(z) - \mathrm{Im}(c)\mathrm{Im}(z) \in {}_{\mathbb{R}}\langle \mathrm{Re}(z), \mathrm{Im}(z)\rangle$.
Analog wird gezeigt: $A(\mathrm{Im}(z)) \in {}_{\mathbb{R}}\langle \mathrm{Re}(z), \mathrm{Im}(z)\rangle$.
(2) Sei $\underline{v}$ eine Basis von h und A die Matrix von h bezüglich $\underline{v}$. Seien x, y reelle n-Spalten wie in (1) und W der von den Vektoren $\underline{v}x$ und $\underline{v}y$ erzeugte Untervektorraum von V. Dann ist $f(\underline{v}x) = \underline{v}Ax =$

$\underline{v}(cx+dy) = c(\underline{v}x) + d(\underline{v}y) \in W$, wobei c, d geeignete reelle Zahlen sind.

Satz 73: *(„Spektralsatz für orthogonale Funktionen")*

(1) *Die Funktion f sei orthogonal. Dann gibt es paarweise aufeinander senkrecht stehende, unter f stabile ein- oder zweidimensionale Untervektorräume $W_1, \ldots, W_k$ von V so, dass*

$$\bigoplus_{i=1}^{k} W_i = V$$

ist.

(2) *Sei $A \in \mathbb{R}^{n\times n}$ eine orthogonale Matrix. Dann gibt es eine orthogonale Matrix $T \in \mathbb{R}^{n\times n}$ so, dass $T^{-1} \cdot A \cdot T$ Blockdiagonalform mit Blockgrössen ≤ 2 und orthogonalen Blöcken in der Diagonale hat.*

Beweis: (2) folgt aus (1). Wir zeigen (1) durch Induktion nach n.
Für $n \leq 2$ ist nichts zu zeigen.
Sei $n > 2$. Nach Hilfssatz 72 gibt es einen ein- oder zweidimensionalen Untervektorraum W_1 von V, der unter f stabil ist. Weil f orthogonal ist, ist auch $W_1^{\perp}$ stabil unter f (Satz 69). Wegen $V = W_1 \oplus W_1^{\perp}$ und $\dim_{\mathbb{R}}\left(W_1^{\perp}\right) < \dim_{\mathbb{R}}(V) = n$ folgt die Behauptung aus der Induktionsannahme.

§3. Spiegelungen

Definition 74: Es sei $h : V \to V$ eine Funktion. Die Menge

$$Fix(h) := \{v \in V \mid h(v) = v\}$$

heißt *Menge der Fixpunkte* (oder kurz: *Fixmenge*) von h.

Beispiel 75: $Fix(-Id_V) = \{0\}, Fix(Id_V) = V$.
Die Fixmenge einer linearen Funktion ist ihr Eigenraum zum Eigenwert 1.
Die Fixmenge einer Translation t_u mit $u \neq 0$ ist leer.

Hilfssatz 76: *Es seien $g : V \to V$ eine lineare Funktion und $u \in V$. Die Fixmenge der affinen Funktion $t_u \circ g$ ist entweder leer oder ein affiner Unterraum von V, dessen paralleler Untervektorraum ist $Fix(g)$.*

Beweis: Die Fixmenge von $t_u \circ g$ ist das Urbild von 0 bezüglich der affinen Funktion $t_u \circ g - Id_V = t_u \circ (g - Id_V)$, also leer oder ein affiner Unterraum. Der zu diesem parallele Untervektorraum ist dann $\mathrm{Kern}(g - Id_V)$, also $Fix(g)$.

Definition 77 : Eine Isometrie von V heißt *Spiegelung in V*, wenn ihre Fixmenge eine Hyperebene in V ist, das heißt: die Dimension ihrer Fixmenge ist $n-1$.
Ist M die Fixmenge einer Spiegelung, dann heißt diese *Spiegelung um M*.

Beispiel 78 : Die Funktion

$$\mathbb{R}^{n\times 1} \to \mathbb{R}^{n\times 1} \quad , \quad x \longmapsto \begin{pmatrix} 1 & & & & 0 \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ 0 & & & & -1 \end{pmatrix} x \, ,$$

ist eine Spiegelung in $\mathbb{R}^n$.

Satz 79 : *Es sei s eine Spiegelung.*

(1) *Der lineare Anteil von s ist auch eine Spiegelung, seine Fixmenge ist der parallele Untervektorraum von $Fix(s)$.*
(2) *Wenn s linear ist, dann ist s orthogonal und es gibt eine ON-Basis von V, bezüglich der s die Matrix*

$$\begin{pmatrix} 1 & & & & 0 \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ 0 & & & & -1 \end{pmatrix}$$

hat. Insbesondere ist $\det(s) = -1$ *und* $Sp_{\mathbb{R}}(s) = \{1, -1\}$.
(3) *Zwei Spiegelungen sind genau dann gleich, wenn ihre Fixmengen gleich sind.*
(4) *Es ist $s^2 = Id_V$.*

Beweis:

(1) Nach Satz 64 ist der lineare Anteil von s eine Isometrie, nach Hilfssatz 76 ist seine Fixmenge der parallele Untervektorraum von $Fix(s)$, insbesondere eine Hyperebene in V.
(2) Wähle eine ON-Basis $(v_1, \ldots, v_{n-1})$ von $Fix(s)$ und ergänze sie zu einer ON-Basis $(v_1, \ldots, v_{n-1}, v_n)$ von V. Nach Satz 63 ist s orthogonal, daraus folgt mit Satz 66 , dass auch

$$(s(v_1), \ldots, s(v_n)) = (v_1, \ldots, v_{n-1}, s(v_n))$$

eine ON-Basis von V ist. Daher muss $s(v_n) = v_n$ oder $s(v_n) = -v_n$ sein. Wegen $Fix(s) \neq V$ muss $s(v_n) = -v_n$ sein.
(3) Es seien $t_v \circ f,\ t_w \circ g$ zwei Isometrien (mit linearen Anteilen f, g), deren Fixmengen gleich sind. Aus (1) folgt, dass $Fix(f) = Fix(g)$

ist. Nun folgt aus (2), dass $f = g$.
Sei u ein Element von $Fix(t_v \circ f) = Fix(t_w \circ g)$. Dann ist
$u = (t_v \circ f)(u) = v + f(u)$ und $u = (t_w \circ f)(u) = w + f(u)$,
also $v = u - f(u) = w$.

(4) Es seien f linear und $s = t_u \circ f$. Nach (2) ist $f \circ f = Id_V$.
Daher ist $s \circ s = t_u \circ f \circ t_u \circ f = t_{u+f(u)}$, somit ist $s \circ s$ eine Translation.
Wegen $\emptyset \neq Fix(s) \subseteq Fix(s \circ s)$ muss $s \circ s = Id_V$ sein.

Satz 80 : *Zu jeder Hyperebene T von V gibt es genau eine Spiegelung s mit $Fix(s) = T$, und zwar*

$$s : V \to V \quad , \quad v \longmapsto 2 \cdot p_T(v) - v \, .$$

Dabei ist p_T die orthogonale Projektion von V auf T.

Beweis: Es ist zu zeigen, dass s eine Isometrie ist. Die Eindeutigkeit folgt aus Aussage (3) in Satz 79.
Es seien $x \in T$, U der parallele Untervektorraum von T und $y := x - p_U(x)$. Dann ist $p_T = t_y \circ p_U$ und $s = 2t_y \circ p_U - Id_V = t_{2y} \circ (2p_U - Id_V)$. Daher genügt es zu zeigen, dass $2p_U - Id_V$ orthogonal ist.

Für alle $v \in V$ ist

$$\begin{aligned}(2p_U - Id_V)(v) &= p_U(v) - (v - p_U(v)),\\ v &= p_U(v) + (v - p_U(v))\end{aligned}$$

und $p_U(v) \in U$, $(v - p_U(v)) \in U^{\perp}$.
Für alle $v, w \in V$ ist daher

$$\begin{aligned}&\langle (2p_U - Id_V)(v), (2p_U - Id_V)(w) \rangle =\\ &\langle p_U(v) - (v - p_U(v)), p_U(w) - (w - p_U(w)) \rangle =\\ &\langle p_U(v), p_U(w) \rangle + \langle v - p_U(v), w - p_U(w) \rangle =\\ &\langle p_U(v) + (v - p_U(v)), p_U(w) + (w - p_U(w)) \rangle = \langle v, w \rangle \, .\end{aligned}$$

Satz 81 : *Jede orthogonale Funktion von V nach V ist Produkt von höchstens n Spiegelungen.*

Beweis: Induktion nach n.

Wenn $n = 1$ ist, dann gibt es nur zwei orthogonale Funktionen, nämlich Id_V und die Spiegelung $-\mathrm{Id}_V$. Die erste ist das Produkt von 0 Spiegelungen, die zweite von einer Spiegelung.

Sei nun $n > 1$ und $f \neq \mathrm{Id}_V$ eine orthogonale Funktion. Dann gibt es einen Vektor v in V mit $f(v) \neq v$.

Es seien $W := \mathbb{R}(f(v) - v)^{\perp}$ und $U := (\mathbb{R}v)^{\perp}$, sowie s_W die Spiegelung mit $Fix(s_W) = W$.

Wegen $f(v) = \frac{1}{2}(f(v) - v) + \frac{1}{2}(f(v) + v) \in W^{\perp} \oplus W$ ist $p_W(f(v)) = \frac{1}{2}(f(v) + v)$ und $s_W(f(v)) = 2p_W(f(v)) - f(v) = v$.

Daher sind $\mathbb{R}v$ und U stabil unter der orthogonalen Funktion $s_W \circ f$.

Nach Induktionsannahme gibt es höchstens $n-1$ Spiegelungen $s_1, \ldots, s_k$ in U so, dass $(s_W \circ f)|_U = s_1 \circ \ldots \circ s_k$.

Wir setzen für $1 \leq i \leq k$ die Spiegelungen s_i durch $s_i(v) := v$ zu Spiegelungen in V fort und bezeichnen diese wieder mit s_i. Dann ist $s_W \circ f = s_1 \circ \ldots \circ s_k$ und $f = s_W \circ s_1 \circ \ldots \circ s_k$.

Definition 82 : Es seien s eine Spiegelung in V und $0 \neq u$ ein Element des parallelen Untervektorraums W von $Fix(s)$.
Dann heißt $t_u \circ s$ *Gleitspiegelung* (um die Hyperebene $Fix(s)$).

Satz 83 : *Es sei f eine Isometrie, deren linearer Anteil g eine Spiegelung ist. Dann gibt es eindeutig bestimmte Vektoren $u \in Fix(g)$ und $v \in Fix(g)^{\perp}$ so, dass $u + v = f(0)$ ist. Es gilt:*

(1) *Wenn $u = 0$ ist, dann ist f eine Spiegelung um $\frac{1}{2}v + Fix(g)$.*
(2) *Wenn $u \neq 0$ ist, dann ist $f = t_u \circ (t_v \circ g)$ eine Gleitspiegelung um die Hyperebene $\frac{1}{2}v + Fix(g)$ und $Fix(f)$ ist leer.*

Beweis:

(1) Es sei $x \in Fix(g)$. Wegen $g(v) = -v$ ist
$$(t_v \circ g)\left(\frac{1}{2}v + x\right) = v + \frac{1}{2}g(v) + x = \frac{1}{2}v + x.$$
Daher ist die Hyperebene $\frac{1}{2}v + Fix(g)$ in $Fix(t_v \circ g)$ enthalten.
Wenn $v = 0$ ist, dann ist $t_v \circ g = g \neq \mathrm{Id}_V$.
Wenn $v \neq 0$ ist, dann ist $(t_v \circ g)(v) = v - v = 0$ und daher $t_v \circ g \neq \mathrm{Id}_V$.
Daher ist $Fix(t_v \circ g) \neq V$ und $Fix(t_v \circ g) = \frac{1}{2}v + Fix(g)$.
(2) Sei $u \neq 0$. Nach (1) ist nur noch zu zeigen, dass $Fix(f)$ leer ist. Sei $s := t_v \circ g$. Wäre $x \in Fix(f)$, dann wäre $f(f(x)) = f(x) = x$. Daher ist $x = (t_u \circ s)((t_u \circ s)(x)) = u + s(u + s(x)) =$
$= u + v + g(u + v + g(x)) = u + v + u - v + x = 2u + x$, also $u = 0$, Widerspruch zu $u \neq 0$.

§4. Isometrien der Ebene

In diesem Abschnitt sei V ein zweidimensionaler orientierter euklidischer Raum.

Wir werden die folgenden Eigenschaften der Funktionen Sinus und Cosinus verwenden:

Hilfssatz 84 :

(1) *Für alle $\alpha \in \mathbb{R}$ ist*

$$\sin(\alpha+2\pi) = \sin(\alpha), \quad \cos(\alpha+2\pi) = \cos(\alpha),$$
$$\sin(-\alpha) = -\sin(\alpha), \quad \cos(-\alpha) = \cos(\alpha)\ .$$

„Sinus und Cosinus sind 2π-periodische Funktionen, Sinus ist eine ungerade und Cosinus eine gerade Funktion“.

(2) *Für alle $\alpha, \beta \in \mathbb{R}$ ist*

$$\sin(\alpha+\beta) = \sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta),$$
$$\cos(\alpha+\beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)\ .$$

(3)

$$\sin(0) = 0\ , \qquad \cos(0) = 1\ , \qquad \sin(\frac{\pi}{2}) = 1\ ,$$
$$\cos(\frac{\pi}{2}) = 0\ , \qquad \sin(\pi) = 0\ , \qquad \cos(\pi) = -1\ .$$

Beweis: Siehe Analysis 1.

Satz 85 : *Es seien $\underline{u} = (u_1, u_2)$ eine positiv orientierte ON-Basis von V und*

$$(-)^{\perp} : V \longrightarrow V \quad , \quad v \longmapsto v^{\perp}\ ,$$

die durch $u_1^{\perp} := u_2$ und $u_2^{\perp} := -u_1$ definierte lineare Funktion.

(1) *Die Funktion $(-)^{\perp}$ ist orthogonal.*

(2) *Für alle Vektoren $v \in V$ mit $\|v\| = 1$ ist $v^{\perp}$ der einzige Vektor mit der Eigenschaft, dass $(v, v^{\perp})$ eine positiv orientierte ON-Basis von V ist.*

(3) *Für alle Vektoren $v, w \in V$ mit $\|v\| = 1 = \|w\|$ gibt es genau eine Zahl $\alpha \in [0, 2\pi[$ so, dass*

$$w = \cos(\alpha)v + \sin(\alpha)v^{\perp}$$

ist. Wenn $\alpha \in [0, \pi]$ bzw. $\alpha \in [\pi, 2\pi[$, dann ist α bzw. $2\pi - \alpha$ der Winkel zwischen v und w.

Beweis:

(1) Die Matrix von $(-)^\perp$ bezüglich $\underline{u}$ ist die orthogonale Matrix $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

(2) Sei $v = au_1 + bu_2$. Dann ist $v^\perp = -bu_1 + au_2$, also
$$(v, v^\perp) = \underline{u} \begin{pmatrix} a & -b \\ b & a \end{pmatrix} .$$
Wegen $\|v\| = 1$ ist $a^2 + b^2 = 1$. Somit ist $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ eine orthogonale Matrix mit Determinante 1 und $(v, v^\perp)$ ist eine positiv orientierte ON-Basis.
Sei $w = cu_1 + du_2$ so, dass (v, w) eine positiv orientierte ON-Basis von V ist. Aus $\langle v, w\rangle = 0$ folgt $(c, d) \in \mathbb{R}(-b, a)$, wegen $\|w\| = 1$ ist dann $(c, d) = (-b, a)$ oder $(c, d) = (b, -a)$. Weil (v, w) positiv orientiert ist, muss $(c, d) = (-b, a)$ sein.

(3) Sei β der Winkel zwischen v und w. Dann ist
$$w = \langle v, w\rangle v + \langle v^\perp, w\rangle v^\perp = \cos(\beta)v + \langle v^\perp, w\rangle v^\perp$$
und
$$\cos^2(\beta) + \langle v^\perp, w\rangle^2 = \|w\|^2 = 1 .$$
Wenn $\langle v^\perp, w\rangle \geq 0$ ist, dann ist $\langle v^\perp, w\rangle = \sin(\beta)$ und $\alpha := \beta$.
Wenn $\langle v^\perp, w\rangle \leq 0$ ist, dann ist $\langle v^\perp, w\rangle = -\sin(\beta)$ und
$\alpha := 2\pi - \beta$.

Definition 86: Es seien $u, v, w \in V$, $v \neq 0$ und $w \neq 0$. Die eindeutig bestimmte Zahl $\alpha \in [0, 2\pi[$ mit
$$\|w\|^{-1}w = \cos(\alpha)\|v\|^{-1}v + \sin(\alpha)\|v\|^{-1}v^\perp$$
heißt *orientierter Winkel von der Halbgeraden $u + \mathbb{R}_{\geq 0}v$ nach $u + \mathbb{R}_{\geq 0}w$* oder kurz *orientierter Winkel von v nach w.*

Satz 87: *Es seien $v, w \in V$, $v \neq 0$, $w \neq 0$ und α der orientierte Winkel von v nach w. Dann ist $2\pi - \alpha$ der orientierte Winkel von w nach v.*

Beweis: Wir können annehmen, dass $\|v\| = 1 = \|w\|$ ist. Dann ist
$$w = \cos(\alpha)v + \sin(\alpha)v^\perp$$
und
$$w^\perp = \cos(\alpha)v^\perp + \sin(\alpha){v^\perp}^\perp = \cos(\alpha)v^\perp - \sin(\alpha)v .$$
Daher ist
$$v = \cos(\alpha)w - \sin(\alpha)w^\perp = \cos(2\pi - \alpha)w + \sin(2\pi - \alpha)w^\perp .$$

Definition 88: Es seien $\alpha \in [0, 2\pi[$, $z \in V$ und $(-)^\perp$ die in Satz 85 definierte Funktion. Dann heißt die Funktion

$$d_\alpha : V \longrightarrow V \quad , \quad v \longmapsto \cos(\alpha)v + \sin(\alpha)v^\perp \ ,$$

Drehung in V um 0 mit *Drehwinkel* α.
Für alle $v \in V \setminus \{0\}$ ist dann α der orientierte Winkel von v nach $d_\alpha(v)$.

Die Funktion

$$d_{z,\alpha} : V \longrightarrow V \quad , \quad x \longmapsto z + d_\alpha(x-z) \ ,$$

heißt *Drehung* (in V) um den *Drehpunkt* z mit *Drehwinkel* α.
Es ist $d_{0,\alpha} = d_\alpha$.

Satz 89:

(1) *Für alle $\alpha \in [0, 2\pi[$ ist die Funktion d_α linear. Ihre Matrix bezüglich jeder positiv orientierten ON-Basis von V ist*

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

Diese Matrix heißt Drehmatrix zum Winkel α.
Insbesondere ist d_α orthogonal und $\det(d_\alpha) = 1$.

(2) *Jede orthogonale Funktion von V nach V mit Determinante* 1 *ist eine Drehung um* 0.

(3) *Für alle $\alpha \in [0, 2\pi[$ und $z \in V$ ist*

$$d_{z,\alpha} = t_{z-d_\alpha(z)} \circ d_\alpha.$$

Jede Drehung ist eine Isometrie.

Beweis: Es sei $\underline{u} = (u_1, u_2)$ eine positiv orientierte *ON*- Basis von V.

(1) Wegen $d_\alpha = \cos(\alpha)\mathrm{Id}_V + \sin(\alpha)(-)^\perp$ ist d_α linear.
Nach Satz 85, (2), ist $u_1^\perp = u_2$ und $u_2^\perp = -u_1$, also ist

$$d_\alpha(u_1) = \cos(\alpha)u_1 + \sin(\alpha)u_2$$

und

$$d_\alpha(u_2) = \cos(\alpha)u_2 - \sin(\alpha)u_1.$$

Daher ist

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

die Matrix von d_α bezüglich $\underline{u}$.
Diese Matrix ist orthogonal und ihre Determinante ist 1.

(2) Es seien $f : V \longrightarrow V$ eine orthogonale Funktion mit $\det(f) = 1$ und α der orientierte Winkel von u_1 nach $f(u_1)$. Dann ist $f(u_1) = d_\alpha(u_1)$. Weil $(f(u_1), f(u_2))$ eine positiv orientierte ON-Basis von V ist, muss

$f(u_2) = f(u_1)^{\perp}$ sein (Satz 85, (2)). Also ist auch $f(u_2) = d_\alpha(u_2)$, somit $f = d_\alpha$.

(3) Nach (1) ist d_α linear und orthogonal, daher ist für alle $x \in V$

$$d_{z,\alpha}(x) = z + d_\alpha(x) - d_\alpha(z) = t_{z-d_\alpha(z)}(d_\alpha(x))$$

und $d_{z,\alpha} = t_{z-d_\alpha(z)} \circ d_\alpha$ eine Isometrie.

Definition 90: Es seien W ein endlich-dimensionaler reeller Vektorraum und $f : W \longrightarrow W$ eine affine Funktion. Dann heißt f *orientierungserhaltend* bzw. *orientierungsumkehrend*, wenn die Determinante des linearen Anteils von f eine positive bzw. negative Zahl ist.

Beispiel 91: Drehungen in V sind orientierungserhaltend, Spiegelungen in V sind orientierungsumkehrend.

Satz 92: *Es sei f eine Isometrie von V.*
Wenn $Fix(f)$ eine Gerade ist, dann ist f eine Spiegelung.
Wenn $Fix(f)$ ein Punkt ist, dann ist f eine Drehung um diesen.
Wenn $Fix(f)$ leer und f orientierungserhaltend ist, dann ist f eine Translation.
Wenn $Fix(f)$ leer und f orientierungsumkehrend ist, dann ist f eine Gleitspiegelung.

Beweis: Sei $u := f(0)$ und g der lineare Anteil von f. Nach Hilfssatz 76 ist $Fix(f)$ entweder leer oder ein affiner Unterraum mit parallelem Vektorraum $Fix(g)$.
Nach Satz 89, (2), und nach Satz 81 ist g eine Drehung, wenn g orientierungserhaltend ist, und eine Spiegelung, wenn g orientierungsumkehrend ist.

Wenn $Fix(f)$ eine Gerade ist, dann ist f wegen $dim_{\mathbb{R}}(V) = 2$ eine Spiegelung.

Wenn $Fix(f) = \{z\}$ ein Punkt ist, ist auch $Fix(g)$ ein Punkt. Dann ist g eine Drehung um 0 und nicht Id_V. Weiters ist $z = f(z) = u + g(z)$. Daher ist für alle $x \in V$

$$f(x) = u + g(x) = z - g(z) + g(x) = z + g(x - z),$$

somit ist f eine Drehung um z mit demselben Drehwinkel wie g.

Wenn $Fix(f)$ leer ist, liegt u nicht im Bild von $\mathrm{Id}_V - g$. Daher ist

$$k := \dim_{\mathbb{R}}(\mathrm{Kern}(Id_V - g)) = 2 - \dim_{\mathbb{R}}(\mathrm{Bild}(Id_V - g)) \geq 1.$$

Wenn $k = 2$ ist, dann ist $g = \mathrm{Id}_V$ und f eine Translation.
Wenn $k = 1$ ist, dann ist g eine Spiegelung und f eine Gleitspiegelung.

Satz 93 : *Für $\alpha,\beta \in [0,2\pi[$ mit $\alpha+\beta \geq 2\pi$ sei $d_{\alpha+\beta} := d_{\alpha+\beta-2\pi}$.
Dann gilt für alle $\alpha,\beta \in [0,2\pi[$:*

$$d_\alpha \circ d_\beta = d_{\alpha+\beta}.$$

Insbesondere ist $d_\alpha^{-1} = d_{2\pi-\alpha}$ und $d_\alpha \circ d_\beta = d_\beta \circ d_\alpha$.

Beweis: Die Matrizen von d_α, d_β bezüglich einer positiv orientierten *ON*-Basis von V sind

$$\begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \text{ und } \begin{pmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{pmatrix}.$$

Daher ist

$$\begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \cdot \begin{pmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{pmatrix} =$$
$$= \begin{pmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta) \\ \sin(\alpha+\beta) & \cos(\alpha+\beta) \end{pmatrix}$$

die Matrix von $d_\alpha \circ d_\beta$.

Beispiel 94 : Es seien v, $w \in V\backslash\{0\}$, α der orientierte Winkel von v nach w und s_v, s_w die Spiegelungen mit $Fix(s_v) = \mathbb{R}v$, $Fix(s_w) = \mathbb{R}w$. O.E.d.A. nehmen wir $\|v\| = \|w\| = 1$ an. Wir berechnen $s_v \circ s_w$. Da $s_v \circ s_w$ linear ist und $\det(s_v \circ s_w) = \det(s_v)\det(s_w) = (-1)^2 = 1$, ist $s_v \circ s_w$ eine Drehung d_β um 0.
Sei $v^\perp \in V$ so, dass $(v, v^\perp)$ eine positiv orientierte ON-Basis ist. Dann ist

$$d_\beta(d_\alpha(v)) = d_\beta(w) = (s_v \circ s_w)(w) = s_v(w) =$$
$$= 2\langle v,w\rangle v - w = 2\cos(\alpha)v - \left(\cos(\alpha)v + \sin(\alpha)v^\perp\right) =$$
$$= \cos(2\pi-\alpha)v + \sin(2\pi-\alpha)v^\perp = d_{2\pi-\alpha}(v) = d_\alpha^{-1}(v).$$

Daher ist $d_\beta \circ d_\alpha = d_\alpha^{-1}$, $s_v \circ s_w = d_\beta = d_{2\alpha}^{-1}$ und $s_w \circ s_v = (s_v \circ s_w)^{-1} = d_{2\alpha}$.

Beispiel 95 : Wir betrachten $\mathbb{C}$ als 2-dimensionalen, durch die Basis $(1,i)$ orientierten euklidischen Raum.
Für $y,z \in \mathbb{C}\backslash\{0\}$ seien α,β die orientierten Winkel von 1 nach y,z. Wir schreiben $e^{i\alpha}$ für $d_\alpha(1) = \cos(\alpha) + i\sin(\alpha)$.
Dann ist $y = |y|\, e^{i\alpha}$, $z = |z|\, e^{i\beta}$ („Polardarstellung von komplexen Zahlen“) und $y \cdot z = |y| \cdot |z| \cdot e^{i(\alpha+\beta)}$.
(„Bei der Multiplikation komplexer Zahlen werden die Beträge multipliziert und die Winkel addiert.“)

Beispiel 96 : Die Punkte $u,v,w \in V$ seien die Eckpunkte eines Dreiecks. Die orientierten Winkel von $v-u$ bzw. $w-v$ bzw. $u-w$ nach $w-u$ bzw. $u-$

v bzw. $v-w$ seien α bzw. β bzw. γ. Wir können annehmen, dass $\alpha,\beta,\gamma \in [0,\pi[$ ist .
Dann ist

$$d_{\alpha+\beta+\gamma}(\|u-w\|^{-1}(u-w)) = (d_\alpha \circ d_\beta \circ d_\gamma)(\|u-w\|^{-1}(u-w)) =$$

$$= d_\alpha(d_\beta(\|v-w\|^{-1}(v-w)) = -d_\alpha(d_\beta(\|w-v\|^{-1}(w-v)) =$$

$$= -d_\alpha(\|u-v\|^{-1}(u-v)) = d_\alpha(\|v-u\|^{-1}(v-u)) =$$

$$= -\|u-w\|^{-1}(u-w) = d_\pi(\|u-w\|^{-1}(u-w))$$

und $\alpha+\beta+\gamma = \pi$.
(„Die Winkelsumme im Dreieck ist π.“)

Beispiel 97 : Wir betrachten in der Ebene, die wir nach Wahl eines Nullpunktes und einer Basis mit $\mathbb{R}^2$ identifizieren, zwei drehbare Scheiben. Die erste ist um den Punkt $(0,0)$ drehbar, die zweite ist auf der ersten im Punkt (a,b) montiert. Die zweite Scheibe wird um den Winkel β, die erste um den Winkel α gedreht.
Dann wird der Punkt (x,y) auf der zweiten Scheibe in den Punkt

$$d_\alpha(d_{(a,b),\beta}(x,y)) = d_\alpha(a,b) + d_{\alpha+\beta}(x-a,y-b) =$$

$$= (a\cos(\alpha) - b\sin(\alpha), a\sin(\alpha) + b\cos(\alpha)) + ((x-a)\cos(\alpha+\beta) -$$

$$-(y-b)\sin(\alpha+\beta), (x-a)\sin(\alpha+\beta) + (y-b)\cos(\alpha+\beta))$$

bewegt.

§5. Zeigerrechnung

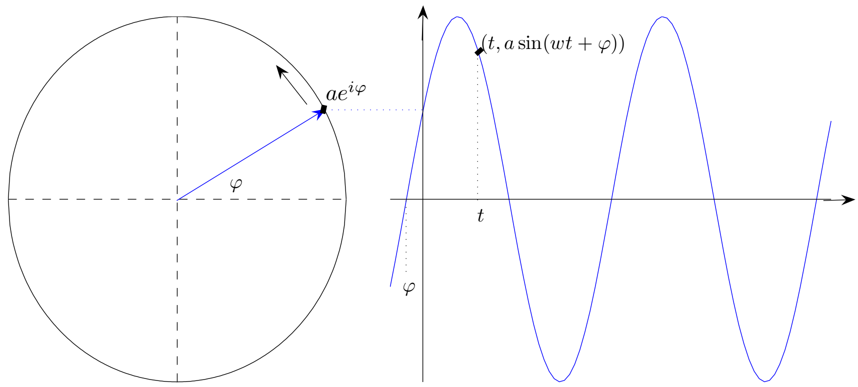
Ein Zeiger (zum Beispiel der Sekundenzeiger einer Uhr) der Länge a dreht sich in einer Ebene mit der Kreisfrequenz ω um einen Punkt $\mathcal{O}$. Wir identifizieren diese Ebene mit $\mathbb{C} = \mathbb{R}^2$ und $\mathcal{O}$ mit $0_\mathbb{C} = (0,0)$. Zur Zeit 0 sei die Spitze des Zeigers im Punkt $ae^{i\varphi} = (a\cos(\varphi), a\sin(\varphi))$. Zur Zeit t befindet sie sich dann im Punkt

$$ae^{i(\omega t+\varphi)} = (a\cos(\omega t+\varphi), a\sin(\omega t+\varphi)).$$

Die Funktion, die der Zeit t den Abstand der Zeigerspitze von der reellen Geraden $\mathbb{R}(1,0)$ zuordnet, ist daher

$$\begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ t & \longmapsto & a\sin(\omega t+\varphi) \end{array} .$$

Diese Funktion beschreibt eine *harmonische Schwingung* mit *Kreisfrequenz* ω, *Amplitude* a und *Anfangsphase* φ.



Drehen sich zwei Zeiger mit derselben Kreisfrequenz ω um denselben Punkt, aber mit eventuell verschiedenen Längen a_1, a_2 und Anfangsphasen φ_1, φ_2, dann heißt $\varphi_1 - \varphi_2$ die *Phasendifferenz* der entsprechenden harmonischen Schwingung. Der Abstand der Summe

$$ae^{i(\omega t+\varphi)} := a_1 e^{i(\omega t+\varphi_1)} + a_2 e^{i(\omega t+\varphi_2)}$$

der Zeigerspitzen von der reellen Geraden $\mathbb{R}(1,0)$ ist die Summe der Abstände der zwei Zeigerspitzen von dieser Geraden, also

$$a\sin(\omega t+\varphi) = a_1\sin(\omega t+\varphi_1) + a_2\sin(\omega t+\varphi_2)\,.$$

Die Summe von zwei harmonischen Schwingungen mit derselben Kreisfrequenz ist daher wieder eine harmonische Schwingung, ihre Amplitude bzw. Anfangsphase ist a bzw. φ. Um diese zu berechnen, muss die *Polardarstellung* $ae^{i(\omega t+\varphi)}$ der Summe der zwei komplexen Zahlen, die den zwei Zeigerspitzen entsprechen, berechnet werden.

Diese Überlegungen führen zur *Zeigerrechnung*, mit der die folgende Aufgabe gelöst wird:
Es seien n eine positive ganze Zahl und $\omega, \varphi_1, \ldots, \varphi_n, a_1, \ldots, a_n$ reelle Zahlen. Gesucht sind reelle Zahlen a und φ so, dass für alle Zahlen $t \in \mathbb{R}$

$$\sum_{\ell=1}^{n} a_\ell \sin(\omega t+\varphi_\ell) = a\sin(\omega t+\varphi)$$

ist.

Sei $t \in \mathbb{R}$. Wir betrachten die komplexen Zahlen

$$z_\ell(t) := a_\ell e^{i(\omega t+\varphi_\ell)} := a_\ell(\cos(\omega t+\varphi_\ell) + i\sin(\omega t+\varphi_\ell)),\ 1 \le \ell \le n.$$

Dann ist $a_\ell \sin(\omega t+\varphi_\ell)$ der Imaginärteil von $z_\ell(t)$ und $\sum_{\ell=1}^{n} a_\ell \sin(\omega t+\varphi_\ell)$ der Imaginärteil von

$$z(t) := \sum_{\ell=1}^{n} z_\ell(t).$$

Es ist

$$z(t) = \sum_{\ell=1}^{n} z_\ell(t) = \sum_{\ell=1}^{n} a_\ell e^{i(\omega t + \varphi_\ell)} = e^{i\omega t} \sum_{\ell=1}^{n} a_\ell e^{i\varphi_\ell} = e^{i\omega t} z(0)\,.$$

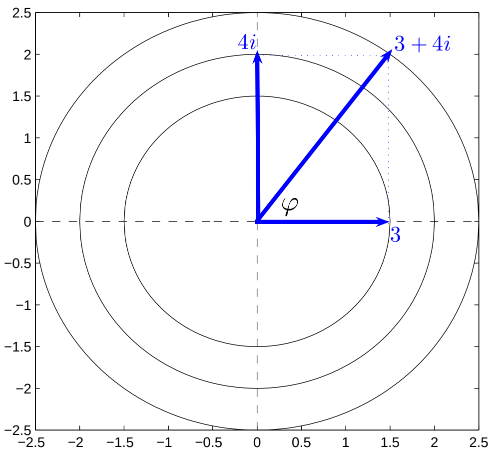
Wir erhalten daher die gesuchten Zahlen a und φ aus der Polardarstellung $z(0) = a(\cos(\varphi) + i\sin(\varphi)) = ae^{i\varphi}$ von $z(0)$, dabei ist a der Abstand zwischen $z(0)$ und 0 und φ der orientierte Winkel von $\mathbb{R}_{\geq 0} \cdot 1$ nach $\mathbb{R}_{\geq 0} \cdot z(0)$.

Beispiel 98 : Berechne reelle Zahlen a und φ so, dass für alle Zahlen $t \in \mathbb{R}$

$$3 \cdot \sin(\omega t) + 4 \cdot \cos(\omega t) = a \cdot \sin(\omega t + \varphi)$$

ist.
Wegen $\cos(\omega t) = \sin(\omega t + \frac{\pi}{2})$ ist $a_1 = 3$, $a_2 = 4$, $\varphi_1 = 0$, $\varphi_2 = \frac{\pi}{2}$. Weiters sind
$z_1(t) = 3(\cos(\omega t) + i\sin(\omega t))$, $z_2(t) = 4(\cos(\omega t + \frac{\pi}{2}) + i\sin(\omega t + \frac{\pi}{2}))$ und $z(t) = z_1(t) + z_2(t)$, also $z(0) = 3 + 4i$.
Der Abstand zwischen $z(0)$ und 0 ist $a = \sqrt{3^2 + 4^2} = 5$, der Tangens des orientierten Winkels von $\mathbb{R}_{\geq 0} \cdot 1$ nach $\mathbb{R}_{\geq 0} \cdot (3 + 4i)$ ist $\frac{4}{3}$. Daher ist für alle $t \in \mathbb{R}$

$$3 \cdot \sin(\omega t) + 4 \cdot \cos(\omega t) = 5 \cdot \sin(\omega t + \arctan(\frac{4}{3}))\,.$$



Beispiel 99 : Berechne reelle Zahlen a und φ so, dass für alle Zahlen $t \in \mathbb{R}$

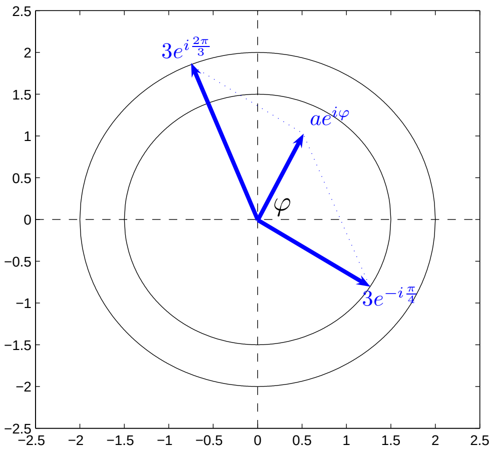
$$3 \cdot \sin(\omega t - \frac{\pi}{4}) + 4 \cdot \sin(\omega t + \frac{2\pi}{3}) = a \cdot \sin(\omega t + \varphi)$$

ist.
Es ist $z_1(0) = 3(\cos(-\frac{\pi}{4}) + i\sin(-\frac{\pi}{4}))$, $z_2(0) = 4(\cos(\frac{2\pi}{3}) + i\sin(\frac{2\pi}{3}))$ und

$z(0) = z_1(0) + z_2(0)$, also

$$z(0) = 3\cos(-\frac{\pi}{4}) + 4\cos(\frac{2\pi}{3}) + i(3\sin(-\frac{\pi}{4}) + 4\sin(\frac{2\pi}{3}))\,.$$

Somit ist $z(0) = \frac{3\sqrt{2}}{2} - 2 - i(\frac{3\sqrt{2}}{2} + 2\sqrt{3})$ und $a \approx 1.34825$, $\varphi \approx 1.48069$.



§6. Isometrien des Raumes

In diesem Abschnitt sei V ein dreidimensionaler orientierter euklidischer Raum.

Definition 100 : Eine orientierungserhaltende Isometrie von V heißt *(eigentliche) Bewegung* in V.

Definition 101 : Es seien E ein zweidimensionaler orientierter Untervektorraum von V, $z \in E$ und d eine Drehung um z in E. Für alle $v \in V$ schreiben wir v_1 bzw. v_2 für die orthogonale Projektion von v auf E bzw. $E^{\perp}$. Die Funktion

$$f : V \to V \quad , \quad v = v_1 + v_2 \longmapsto d\,(v_1) + v_2\,,$$

heißt *Drehung (in V)* um die *Drehachse* $z + E^{\perp}$ und mit *Drehebene* E.
Sei $0 \neq u \in E^{\perp}$ und $f \neq \mathrm{Id}_V$. Dann heißt die Funktion $t_u \circ f$ *Schraubung* um die Drehachse $z + E^{\perp}$.
Sei s eine Spiegelung in V, deren Fixmenge zu E parallel ist. Dann heißt die Funktion $s \circ f$ *Drehspiegelung* um die Drehachse $z + E^{\perp}$ und mit Spiegelungsebene $Fix\,(s)$.
Der Drehwinkel von d wird auch *Drehwinkel von f bzw. $t_u \circ f$ bzw. $s \circ f$* genannt.

Satz 102 :

(1) *Drehungen, Schraubungen und Drehspiegelungen sind Isometrien.*
(2) *Die Fixmenge einer Drehung* $\neq \mathrm{Id}_V$ *ist ihre Drehachse. Die Fixmenge einer Schraubung ist leer.*
(3) *Sei* g *eine Drehung mit* $g(0) = 0$. *Ist* $\underline{v} = (v_1, v_2, v_3)$ *eine ON-Basis von* V *so, dass* (v_1, v_2) *eine ON-Basis der Drehebene* E *von* g *ist, dann ist*

$$\begin{pmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

die Matrix von g *bezüglich* $\underline{v}$. *Dabei ist* α *der Drehwinkel von* $g|_E$ *bezüglich der durch* (v_1, v_2) *gegebenen Orientierung von* E. *Insbesondere ist* $\cos(\alpha) = \frac{1}{2}(\mathrm{spur}(g) - 1)$.
(4) *Drehungen und Schraubungen sind orientierungserhaltend, Drehspiegelungen sind orientierungsumkehrend.*

Beweis:

(1) Es seien f eine Drehung mit Drehebene E und $d := f|_E$. Für $v, w \in V$ ist $d(v_1) - d(w_1) \in E$ und $v_2 - w_2 \in E^\perp$. Daher ist

$$\begin{aligned} \|f(v) - f(w)\|^2 &= \|d(v_1) - d(w_1) + v_2 - w_2\|^2 \\ &= \|d(v_1) - d(w_1)\|^2 + \|v_2 - w_2\|^2 \\ &= \|v_1 - w_1\|^2 + \|v_2 - w_2\|^2 \\ &= \|v - w\|^2 . \end{aligned}$$

Somit ist f eine Isometrie.
(2) nachprüfen.
(3) Wegen (1) und $g(0) = 0$ folgt aus Satz 63 , dass g linear ist. Die Behauptung folgt daher aus Satz 89 .
(4) Folgt aus (3).

Satz 103 :

(1) *Jede Bewegung in* V *ist eine Drehung, eine Schraubung oder eine Translation.*
(2) *Jede orientierungsumkehrende Isometrie in* V *ist eine Spiegelung, Gleitspiegelung oder Drehspiegelung.*

Beweis: Es sei $g : V \to V$ eine orthogonale Funktion. Nach Satz 73 gibt es einen 1-dimensionalen, unter g stabilen Untervektorraum W von V. Aus Satz 67 folgt, dass $g|_W = \mathrm{Id}_W$ oder $g|_W = -\mathrm{Id}_W$.

Nach Satz 69 ist auch $W^{\perp}$ stabil unter g. Nach Satz 92 ist die Einschränkung von g auf $W^{\perp}$ eine Drehung oder eine Spiegelung. Es ist
$\det(g) = \det(g\,|_W) \cdot \det(g\,|_{W^{\perp}})$.

(1) Sei $\det(g) = 1$. Dann ist entweder
$g\,|_W = \mathrm{Id}_V$ und $g\,|_{W^{\perp}}$ eine Drehung oder
$g\,|_W = -\mathrm{Id}_V$ und $g\,|_{W^{\perp}}$eine Spiegelung.
Im ersten Fall ist g eine Drehung um die Drehachse W, im zweiten die Drehung um die Drehachse $Fix(g\,|_{W^{\perp}})$ mit
Drehwinkel π.
Sei $u \in Fix(g)$ und $v \in Fix(g)^{\perp}$. Die Bewegung $t_{u+v} \circ g$ ist
eine Translation, wenn $g = \mathrm{Id}_{V,}$
eine Drehung, wenn $u = 0$,
und eine Schraubung, wenn $u \neq 0$ und $g \neq \mathrm{Id}_V$.

(2) Sei $\det(g) = -1$. Dann ist entweder
$g\,|_W = \mathrm{Id}_V$ und $g\,|_{W^{\perp}}$eine Spiegelung oder
$g\,|_W = -\mathrm{Id}_V$ und $g\,|_{W^{\perp}}$ eine Drehung.
Im ersten Fall ist g die Spiegelung um $W \oplus Fix(g\,|_{W^{\perp}})$, im zweiten Fall ist g eine Drehspiegelung mit Drehachse W. Das Zusammensetzen von g mit einer Translation ergibt nach Satz 83 eine Spiegelung, Gleitspiegelung oder Drehspiegelung.

Beispiel 104 : Die Funktion

$$f : \mathbb{R}^3 \to \mathbb{R}^3 \quad , \quad (a,b,c) \longmapsto (c,a,b) \ ,$$

ist eine Isometrie, wegen $f(0) = 0$ sogar orthogonal.
Es ist $\det(f) = 1$, $\mathrm{spur}(f) = 0$ und der Eigenraum von f zum Eigenwert 1 ist $\mathbb{R}(1,1,1)$.
Daher ist f eine Drehung um die Drehachse $\mathbb{R}(1,1,1)$ mit Drehwinkel α, wobei $\cos(\alpha) = -\frac{1}{2}$.

Beispiel 105 : Die Funktion

$$f : \mathbb{R}^3 \to \mathbb{R}^3 \quad , \quad (a,b,c) \longmapsto (-b,a,-c) \ ,$$

ist eine Isometrie, wegen $f(0) = 0$ sogar orthogonal.
Es ist $\det(f) = -1$, $\mathrm{spur}(f) = -1$ und der Eigenraum von f zum Eigenwert -1 ist $\mathbb{R}(0,0,1)$.
Daher ist f eine Drehspiegelung um die Drehachse $\mathbb{R}(0,0,1)$
mit Drehwinkel α, wobei $\cos(\alpha) = 0$.

Satz 106 : *Es seien $\underline{v} = (v_1, v_2, v_3)$ eine ON-Basis von V und f eine Bewegung in V mit $f(0) = 0$. Dann gibt es Drehungen g_α, g_γ mit Drehachse $\mathbb{R}v_1$ und Drehwinkeln $\alpha, \gamma \in [0, 2\pi[$ und eine Drehung h_β mit Drehachse $\mathbb{R}v_3$*

und Drehwinkel $\beta \in [0,\pi]$ *so, dass*

$$f = g_\alpha \circ h_\beta \circ g_\gamma$$

ist. Die Zahlen α, β, γ *heißen* Euler-Winkel *von* f.
Für Matrizen bedeutet dieser Satz: Jede orthogonale Matrix $A \in \mathbb{R}^{3\times 3}$ *mit* $\det(A) = 1$ *kann als Produkt*

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{pmatrix} \cdot \begin{pmatrix} \cos(\beta) & -\sin(\beta) & 0 \\ \sin(\beta) & \cos(\beta) & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot$$
$$\cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\gamma) & -\sin(\gamma) \\ 0 & \sin(\gamma) & \cos(\gamma) \end{pmatrix}$$

geschrieben werden, wobei $\alpha, \gamma \in [0, 2\pi[$ *und* $\beta \in [0,\pi]$ *sind.*

Beweis: Sei $w := f^{-1}(v_1) \in V$. Sei g_γ eine Drehung um die Drehachse $\mathbb{R}v_1$ so, dass $g_\gamma(w) \in (\mathbb{R}v_3)^\perp$. Dann gibt es eine Drehung h_β mit Drehachse $\mathbb{R}v_3$ und Drehwinkel $\beta \in [0,\pi]$ so, dass $h_\beta\left(g_\gamma(w)\right) = v_1$. Die Funktion $h_\beta \circ g_\gamma \circ f^{-1}$ ist eine Drehung, wegen $\left(h_\beta \circ g_\gamma \circ f^{-1}\right)(v_1) = v_1$ ist ihre Drehachse $\mathbb{R}v_1$. Setze $g_\alpha := \left(h_\beta \circ g_\gamma \circ f^{-1}\right)^{-1}$.

§7. Symmetriegruppen

Definition 107: Es seien $M_1, \ldots, M_k$ Teilmengen von V. Dann heißt die Menge $Sym_V(M_1, \ldots, M_k) :=$

$$= \{f : V \to V \mid f \text{ Isometrie}, f(M_i) = M_i, i = 1, \ldots, k\}$$

Symmetriegruppe von $M_1, \ldots, M_k$. Ihre Elemente heißen *Symmetrieoperationen* von $M_1, \ldots, M_k$.

Satz 108:

(1) *Die Symmetriegruppe von* $M_1, \ldots, M_k$ *ist mit der Zusammensetzung von Funktionen eine Gruppe.*
(2) *Ist eine der Mengen* $M_1, \ldots, M_k$ *endlich, dann ist ihr Schwerpunkt ein Fixpunkt für alle Symmetrieoperationen von* $M_1, \ldots, M_k$.

Beweis:

(1) kann leicht nachgeprüft werden.
(2) Sei M_1 endlich. Nach Satz 34 ist für jede Symmetrieoperation f das Bild des Schwerpunktes von M_1 der Schwerpunkt des Bildes $f(M_1) = M_1$.

Satz 109 : *Es seien $M_1,...,M_k$ Teilmengen von V. M_1 sei eine endliche, nicht-leere Menge, deren affine Hülle eine Hyperebene in V oder gleich V ist. Dann ist $Sym_V(M_1,...,M_k)$ eine endliche Menge.*
Diese kann wie folgt berechnet werden:
Sei z der Schwerpunkt von M_1 und W die affine Hülle von M_1.

Fall *1: $z = 0$. Dann ist W ein Untervektorraum von V.*
Wähle eine Basis $(v_1,...,v_n)$ von W aus Elementen von M_1. Wenn $V \neq W$, ergänze diese durch ein Element $v_{n+1} \in W^{\perp}$ zu einer Basis von V.
$Sym_V(M_1,...,M_k)$ ist dann in der endlichen Menge
$\{g \in \mathrm{GL}_{\mathbb{R}}(V) \mid g(v_i) \in M_1, 1 \leq i \leq n, g(v_{n+1}) \in \{v_{n+1}, -v_{n+1}\}\}$
enthalten.
Wähle aus dieser die Elemente von $Sym_V(M_1,...,M_k)$ aus.

Fall *2: $z \neq 0$.*
Dann ist 0 der Schwerpunkt von $t_{-z}(M_1)$.
Berechne $Sym_V(t_{-z}(M_1),...,t_{-z}(M_k))$ wie in Fall 1. Dann ist
$Sym_V(M_1,...,M_k) = t_z \circ Sym_V(t_{-z}(M_1),...,t_{-z}(M_k)) \circ t_{-z}$.

Beweis: Übung.

Beispiel 110 : Wir stellen das Wassermolekül H_2O durch zwei Teilmengen $M_1 = \{O\}, M_2 = \{H, H'\}$ eines 3-dimensionalen euklidischen Raumes V dar. Der Winkel zwischen H und H' liegt in $]0, \pi[$.
Die Symmetriegruppe dieses Moleküls ist $Sym_V(M_1, M_2)$.
Es seien T die affine Hülle von O, H, H' und G die Gerade durch O und $\frac{1}{2}(H + H')$. Dann ist $Sym_V(M_1, M_2) =$

$$\left\{ \begin{array}{c} \mathrm{Id}_V,\ \text{Drehung um die Drehachse } G \text{ mit Drehwinkel } \pi, \\ \text{Spiegelung um } T,\ \text{Spiegelung um die Ebene } O + (H - H')^{\perp} \end{array} \right\}.$$

KAPITEL 4

Multilineare Funktionen

Es seien K ein Körper mit $1_K + 1_K \neq 0_K$ und V ein n-dimensionaler Vektorraum über K.

§1. Multilineare Funktionen

Definition 111 : Seien $V_1, \ldots, V_\ell$ und W Vektorräume über K. Eine Funktion

$$f : V_1 \times \cdots \times V_\ell \to W \,, \, (x_1, \ldots, x_\ell) \mapsto f(x_1, \ldots, x_\ell) \,,$$

heißt *K-multilinear*, wenn f in jeder Komponente K-linear ist, d.h. für alle $(x_1, \ldots, x_\ell) \in V_1 \times \cdots \times V_\ell$, für alle $k \in \{1, \ldots, \ell\}$, für alle $y \in V_k$ und für alle $c \in K$ gilt

$$\begin{aligned} &f(x_1, \ldots, x_{k-1}, x_k + y, x_{k+1}, \ldots, x_\ell) = \\ &\qquad f(x_1, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_\ell) + f(x_1, \ldots, x_{k-1}, y, x_{k+1}, \ldots, x_\ell) \end{aligned}$$

und

$$\begin{aligned} &f(x_1, \ldots, x_{k-1}, c \cdot x_k, x_{k+1}, \ldots, x_\ell) = \\ &\qquad\qquad c \cdot f(x_1, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_\ell) \,. \end{aligned}$$

Sei

$$\mathrm{Mult}_K(V_1 \times \cdots \times V_\ell, W) := \{f : V_1 \times \cdots \times V_\ell \to W \mid f \text{ } K\text{-multilinear}\}.$$

Für $\ell = 1$ ist „multilinear" gleich „linear", für $\ell = 2, 3$ sagt man anstelle von „multilinear" auch „bilinear" bzw. „trilinear".

Beispiel 112 : Die Funktion

$$K^\ell \to K \,, \, (x_1, \ldots, x_\ell) \mapsto \prod_{i=1}^{\ell} x_i \,,$$

ist für $\ell \geq 2$ nicht K-linear, weil im Allgemeinen

$$\prod_{i=1}^{\ell} (a_i + b_i) \neq \left(\prod_{i=1}^{\ell} a_i \right) + \left(\prod_{i=1}^{\ell} b_i \right)$$

ist, aber K-multilinear, weil für alle $(x_1,\dots,x_\ell)\in K^\ell$, $k\in\{1,\dots,\ell\}$, $y\in K$ und $c\in K$ sowohl

$$x_1\dots x_{k-1}(x_k+y)x_{k+1}\dots x_\ell = x_1\dots x_{k-1}x_kx_{k+1}\dots x_\ell + x_1\dots x_{k-1}yx_{k+1}\dots x_\ell$$

als auch

$$x_1\dots x_{k-1}(cx_k)x_{k+1}\dots x_\ell = c(x_1\dots x_{k-1}x_kx_{k+1}\dots x_\ell)$$

gilt.

Satz 113 : $\mathrm{Mult}_K(V_1\times\dots\times V_\ell,W)$ *ist mit der punktweisen Addition und Skalarmultiplikation ein Untervektorraum von* $\mathcal{F}(V_1\times\dots\times V_\ell,W)$.

Beweis: Übung.

Satz 114 : *Die Funktion*

$$D:\left(K^{n\times 1}\right)^n\to K\,,\ (x_1,\dots,x_n)\mapsto \det(x_1|\dots|x_n)\,,$$

ist K-multilinear, d.h. die Determinante ist multilinear in den Spalten.
Wegen $\det(A)=\det(A^T)$ *ist die Determinante daher auch multilinear in den Zeilen.*
Insbesondere gilt für $A\in K^{n\times n}$ *und* $c\in K$:

$$\det(cA)=c^n\det(A)\,.$$

Beweis: Wegen

$$D(x_1,\dots,x_n)=\sum_{\sigma\in S_n}\mathrm{sign}(\sigma)(x_1)_{\sigma(1)}\dots(x_n)_{\sigma(n)}$$

genügt es nach Satz 113 zu zeigen, dass für jedes $\sigma\in S_n$ die Funktion

$$D_\sigma:(K^{n\times 1})^n\to K\,,\ (x_1,\dots,x_n)\mapsto (x_1)_{\sigma(1)}\dots(x_n)_{\sigma(n)}\,,$$

multilinear ist. Für $k\in\{1,\dots,n\}$, $y\in K^{n\times 1}$ und $c\in K$ ist
$D_\sigma(x_1,\dots,x_{k-1},x_k+y,x_{k+1},\dots,x_n)=$
$(x_1)_{\sigma(1)}\cdots(x_{k-1})_{\sigma(k-1)}((x_k)_{\sigma(k)}+y_{\sigma(k)})(x_{k+1})_{\sigma(k+1)}\cdots(x_n)_{\sigma(n)}=$
$D_\sigma(x_1,\dots,x_n)+D_\sigma(x_1,\dots,x_{k-1},y,x_{k+1},\dots,x_n)$
und
$D_\sigma(x_1,\dots,x_{k-1},(cx_k)x_{k+1},\dots,x_n)=$
$(x_1)_{\sigma(1)}\cdots(x_{k-1})_{\sigma(k-1)}(c(x_k)_{\sigma(k)})(x_{k+1})_{\sigma(k+1)}\cdots(x_n)_{\sigma(n)}=$
$cD_\sigma(x_1,\dots,\dots,x_n)$.

Satz 115 : *Seien $V_1,\dots,V_\ell$ und W Vektorräume über K und $(v_{1i_1})_{i_1\in I_1},\dots,(v_{\ell i_\ell})_{i_\ell\in I_\ell}$ Basen von $V_1,\dots,V_\ell$. Sei $(w_i)_{i\in I}\in\mathcal{F}(I,W)$, wobei*

$$I := I_1\times\cdots\times I_\ell$$

ist. Dann gibt es genau eine multilineare Funktion

$$f: V_1\times\cdots\times V_\ell\to W$$

mit

$$f(v_{1i_1},\dots,v_{\ell i_\ell}) = w_i$$

für alle $i=(i_1,\dots,i_\ell)\in I$. Somit kann eine multilineare Funktion durch Vorgabe der Bilder aller möglichen ℓ-Tupel von Basisvektoren eindeutig definiert werden.

Beweis: Wenn eine derartige Funktion f existiert, dann ist für Vektoren $x_1=\sum_{i_1\in I_1} c_{1i_1}v_{1i_1}\in V_1,\dots,x_\ell=\sum_{i_\ell\in I_\ell} c_{\ell i_\ell}v_{\ell i_\ell}\in V_\ell$

$$\begin{aligned}
f(x_1,\dots,x_\ell) &= f(\sum_{i_1\in I_1} c_{1i_1}v_{1i_1},\dots,\sum_{i_\ell\in I_\ell} c_{\ell i_\ell}v_{\ell i_\ell})\\
&= \sum_{i_1\in I_1}\cdots\sum_{i_\ell\in I_\ell} c_{1i_1}\dots c_{\ell i_\ell} f(v_{1i_1},\dots,v_{\ell i_\ell})\\
&= \sum_{i_1\in I_1}\cdots\sum_{i_\ell\in I_\ell} c_{1i_1}\dots c_{\ell i_\ell} w_{(i_1,\dots,i_\ell)},
\end{aligned}$$

was die Eindeutigkeit der Funktion beweist. Um die Existenz zu zeigen, definieren wir eine Funktion $f: V_1\times\cdots\times V_\ell\to W$ durch

$$f(x_1,\dots,x_\ell) := \sum_{i_1\in I_1}\cdots\sum_{i_\ell\in I_\ell} c_{1i_1}\cdots\ell i_\ell\, w_{(i_1,\dots,i_\ell)},$$

wobei $(c_{ji_j})_{i_j\in I_j}$ für $1\le j\le\ell$ die Koordinatenfamilie von $x_j\in V_j$ bezüglich der Basis $(v_{ji_j})_{i_j\in I_j}$ ist. Dann ist leicht nachzuprüfen, dass f multilinear und $f(v_{1i_1},\dots,v_{\ell i_\ell})=w_i$ für alle $i\in I$ ist.

Beispiel 116 : Seien $(E_{ij})_{i,j}$, $(F_{k\ell})_{k,\ell}$ und $(G_{rs})_{r,s}$ die Standardbasen von $K^{m\times n}$, $K^{n\times p}$ bzw. $K^{m\times p}$. Dann gibt es nach Satz 115 genau eine bilineare Funktion f: $K^{m\times n}\times K^{n\times p}\to K^{m\times p}$ mit

$$f(E_{ij},F_{k\ell}) = \delta_{jk}G_{i\ell}$$

Für Matrizen $A\in K^{m\times n}$ und $B\in K^{n\times p}$ ist wegen $A=\sum_{i,j}A_{ij}E_{ij}$ und $B=\sum_{k,\ell}B_{k\ell}F_{k\ell}$

$$\begin{aligned}
f(A,B) &= \sum_{i,j}\sum_{k,\ell}A_{ij}B_{k\ell}\delta_{jk}G_{i\ell} = \sum_{i,\ell}\left(\sum_j A_{ij}B_{j\ell}\right)G_{i\ell}\\
&= \sum_{i,\ell}(AB)_{i\ell}G_{i\ell} = AB,
\end{aligned}$$

also f das Matrizenprodukt.

§2. Alternierende Funktionen

Definition 117 : Seien V und W Vektorräume über K. Eine Funktion

$$f : V^\ell \to W \, , \, (x_1, \dots, x_\ell) \mapsto f(x_1, \dots, x_\ell) \, ,$$

heißt K-*alternierend*, wenn f K-multilinear ist und alle Elemente von V^ℓ mit zwei gleichen Komponenten auf 0 abbildet, d.h. für alle $(x_1, \dots, x_\ell) \in V^\ell$ mit $x_i = x_k$ für irgendwelche $i \neq k$ ist $f(x_1, \dots, x_\ell) = 0$. Sei

$$\mathrm{Alt}_K(V^\ell, W) := \{ f : V^\ell \to W \mid f \ K\text{-alternierend} \} \, .$$

Satz 118 : $\mathrm{Alt}_K(V^\ell, W)$ *ist mit der punktweisen Addition und Skalarmultiplikation ein Untervektorraum von* $\mathrm{Mult}(V^\ell, W)$.

Beweis: Übung.

Satz 119 : *Die Funktion*

$$D : \left(K^{n \times 1}\right)^n \to K \, , \, (x_1, \dots, x_n) \mapsto \det(x_1 | \dots | x_n) \, ,$$

ist K-alternierend, d.h. die Determinante ist alternierend in den Spalten. (und damit auch in den Zeilen).

Beweis: Sei $(x_1, \dots, x_\ell) \in (K^{n \times 1})^\ell$ mit $x_i = x_k$ für gewisse $i \neq k$, und sei

$$\tau := (i, k) \in S_n$$

die Vertauschung von i und k. Für $\rho \in S_n$ mit $\mathrm{sign}(\rho) = 1$ ist

$$\mathrm{sign}(\rho \tau) = \mathrm{sign}(\rho) \cdot \mathrm{sign}(\tau) = 1 \cdot (-1) = -1 \, .$$

Somit erhalten wir eine Funktion

$$\{ \rho \in S_n \mid \mathrm{sign}(\rho) = 1 \} \to \{ \sigma \in S_n \mid \mathrm{sign}(\sigma) = -1 \} \, , \, \rho \mapsto \rho \tau \, .$$

Deren Umkehrfunktion ist

$$\{ \sigma \in S_n \mid \mathrm{sign}(\sigma) = -1 \} \to \{ \rho \in S_n \mid \mathrm{sign}(\rho) = 1 \} \, , \, \sigma \mapsto \sigma \tau \, ,$$

weil $\mathrm{sign}(\sigma \tau) = 1$ und $\tau \tau = \mathrm{Id}_n$ ist. Für $\sigma \in S_n$ sei

$$D_\sigma : (K^{n \times 1})^n \to K \, , \, (y_1, \dots, y_n) \mapsto (y_1)_{\sigma(1)} \cdots (y_n)_{\sigma(n)} \, .$$

Dann ist

$$D = \sum_{\sigma \in S_n} D_\sigma = \sum_{\substack{\sigma \in S_n \\ \mathrm{sign}(\sigma) = 1}} (D_\sigma - D_{\sigma \tau}) \, .$$

Wegen $D_\sigma(x) = D_{\sigma \tau}(x)$ folgt

$$D(x) = \sum_{\substack{\sigma \in S_n \\ \mathrm{sign}(\sigma) = 1}} (D_\sigma(x) - D_{\sigma \tau}(x)) = 0 \, .$$

Satz 120: *Seien V, W Vektorräume über K, $f : V^\ell \to W$ alternierend und $(x_1, \ldots, x_\ell) \in V^\ell$. Dann ist für $\sigma \in S_n$*

$$f(x_{\sigma(1)}, \ldots, x_{\sigma(\ell)}) = \operatorname{sign}(\sigma) \cdot f(x_1, \ldots, x_\ell)\,.$$

Insbesondere ist für eine Transposition $(i,k) \in S_n$, $i < k$,

$$f(x_1, \ldots, x_{i-1}, x_k, x_{i+1}, \ldots, x_{k-1}, x_i, x_{k+1}, \ldots, x_\ell) = -f(x_1, \ldots, x_\ell)\,.$$

Beweis: Da jede Permutation Produkt von Transpositionen ist, genügt es zu zeigen, dass für Transpositionen $\tau_1, \ldots, \tau_r \in S_\ell$

$$f(x_{\tau_1 \ldots \tau_r(1)}, \ldots, x_{\tau_1 \ldots \tau_r(\ell)}) = (-1)^r \cdot f(x_1, \ldots, x_\ell)$$

ist. Dies zeigen wir durch Induktion nach r. Für $r = 0$ ist die Aussage offenbar richtig. Für $r = 1$ schreiben wir $\tau_1 = (i,k)$ mit $i < k$ und folgern aus f alternierend

$$\begin{aligned}
0 &= f(x_1, \ldots, x_{i-1}, x_i + x_k, x_{i+1}, \ldots, x_{k-1}, x_i + x_k, x_{k+1}, \ldots, x_\ell) \\
&= f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_{k-1}, x_i, x_{k+1}, \ldots, x_\ell) \\
&\quad + f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_\ell) \\
&\quad + f(x_1, \ldots, x_{i-1}, x_k, x_{i+1}, \ldots, x_{k-1}, x_i, x_{k+1}, \ldots, x_\ell) \\
&\quad + f(x_1, \ldots, x_{i-1}, x_k, x_{i+1}, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_\ell) \\
&= f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_\ell) \\
&\quad + f(x_1, \ldots, x_{i-1}, x_k, x_{i+1}, \ldots, x_{k-1}, x_i, x_{k+1}, \ldots, x_\ell)\,,
\end{aligned}$$

also

$$\begin{aligned}
&f(x_1, \ldots, x_{i-1}, x_k, x_{i+1}, \ldots, x_{k-1}, x_i, x_{k+1}, \ldots, x_\ell) \\
&\qquad = -f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_\ell)\,.
\end{aligned}$$

Sei nun $r > 1$ und wir nehmen an, die Aussage gelte für $r - 1$. Setzt man $y_j := x_{\tau_1(j)}$ für $1 \le j \le \ell$, dann folgt

$$\begin{aligned}
f(x_{\tau_1 \ldots \tau_r(1)}, \ldots, x_{\tau_1 \ldots \tau_r(\ell)}) &= f(x_{\tau_1(\tau_2 \ldots \tau_r(1))}, \ldots, x_{\tau_1(\tau_2 \ldots \tau_r(\ell))}) \\
&= f(y_{\tau_2 \ldots \tau_r(1)}, \ldots, y_{\tau_2 \ldots \tau_r(\ell)}) \\
&= (-1)^{r-1} f(y_1, \ldots, y_\ell) \\
&= (-1)^{r-1} f(x_{\tau_1(1)}, \ldots, x_{\tau_1(\ell)}) \\
&= (-1)^r f(x_1, \ldots, x_\ell)\,.
\end{aligned}$$

Hilfssatz 121: *Seien V, W Vektorräume über K und $f : V^\ell \to W$ alternierend. Sei $\underline{v} = (v_1, \ldots, v_n) \in V^n$, $T \in K^{n \times \ell}$ und $\underline{w} = \underline{v}T \in V^\ell$. Dann ist*

$$f(w_1, \ldots, w_\ell) = \sum_{(i_1 \ldots i_\ell) \in I} \det(T_{(i_1 \ldots i_\ell)}) f(v_{i_1}, \ldots, v_{i_\ell})\,,$$

wobei

$$I := \{(i_1, \ldots, i_\ell) \mid 1 \le i_1 < \ldots i_\ell \le n\}$$

ist und die Matrix

$$T_{(i_1\dots i_\ell)} := \begin{pmatrix} T_{i_1 1} & T_{i_1 2} & \dots & T_{i_1 \ell} \\ T_{i_2 1} & T_{i_2 2} & \dots & T_{i_2 \ell} \\ \vdots & & & \vdots \\ T_{i_\ell 1} & T_{i_\ell 2} & \dots & T_{i_\ell \ell} \end{pmatrix} \in K^{\ell\times\ell}$$

aus den Zeilen von T mit Indizes $i_1, i_2, \dots, i_\ell$ besteht.

Im Spezialfall $\ell = n$ enthält $I = \{(1,2,\dots,n)\}$ nur ein Element und die Behauptung vereinfacht sich zu

$$f(w_1,\dots,w_n) = \det(T)\cdot f(v_1,\dots,v_n)\,.$$

Beweis: Da f alternierend ist, gilt

$$\begin{aligned} f(w_1,\dots,w_\ell) &= f(\sum_{j_1=1}^n T_{j_1 1} v_{j_1},\dots,\sum_{j_\ell=1}^n T_{j_\ell \ell} v_{j_\ell}) \\ &= \sum_{j_1=1}^n \cdots \sum_{j_\ell=1}^n T_{j_1 1}\dots T_{j_\ell \ell}\, f(v_{j_1},\dots,v_{j_\ell}) \\ &= \sum_{j\in J} T_{j_1 1}\dots T_{j_\ell \ell}\, f(v_{j_1},\dots,v_{j_\ell}) \end{aligned}$$

mit $J := \{(j_1,\dots,j_\ell) \mid j_1,\dots,j_\ell \in \{1,\dots,\ell\}$ paarweise verschieden$\}$.
Da man jedes Tupel $(j_1,\dots,j_\ell) \in J$ der Größe nach ordnen kann, ist die Funktion

$$I \times S_\ell \to J\,,\; ((i_1,\dots,i_\ell),\sigma) \mapsto (i_{\sigma(1)},\dots,i_{\sigma(\ell)})\,,$$

surjektiv. Die Funktion ist auch injektiv, weil i_1 die kleinste Komponente von $(j_1,\dots,j_\ell)$ ist, i_2 die zweitkleinste Komponente von $(j_1,\dots,j_\ell)$ usw. und $\sigma(1),\dots,\sigma(\ell)$ die Positionen von $j_1,\dots,j_\ell$ in $(i_1,\dots,i_\ell)$ angibt. Umordnen der Summe mittels dieser Bijektion und Anwenden von Satz 120 liefert

$$\begin{aligned} f(w_1,\dots,w_\ell) &= \sum_{i\in I}\sum_{\sigma\in S_\ell} T_{i_{\sigma(1)}1}\dots T_{i_{\sigma(\ell)}\ell}\, f(v_{i_{\sigma(1)}},\dots,v_{i_{\sigma(\ell)}}) \\ &= \sum_{i\in I}\left(\sum_{\sigma\in S_\ell} \operatorname{sign}(\sigma) T_{i_{\sigma(1)}1}\dots T_{i_{\sigma(\ell)}\ell}\right) f(v_{i_1},\dots,v_{i_\ell}) \\ &= \sum_{i\in I} \det(T_{(i_1\dots i_\ell)})\, f(v_{i_1},\dots,v_{i_\ell})\,, \end{aligned}$$

was zu zeigen war.

Satz 122: *Seien V, W Vektorräume über K, $(v_1,\dots,v_n)$ eine Basis von V und $\ell \in \mathbb{N}$. Sei $(w_i)_{i\in I} \in \mathcal{F}(I,W)$, wobei*

$$I := \{(i_1,\dots,i_\ell) \mid 1 \le i_1 < \dots i_\ell \le n\}$$

ist. Dann gibt es genau eine alternierende Funktion $f: V^\ell \to W$ mit

$$f(v_{i_1}, \dots, v_{i_\ell}) = w_i$$

für alle $i = (i_1, \dots, i_\ell) \in I$. Somit kann eine alternierende Funktion durch Vorgabe der Bilder aller geordneten ℓ-Tupel von Basisvektoren eindeutig definiert werden.
Insbesondere: Für $\ell = n$, $V := K^{n\times 1}$ und $W := K$ erhalten wir eine neue Charakterisierung der Determinante: Die Determinante

$$D: \left(K^{n\times 1}\right)^n \to K\,,\ (x_1, \dots, x_n) \mapsto \det(x_1|\dots|x_n)\,,$$

ist die eindeutig bestimmte alternierende Funktion von $\left(K^{n\times 1}\right)^n$ nach K mit $D(e_1, e_2, \dots, e_n) = 1$.

Beweis: Wenn eine derartige Funktion f existiert, dann ist für Vektoren $x_1, \dots, x_\ell \in V$ mit Koordinatenspalten $c_1, \dots, c_\ell \in K^{n\times 1}$

$$(x_1, \dots, x_\ell) = (v_1, \dots, v_n)T$$

mit $T := (c_1|\dots|c_\ell) \in K^{n\times\ell}$ und nach Hilfssatz 121

$$f(x_1, \dots, x_\ell) = \sum_{(i_1\dots i_\ell)\in I} \det(T_{i_1\dots i_\ell})\, f(v_{i_1}, \dots, v_{i_\ell})\,,$$

was die Eindeutigkeit der Funktion beweist. Um die Existenz einer derartigen Funktion zu zeigen, definieren wir eine Funktion
$f: V^\ell \to W$ durch

$$f(x_1, \dots, x_\ell) := \sum_{(i_1\dots i_\ell)\in I} \det(T_{i_1\dots i_\ell})\, w_i\,,$$

wobei $T := (c_1|\dots|c_\ell) \in K^{n\times\ell}$ und $c_j \in K^{n\times 1}$ die Koordinatenspalte von x_j bzgl. der Basis $(v_1, \dots, v_n)$ ist. Dann ist nach Satz 119 und Satz 118 die Funktion f alternierend. Wegen $\det(I_\ell) = 1$ gilt auch $f(v_{i_1}, \dots, v_{i_\ell}) = w_i$ für alle $i \in I$.

Beispiel 123: Sei $V = W = \mathbb{R}^3$, $(v_1, v_2, v_3) = (e_1, e_2, e_3)$ die Standardbasis von $\mathbb{R}^3$ und $\ell = 2$. Dann ist

$$I = \{(1,2), (1,3), (2,3)\},$$

und es gibt genau eine alternierende Funktion $f: \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ mit

$$f(e_1, e_2) = e_3 \quad , \quad f(e_1, e_3) = -e_2 \quad \text{und} \quad f(e_2, e_3) = e_1\,.$$

Für Tripel $x = (x_1, x_2, x_3) = x_1e_1 + x_2e_2 + x_3e_3$, $y = (y_1, y_2, y_3) = y_1e_1 + y_2e_2 + y_3e_3 \in \mathbb{R}^3$ ist

$$f(x,y) = det\begin{pmatrix} x_2 & y_2 \\ x_3 & y_3 \end{pmatrix} e_1 - det\begin{pmatrix} x_1 & y_1 \\ x_3 & y_3 \end{pmatrix} e_2 + det\begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix} e_3\,,$$

also ist

$$x \times y := f(x,y)$$

das *Vektorprodukt* von x und y.

§3. Entwicklung von Determinanten

Definition 124: Sei $A \in K^{n\times n}$. Für $1 \leq i,j \leq n$ heißt die Matrix

$$A^{(i,j)} \in K^{(n-1)\times(n-1)},$$

die man aus A durch Weglassen der i-ten Zeile und j-ten Spalte erhält, die *i-j-te Streichungsmatrix*.

Satz 125: *Sei $A \in K^{n\times n}$. Dann gilt:*

(1) *Für $1 \leq j \leq n$ ist*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} A_{ij} \det(A^{(i,j)})$$

*(*Entwicklung der Determinante nach der j-ten Spalte*).*

(2) *Für $1 \leq i \leq n$ ist*

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} A_{ij} \det(A^{(i,j)})$$

*(*Entwicklung der Determinante nach der i-ten Zeile*).*

Wenn eine Matrix eine Spalte oder Zeile mit vielen Nullen besitzt, dann ist zur Berechnung der Determinante die Entwicklung nach dieser Spalte oder Zeile zu empfehlen.

Beweis:

(1) Nach Satz 119 ist die Determinante alternierend in den Spalten der Matrix. Wegen

$$A_{-j} = \sum_{i=1}^{n} A_{ij} e_i\,,$$

wobei $(e_1,\ldots,e_n)$ die Standardbasis von $K^{n\times 1}$ ist, folgt

$$\det(A) = \sum_{i=1}^{n} A_{ij} \det(A_{-1}|\ldots|A_{-(j-1)}|e_i|A_{-(j+1)}|\ldots|A_{-n})\,.$$

Sukzessives Vertauschen der j-ten Spalte mit der $(j+1)$-ten Spalte, der $(j+1)$-ten Spalte mit der $(j+2)$-ten Spalte usw. gibt

$$\det(A) = \sum_{i=1}^{n} (-1)^{n-j} A_{ij} \det(A_{-1}|\ldots|A_{-(j-1)}|A_{-(j+1)}|\ldots|A_{-n}|e_i)\,.$$

Sukzessives Vertauschen der i-ten Zeile mit der $(i+1)$-ten Zeile, der $(i+1)$-ten Zeile mit der $(i+2)$-ten Zeile usw. gibt weiters

$$\det(A) = \sum_{i=1}^{n} (-1)^{n-j+n-i} A_{ij} \det(B)$$

mit

$$B := \begin{pmatrix} & & & 0 \\ & A^{(i,j)} & & \vdots \\ & & & 0 \\ A_{i1} & \dots & A_{i,n-1} & 1 \end{pmatrix} \in K^{n\times n}.$$

Wegen

$$\begin{aligned}\det(B) &= \sum_{\sigma\in S_n} \operatorname{sign}(\sigma) B_{\sigma(1)1} \dots B_{\sigma(n-1)(n-1)} B_{\sigma(n)n} \\ &= \sum_{\substack{\sigma\in S_n \\ \sigma(n)=n}} \operatorname{sign}(\sigma) B_{\sigma(1)1} \dots B_{\sigma(n-1)(n-1)} \\ &= \sum_{\rho\in S_{n-1}} \operatorname{sign}(\rho) B_{\rho(1)1} \dots B_{\rho(n-1)(n-1)} = \det(A^{(i,j)})\end{aligned}$$

folgt

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} A_{ij} \det(A^{(i,j)}).$$

(2) Entwickeln der Determinante der transponierten Matrix nach der i-ten Spalte gibt

$$\begin{aligned}\det(A) = \det(A^T) &= \sum_{j=1}^{n} (-1)^{j+i} (A^T)_{ji} \det((A^T)^{(j,i)}) \\ &= \sum_{j=1}^{n} (-1)^{j+i} A_{ij} \det((A^{(i,j)})^T) = \sum_{j=1}^{n} (-1)^{i+j} A_{ij} \det(A^{(i,j)}).\end{aligned}$$

Definition 126: Sei K ein beliebiger Körper und $A \in K^{n\times n}$.

(1) Sei $p \in \mathbb{N}$ mit $p \geq 1$ sowie $n_1, \dots, n_p \in \mathbb{N}$ mit $n_1, \dots, n_p \geq 1$ und $n_1 + \dots + n_p = n$. Dann heißt die Darstellung

$$A = \begin{pmatrix} B_{11} & B_{12} & \dots & B_{1p} \\ B_{21} & B_{22} & \dots & B_{2p} \\ \vdots & \vdots & & \vdots \\ B_{p1} & B_{p2} & \dots & B_{p,p} \end{pmatrix},$$

wobei die Matrix $B_{k\ell} \in K^{n_k \times n_\ell}$ ist, eine *Blockzerlegung* von A mit *Blockgrößen* $(n_1, \dots, n_p)$ und *Blöcken* $B_{k\ell}$.
Es ist

$$(B_{k\ell})_{ij} := A_{n_1+\dots+n_{k-1}+i, n_1+\dots+n_{\ell-1}+j}, 1 \leq i \leq n_k, 1 \leq j \leq n_\ell.$$

(2) Die Matrix A hat *obere Blockdreiecksform mit Blockgrößen* $(n_1,\dots,n_p)$, wenn es eine Blockzerlegung von A mit Blockgrößen $(n_1,\dots,n_p)$ und Blöcken $B_{k\ell}$ gibt, sodass für alle Indizes $1 \le k,\ell \le p$ mit $k > \ell$ die Matrix $B_{k\ell} = 0$ ist, d.h.

$$A = \begin{pmatrix} B_{11} & B_{12} & \dots & B_{1p} \\ 0 & B_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & B_{p-1,p} \\ 0 & \dots & 0 & B_{pp} \end{pmatrix}$$

ist, wobei die Nullen für die Nullmatrizen entsprechender Größe stehen.

(3) Die Matrix A hat *untere Blockdreiecksform mit Blockgrößen* $(n_1,\dots,n_p)$, wenn es eine Blockzerlegung von A mit Blockgrößen $(n_1,\dots,n_p)$ und Blöcken $B_{k\ell}$ gibt, sodass für alle Indizes $1 \le k,\ell \le p$ mit $k < \ell$ die Matrix $B_{k\ell} = 0$ ist.

(4) Die Matrix A hat *Blockdiagonalform mit Blockgrößen* $(n_1,\dots,n_p)$, wenn es eine Blockzerlegung von A mit Blockgrößen $(n_1,\dots,n_p)$ und Blöcken $B_{k\ell}$ gibt, sodass für alle Indizes $1 \le k,\ell \le p$ mit $k \ne \ell$ die Matrix $B_{k\ell} = 0$ ist.

Satz 127 : *Sei $A \in K^{n\times n}$ in oberer oder unterer Blockdreiecksform mit Blockgrößen $(n_1,\dots,n_p)$ und Blöcken $B_{k\ell}$. Dann ist*

$$\det(A) = \det(B_{11})\det(B_{22})\dots\det(B_{pp})\,,$$

d.h. die Determinante einer Matrix in Blockdreiecksform berechnet sich als Produkt der Determinanten ihrer Diagonalblöcke.

Beweis: Wir beweisen die Behauptung durch Induktion nach n. Für $n = 1$ ist $p = 1$ und wegen $A = B_{11}$ nichts zu zeigen. Für $n \ge 2$ entwickeln wir die Determinante von A nach der ersten Spalte und erhalten nach Satz 125

$$\det(A) = \sum_{i=1}^{n_1} (-1)^{i+1} A_{i1} \det(A^{(i,1)})\,.$$

Da die Streichungsmatrix $A^{(i,1)}$ wieder Blockdreiecksform besitzt, können wir die Induktionsannahme anwenden und erhalten nach Satz 125

$$\begin{aligned} \det(A) &= \sum_{i=1}^{n_1} (-1)^{i+1} A_{i1} \det((B_{11})^{(i,1)}) \det(B_{22})\dots\det(B_{pp}) \\ &= \left(\sum_{i=1}^{n_1} (-1)^{i+1} (B_{11})_{i1} \det((B_{11})^{(i,1)})\right) \det(B_{22})\dots\det(B_{pp}) \\ &= \det(B_{11})\det(B_{22})\dots\det(B_{pp})\,. \end{aligned}$$

§4. Die adjungierte Matrix

Satz 128: *Sei $A \in K^{n\times n}$. Dann heißt die Matrix* $\mathrm{Ad}(A) \in K^{n\times n}$ *, definiert durch*

$$\mathrm{Ad}(A)_{ij} := (-1)^{i+j} \det(A^{(j,i)})$$

für $1 \leq i, j \leq n$, die zu A adjungierte Matrix*, und es gilt*

$$A \cdot \mathrm{Ad}(A) = \mathrm{Ad}(A) \cdot A = \det(A) \cdot I_n .$$

Beweis: Für $1 \leq i, j \leq n$ ist wegen Satz 125,(2)

$$\begin{aligned}(\mathrm{Ad}(A)\cdot A)_{ji} &= \sum_{k=1}^{n} \mathrm{Ad}(A)_{jk} A_{ki} = \sum_{k=1}^{n} A_{ki}(-1)^{j+k} \det(A^{(k,j)}) \\ &= \det(A_{-1}|\dots|A_{-(j-1)}|A_{-i}|A_{-(j+1)}|\dots|A_{-n}),\end{aligned}$$

wobei A_{-i} in der j-ten Spalte steht. Daher ist

$$(A\cdot \mathrm{Ad}(A))_{ij} = \begin{cases} \det(A) & \text{falls } i = j, \\ 0 & \text{falls } i \neq j, \end{cases}$$

also $A \cdot \mathrm{Ad}(A) = \det(A) \cdot I_n$. Anwenden auf A^T gibt
$A^T \cdot \mathrm{Ad}(A^T) = \det(A^T) \cdot I_n$, also $A^T \cdot \mathrm{Ad}(A)^T = \det(A) \cdot I_n$
und $\mathrm{Ad}(A) \cdot A = \det(A) \cdot I_n$.

Satz 129: *Eine Matrix $A \in K^{n\times n}$ ist genau dann invertierbar, wenn ihre Determinante ungleich 0 ist. In diesem Fall gilt*

$$A^{-1} = \det(A)^{-1} \mathrm{Ad}(A)$$

und

$$\det(A^{-1}) = \det(A)^{-1} .$$

Beweis: Wenn A invertierbar ist, dann folgt aus $A \cdot A^{-1} = I_n$ durch Anwenden der Determinante $\det(A) \cdot \det(A^{-1}) = 1$, also $\det(A) \neq 0$ und $\det(A^{-1}) = \det(A)^{-1}$. Wenn umgekehrt $\det(A) \neq 0$ ist, dann ist nach Satz 128

$$A(\det(A)^{-1} \mathrm{Ad}(A)) = \det(A)^{-1} \det(A) I_n = I_n$$

und

$$(\det(A)^{-1} \mathrm{Ad}(A))A = \det(A)^{-1} \det(A) I_n = I_n ,$$

also $A^{-1} = \det(A)^{-1} \mathrm{Ad}(A)$.

Beispiel 130: Sei

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in K^{2\times 2} .$$

Dann ist $\det(A) = ad - bc$ und, falls $\det(A) \neq 0$ ist,

$$A^{-1} = \det(A)^{-1} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Satz 131 : *(*Cramersche Regel*) Sei $A \in K^{n\times n}$ und $b \in K^{n\times 1}$. Dann ist das System linearer Gleichungen (A,b) genau dann eindeutig lösbar, wenn* $\det(A) \neq 0$ *ist. In diesem Fall ist die eindeutige Lösung $x \in K^{n\times 1}$ gegeben durch*

$$x_i = \det(A_{-1}|\dots|A_{-(i-1)}|\,b\,|A_{-(i+1)}|\dots|A_{-n}) / \det(A)$$

für $i = 1,\dots,n$.

Beweis: Wenn $\det(A) \neq 0$ ist, dann ist nach Satz 129 die Matrix A invertierbar und $A^{-1}b$ die eindeutige Lösung des linearen Gleichungssystems. Wenn umgekehrt das System (A,b) eindeutig lösbar ist, dann ist $L(A,0) = \{0\}$, folglich $\operatorname{rg}(A) = n - \dim_K(L(A,0)) = n$, somit A äquivalent zu I_n und $\det(A) \neq 0$.

In diesem Fall gilt für die Lösung $x \in K^{n\times 1}$ und für $1 \leq i \leq n$

$$\begin{aligned} \det(A_{-1}|\dots|A_{-(i-1)}|\,b\,|A_{-(i+1)}|\dots|A_{-n}) &= \\ \det(A_{-1}|\dots|A_{-(i-1)}|Ax|A_{-(i+1)}|\dots|A_{-n}) &= \\ \det(A_{-1}|\dots|A_{-(i-1)}|\sum_{k=1}^{n} x_k A_{-k}|A_{-(i+1)}|\dots|A_{-n}) &= \\ x_i \det(A_{-1}|\dots|A_{-(i-1)}|A_{-i}|A_{-(i+1)}|\dots|A_{-n}) &= x_i \det(A), \end{aligned}$$

woraus die Behauptung folgt.

§5. Symmetrische Bilinearformen

Definition 132 : Eine bilineare Funktion $b : V \times V \to K$ heißt *Bilinearform auf* V.
Eine Bilinearform b ist *symmetrisch*, wenn für alle $v,w \in V$

$$b(v,w) = b(w,v)$$

ist.
Für eine Basis $\underline{v}$ von V heißt

$$M(b,\underline{v}) := \left(b\left(v_i,v_j\right)\right)_{1\leq i,j\leq n}$$

die *Matrix von b* bezüglich $\underline{v}$. Eine Matrix $A \in K^{n\times n}$ ist *symmetrisch*, wenn

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

ist.

Beispiel 133 : Wenn V ein reeller Vektorraum ist, dann ist jedes Skalarprodukt auf V eine symmetrische Bilinerarform. Die Matrix jedes Skalarproduktes bezüglich jeder ON-Basis ist I_n.

Beispiel 134 : Sei $A \in K^{n\times n}$. Dann ist die Funktion

$$K^{n\times 1} \times K^{n\times 1} \to K \quad , \quad (x,y) \longmapsto x^\top \cdot A \cdot y \ ,$$

eine Bilinearform und ihre Matrix bezüglich der Standardbasis ist A.

Satz 135 : *Es seien b eine Bilinearform auf V, $\underline{v}$ eine Basis von V und A die Matrix von b bezüglich $\underline{v}$. Dann gilt:*

(1) *Für alle $x,y \in K^{n\times 1}$ ist*
$$b\left(\underline{v}x, \underline{v}y\right) = x^\top \cdot A \cdot y \ .$$
(2) *Die Matrix A ist genau dann symmetrisch, wenn b symmetrisch ist.*
(3) *Für alle $S \in \mathrm{GL}_n(K)$ ist*
$$M\left(b, \underline{v}S\right) = S^\top \cdot A \cdot S \ .$$

Beweis:

(1)
$$b\left(\underline{v}x, \underline{v}y\right) = b\left(\sum_{i=1}^n x_i v_i, \sum_{j=1}^n y_j v_j\right) = \sum_{i,j=1}^n x_i y_j \cdot b\left(v_i, v_j\right)$$
$$= \sum_{i,j=1}^n x_i A_{ij} y_j = x^\top \cdot A \cdot y.$$

(2) Aus b symmetrisch folgt
$A_{ij} = b\left(v_i, v_j\right) = b\left(v_j, v_i\right) = A_{ji}, 1 \le i, j \le n.$
Umgekehrt folgt aus A symmetrisch, dass
$x^\top \cdot A \cdot y = \left(x^\top \cdot A \cdot y\right)^\top = y^\top \cdot A^\top \cdot x = y^\top \cdot A \cdot x$
ist, nach (1) ist b daher symmetrisch.

(3) Für $1 \le i, j \le n$ ist
$$b\left(\underline{v}S_{-i}, \underline{v}S_{-j}\right) = \sum_{k,\ell=1}^n S_{ki} A_{k\ell} S_{\ell j} = \left(S^\top \cdot A \cdot S\right)_{ij} \ .$$

Definition 136 : Zwei Matrizen $A, B \in K^{n\times n}$ heißen *kongruent*, wenn es eine Matrix $P \in \mathrm{GL}_n(K)$ mit
$$B = P^\top \cdot A \cdot P$$
gibt. (Zwei kongruente Matrizen beschreiben dieselbe Bilinearform bezüglich zweier Basen von V).

Definition 137 : Für $d_1, \dots, d_n \in K$ sei

$$Diag\,(d_1, \dots, d_n) := \begin{pmatrix} d_1 & & & 0 \\ & d_2 & & \\ & & \ddots & \\ 0 & & & d_n \end{pmatrix} \in K^{n \times n}$$

die Matrix in Diagonalform mit Diagonalelementen $d_1, \dots, d_n$.

Satz 138 : *Es seien $A \in K^{n\times n}$ eine symmetrische Matrix und r der Rang von A. Dann gibt es $d_1, \dots, d_r \in K \setminus \{0\}$ so, dass A und*

$$Diag\,(d_1, d_2, \dots, d_r, 0, \dots, 0) \in K^{n \times n}$$

kongruent sind. Mit dem folgenden Verfahren können $P \in \mathrm{GL}_n(K)$ und $d_1, \dots, d_r \in K$ so berechnet werden, dass

$$P^\top \cdot A \cdot P = Diag\,(d_1, d_2, \dots, d_r, 0, \dots, 0)$$

ist:

(1) *Setze $C := (A \mid I_n)$ und $i := 1$.*
(2) *Wenn $i > n$ oder $C_{k\ell} = 0$ für alle $k, \ell \in \{i, \dots, n\}$ ist, dann ist*

$$C = \left(Diag\,(d_1, \dots, d_{i-1}, 0, \dots, 0) \mid P^\top \right)$$

und das Verfahren zu Ende.
(3) *Falls $C_{jj} = 0$ ist für alle $j \in \{i, \dots, n\}$ und $C_{k\ell} \neq 0$ ist für Indizes $k, \ell \in \{i, \dots, n\}$, addiere die k-te Zeile von C zur ℓ-ten und anschließend die k-te Spalte zur ℓ-ten. Nenne die neue Matrix wieder C.*
(4) *Falls $C_{ii} = 0$ ist, aber $C_{jj} \neq 0$ ist für ein $j \in \{i+1, \dots, n\}$, vertausche die j-te mit der i-ten Zeile von C und dann die j-te mit der i-ten Spalte. Nenne die neue Matrix wieder C.*
(5) *Falls $C_{ii} \neq 0$ ist, subtrahiere für $j = i+1, \dots, n$ die $(C_{ii}^{-1} \cdot C_{ji})$-fache i-te Zeile von C von der j-ten und dann die $(C_{ii}^{-1} \cdot C_{ji})$-fache i-te Spalte von der j-ten. Nenne die neue Matrix wieder C, setze $i := i+1$ und fahre mit (2) fort.*

Beweis: Die im Verfahren angegebenen Umformungen von C bedeuten jeweils, dass A und I_n von rechts mit geeigneten Elementarmatrizen Q und von links mit $Q^\top$ multipliziert werden. Insbesondere sind alle Matrizen, die von den ersten n Spalten der mit C bezeichneten Matrizen gebildet werden, symmetrisch.
Mit $r = \mathrm{rg}\,(A) = \mathrm{rg}\left(P^\top \cdot A \cdot P\right)$ folgt daraus die Behauptung.

Satz 139 : *„Trägheitssatz von Sylvester“*

(1) *Jede komplexe symmetrische Matrix A ist kongruent zu*

$$Diag\,(1,1,...,1,0,...,0)\ ,$$

wobei die Anzahl der Einsen in der Diagonalmatrix gleich dem Rang von A ist. Insbesondere sind zwei komplexe symmetrische Matrizen genau dann kongruent, wenn sie den gleichen Rang haben.

(2) *Jede reelle symmetrische Matrix A ist kongruent zu genau einer der Matrizen*

$$I_n^{s,t} := Diag(1,...,1,-1,...,-1,0,...,0),$$

wobei s bzw. t die Anzahl der Einträge 1 *bzw.* -1 *in* $I_n^{s,t}$ *ist. Es gilt* $s+t = \mathrm{rg}\,(A)$ *. Das Paar* (s,t) *heißt* Signatur von A.
Insbesondere sind zwei reelle symmetrische Matrizen genau dann kongruent, wenn sie die gleiche Signatur haben.

Beweis: Nach Satz 138 gibt es $P \in \mathrm{GL}_n(K)$ und $d_1,\ldots,d_r \in K\backslash\{0\}$ so, dass $P^T \cdot A \cdot P = Diag(d_1,\ldots,d_r,0,\ldots 0)$ ist.

(1) Wähle Zahlen $c_1,\ldots,c_r \in \mathbb{C}$ so, dass $c_i^2 = d_i^{-1}, 1 \le i \le r$, und setze $Q := P \cdot Diag(c_1,\ldots,c_r,0,\ldots,0)$.
Dann ist $Q^T \cdot A \cdot Q = Diag(1,\ldots,1,0,\ldots,0)$.

(2) Wir können annehmen, dass $d_1 \ge d_2 \ge \ldots \ge d_r$ ist.
Wähle Zahlen $c_1,\ldots,c_r \in \mathbb{R}$ so, dass $c_i^2 = |d_i|^{-1}$ für $1 \le i \le r$ ist, und setze $Q := P \cdot Diag(c_1,\ldots,c_r,0,\ldots.,0)$. Dann ist $Q^T \cdot A \cdot Q = I_n^{s,t}$, wobei s bzw. t die Anzahl der Indizes i mit $d_i > 0$ bzw. $d_i < 0$ ist.
Es ist noch zu zeigen, dass für $S \in \mathrm{GL}_n(\mathbb{R})$ aus $S^T \cdot I_n^{s',t'} \cdot S = I_n^{s,t}$ folgt, dass $s' = s$ und $t' = t$ ist. Wegen $s+t = \mathrm{rg}(I_n^{s,t})$ genügt es, $s' = s$ zu zeigen. Wir können annehmen, dass $s' \ge s$ gilt.
Sei $U := \{y \in \mathbb{R}^{n\times 1} \mid y_1 = \ldots = y_s = 0\}$ und
$W := \{y \in \mathbb{R}^{n\times 1} \mid (Sy)_{s'+1} = (Sy)_{s'+2} = \ldots = (Sy)_n = 0\}$.
Dann ist $\dim_{\mathbb{R}}(U) = n-s$ und wegen $S \in \mathrm{GL}_n(\mathbb{R})$ auch $\dim_{\mathbb{R}}(W) = s'$.
Wäre $s' > s$, dann wäre

$$\dim_{\mathbb{R}}(U) + \dim_{\mathbb{R}}(W) = n - s + s' > n$$

und daher $U \cap W \ne \{0\}$.
Für $0 \ne x \in U \cap W$ würde dann

$$0 \ge -\sum_{i=s+1}^{s+t} x_i^2 = x^T \cdot I_n^{s,t} \cdot x =$$

$$= (Sx)^T \cdot I_n^{s',t'} \cdot (Sx) = \sum_{i=1}^{s'} (Sx)_i^2 > 0$$

sein. Widerspruch.

§6. Positiv definite Matrizen

Definition 140 : Eine Bilinearform b auf einem reellen Vektorraum V heißt *positiv definit* bzw. *positiv semidefinit*, wenn für alle $v \in V \backslash \{0\}$

$$b(v,v) > 0 \quad \text{bzw.} \quad b(v,v) \geq 0$$

ist. Eine Matrix $A \in \mathbb{R}^{n\times n}$ heißt *positiv definit* bzw. *positiv semidefinit*, wenn für alle $x \in \mathbb{R}^{n\times 1}\backslash\{0\}$

$$x^\top \cdot A \cdot x > 0 \quad \text{bzw.} \quad x^\top \cdot A \cdot x \geq 0$$

ist.

Beispiel 141 : Die Matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ ist positiv definit und damit auch positiv semidefinit.
Die Matrix $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ ist positiv semidefinit, aber nicht positiv definit.

Satz 142 :

(1) *Eine Bilinearform ist genau dann positiv definit bzw. positiv semidefinit, wenn ihre Matrix bezüglich einer Basis von V positiv definit bzw. positiv semidefinit ist.*
(2) *Wenn eine Matrix positiv definit bzw. positiv semidefinit ist, dann sind dies auch alle zu ihr kongruenten Matrizen.*

Beweis: Übung.

Satz 143 : *Für eine symmetrische Matrix $A \in \mathbb{R}^{n\times n}$ sind die folgenden Aussagen äquivalent:*

(1) *A ist positiv definit.*
(2) *Es gibt eine Matrix $B \in \mathrm{GL}_n(\mathbb{R})$ so, dass $A = B^\top \cdot B$ ist.*
(3) *Die Signatur von A ist $(n,0)$.*

Beweis: Eine reelle Diagonalmatrix ist genau dann positiv definit, wenn alle ihre Einträge in der Diagonale positiv sind. Nach Satz 139 gibt es eine invertierbare Matrix P so, dass $P^\top \cdot A \cdot P =: D$ eine Diagonalmatrix ist, deren Einträge $0, 1$ oder -1 sind.

(1) $\Rightarrow$ (2) : Wenn A positiv definit ist, muss D nach Satz 142 auch positiv definit sein, also $D = I_n$ sein. Mit $B := P^{-1}$ folgt (2).
(2) $\Rightarrow$ (3) : Nach (2) ist A kongruent zu I_n.
(3) $\Rightarrow$ (1) : Nach (3) und Satz 139 ist A kongruent zu I_n. Nach Satz 142 ist A positiv definit.

Mit Aussage (3) kann überprüft werden, ob eine symmetrische Matrix positiv definit ist.
Mit Aussage (2) können Beispiele für positiv definite symmetrische Matrizen (und damit für Skalarprodukte) konstruiert werden.
Eine Anwendung in der Analysis: Es seien $f : \mathbb{R}^n \to \mathbb{R}$ eine 2-mal stetig differenzierbare Funktion, deren Ableitung im Punkt p Null ist, und A ihre Hesse'sche Matrix in p. Wenn A bzw. $-A$ positiv definit ist, dann hat f in p ein isoliertes Minimum bzw. Maximum.

Satz 144 : *Für eine symmetrische Matrix $A \in \mathbb{R}^{n\times n}$ sind die folgenden Aussagen äquivalent:*

(1) *A ist positiv semidefinit.*
(2) *Es gibt eine Matrix $B \in \mathbb{R}^{n\times n}$ so, dass $A = B^\top \cdot B$ ist.*
(3) *Die Signatur von A ist $(r,0)$, wobei r der Rang von A ist.*

Beweis: Analog dem Beweis von Satz 143.

Satz 145 : *Eine symmetrische reelle Matrix A ist genau dann positiv definit, wenn alle ihre* Hauptminoren

$$\det\begin{pmatrix} A_{11} & \cdots & A_{1k} \\ \vdots & \ddots & \vdots \\ A_{k1} & \cdots & A_{kk} \end{pmatrix},\ 1 \le k \le n,$$

positive Zahlen sind.

Beweis: Für $1 \le k \le n$ sei

$$A(k) := \begin{pmatrix} A_{11} & \cdots & A_{1k} \\ \vdots & \ddots & \vdots \\ A_{k1} & \cdots & A_{kk} \end{pmatrix} \in \mathbb{R}^{k\times k}$$

Wenn A positiv definit ist, dann ist für alle $x \in \mathbb{R}^{k\times 1} \setminus \{0\}$

$$x^\top \cdot A(k) \cdot x = \begin{pmatrix} x \\ 0 \end{pmatrix}^\top \cdot A \cdot \begin{pmatrix} x \\ 0 \end{pmatrix} > 0\,,$$

daher ist $A(k)$ positiv definit. Aus Satz 143, (2) folgt dann, dass $\det(A(k)) > 0$ ist.

Seien nun alle Hauptminoren positiv. Wie zeigen durch Induktion nach n, dass A positiv definit ist. Nach Induktionsannahme ist $A(n-1)$ positiv definit, nach Satz 143, $(1) \Rightarrow (2)$, daher invertierbar.
Somit bilden die Zeilen von $A(n-1)$ eine Basis von $\mathbb{R}^{1\times(n-1)}$, deshalb gibt

es Zahlen $c_1, \ldots, c_{n-1} \in \mathbb{R}$ so, dass

$$(A_{n1}, \ldots, A_{n(n-1)}) = \sum_{i=1}^{n-1} c_i A(n-1)_{i-}$$

ist. Wird das c_i-fache der i-ten Zeile von A von der n-ten Zeile subtrahiert und das c_i-fache der i-ten Spalte von der n-ten Spalte, $1 \leq i \leq n$, dann erhalten wir eine zu A kongruente Matrix

$$C := \begin{pmatrix} A(n-1) & 0 \\ 0 & d \end{pmatrix} \in \mathbb{R}^{n \times n}$$

mit $d \in \mathbb{R}$. Da wir C aus A erhalten, indem wir A von links und von rechts mit Elementarmatrizen, deren Determinante 1 ist, multiplizieren, ist
$\det(A) = \det(C) = \det(A(n-1)) \cdot d$.
Aus $\det(A) > 0$ und $\det(A(n-1)) > 0$ folgt $d > 0$.
Da $A(n-1)$ positiv definit ist, gibt es nach Satz 143 eine Matrix
$B \in GL_{n-1}(\mathbb{R})$ so, dass $A(n-1) = B^\top \cdot B$ ist. Die Matrix

$$M := \begin{pmatrix} B & 0 \\ 0 & \sqrt{d} \end{pmatrix}$$

ist invertierbar und $C = M^\top \cdot M$. Nach Satz 143 ist C positiv definit, daher auch die zu C kongruente Matrix A.

KAPITEL 5

Quadratische Funktionen und Quadriken

Es seien K ein Körper mit $1_K + 1_K \neq 0_K$ und V ein n-dimensionaler Vektorraum über K.

§1. Linearformen

Definition 146 : Eine lineare Abbildung von V nach K heißt *Linearform auf* V. Die Menge $\mathrm{Lin}_K(V,K)$ aller linearen Abbildungen von V nach K bezeichnen wir kurz mit V^*.
V^* mit der punktweisen Addition und Skalarmultiplikation heißt *Vektorraum der Linearformen auf V* oder *zu V dualer Vektorraum.*

Satz 147 : *Es seien $\underline{v} = (v_1, \ldots, v_n)$ eine Basis von V und $1 \leq i \leq n$. Die Abbildung*

$$X_i : V \longrightarrow K \quad , w \longmapsto \quad \textit{Koordinate von } w \textit{ bei } v_i \ ,$$

ist eine Linearform und heißt i-te Koordinatenfunktion bezüglich $\underline{v}$. *Für $1 \leq i, j \leq n$ ist $X_i(v_j) = \delta_{ij}$.*
Das n-Tupel $\underline{X} := (X_1, \ldots, X_n)$ ist eine Basis von V^* und heißt die zu $\underline{v}$ duale Basis von V^*. *Für $f \in V^*$ ist*

$$f = \sum_{i=1}^{n} f(v_i) X_i \ .$$

Insbesondere haben V und V^* dieselbe Dimension und sind daher isomorph.

Beweis: Es ist leicht nachzuprüfen, dass die Koordinatenfunktionen linear sind.
Aus $\sum_{i=1}^{n} c_i X_i = 0$ folgt für alle $j \in \{1, \ldots, n\}$

$$0 = (\sum_{i=1}^{n} c_i X_i)(v_j) = \sum_{i=1}^{n} c_i X_i(v_j) = c_j \ .$$

Daher ist $(X_1, \ldots, X_n)$ linear unabhängig.
Sei $f \in V^*$, dann ist für alle $j \in \{1, \ldots, n\}$

$$\left(\sum_{i=1}^{n} f(v_i) X_i \right) (v_j) = \sum_{i=1}^{n} f(v_i) X_i(v_j) = f(v_j) \ ,$$

also $f = \sum_{i=1}^{n} f(v_i) X_i$.

Satz 148: *Es seien V ein reeller oder komplexer Vektorraum mit Skalarprodukt $\langle -,- \rangle$, $\underline{v} = (v_1, \dots, v_n)$ eine Basis von V und $\underline{X} := (X_1, \dots, X_n)$ die dazu duale Basis von V^*. Für jeden Vektor $v \in V$ ist die Abbildung*

$$\langle v, - \rangle : V \longrightarrow K \quad , \quad w \longmapsto \langle v, w \rangle \; ,$$

linear. Die Abbildung

$$V \longrightarrow V^* \quad , \quad v \longmapsto \langle v, - \rangle \; ,$$

ist ein Isomorphismus von Vektorräumen. Wenn $\underline{v}$ eine ON-Basis ist, dann ist $\langle v_i, - \rangle = X_i$, für $1 \leq i \leq n$.

Beweis: Übung.

Satz 149: *Es seien $S \in GL_n(K)$, $\underline{w} := \underline{v}S$ und $\underline{X}$ bzw. $\underline{Y}$ die zu $\underline{v}$ bzw. $\underline{w}$ dualen Basen von V^*. Dann ist*

$$\underline{Y} = \underline{X}(S^\top)^{-1} \; .$$

(„Transformieren sich die Basen mit S, dann transformieren sich die dazu dualen Basen mit $(S^\top)^{-1}$“).
Wenn $K = \mathbb{R}$ und $S \in \mathcal{O}_n$ ist, dann gilt $\underline{Y} = \underline{X}S$. („Bei orthogonalem Basiswechsel transformieren sich die dualen Basen wie die Basen“).

Beweis: Für alle $j \in \{1, \dots, n\}$ ist

$$Y_j = \sum_{i=1}^{n} Y_j(v_i) X_i = \sum_{i=1}^{n} Y_j \left((\underline{w}S^{-1})_i \right) X_i =$$

$$= \sum_{i=1}^{n} (S^{-1})_{ji} X_i = \sum_{i=1}^{n} ((S^\top)^{-1})_{ij} X_i = \left(\underline{X}(S^\top)^{-1} \right)_j \; .$$

Satz 150: *Es seien $\underline{v}$ eine Basis von V, $\underline{X}$ die dazu duale Basis von V^* und*

$$\mathbf{1} : V \longrightarrow K \quad , \quad w \longmapsto 1 \; .$$

(1) *Die Menge der affinen Abbildungen von V nach K ist mit der punktweisen Addition und Skalarmultiplikation ein Vektorraum.*
Das $(n+1)$-Tupel $(\mathbf{1}, X_1, \dots X_n)$ ist eine Basis dieses Vektorraums. Für eine affine Abbildung $f : V \longrightarrow K$ ist

$$f = f(0)\mathbf{1} + \sum_{i=1}^{n} (f(v_i) - f(0)) X_i \; .$$

(2) *Es seien $g : V \longrightarrow V$ eine lineare Abbildung, A ihre Matrix bezüglich $\underline{v}$, $u = \sum_{i=1}^{n} z_i v_i \in V$ und $z := (z_1, \dots, z_n) \in K^n$. Dann ist für alle $j \in \{1, \dots, n\}$*

$$X_j \circ g = \sum_{i=1}^{n} A_{ji} X_i$$

(kurz: $\underline{X} \circ g = \underline{X} A^\top$),

$$X_j \circ t_u = X_j + z_j \mathbf{1}$$

(kurz: $\underline{X} \circ t_u = z\mathbf{1} + \underline{X}$) und

$$X_j \circ (t_u \circ g) = z_j \mathbf{1} + \sum_{i=1}^{n} A_{ji} X_i$$

(kurz: $\underline{X} \circ (t_u \circ g) = z\mathbf{1} + \underline{X} A^\top$).

Beweis: Übung.

Definition 151 : Es sei f eine beliebige Abbildung von V nach K. Die Menge

$$\mathcal{N}(f) := \{v \in V \mid f(v) = 0\}$$

heißt *Nullstellenmenge von f*.

Satz 152 : *Es seien f eine Abbildung von V nach K, W ein Vektorraum und h eine surjektive Abbildung von W nach V. Dann ist*

$$h^{-1}(\mathcal{N}(f)) = \mathcal{N}(f \circ h) .$$

Beweis: Für $w \in W$ ist $(f \circ h)(w) = 0$ genau dann, wenn $h(w) \in \mathcal{N}(f)$.

Dieser Satz liefert eine Idee zur Beschreibung von Nullstellenmengen gewisser Abbildungen f: Finde eine „einfache“ surjektive Abbildung h so, dass die Nullstellenmenge von $f \circ h$ „gut bekannt“ ist. Dann kann $\mathcal{N}(f)$ als Bild von $\mathcal{N}(f \circ h)$ unter h beschrieben werden.

Beispiel 153 : Seien $f : V \longrightarrow K$ eine affine Abbildung, $f \notin K \cdot \mathbf{1}$, $\underline{v}$ eine Basis von V und $\underline{X}$ die dazu duale Basis. Dann ist

$$f = f(0)\mathbf{1} + \sum_{i=1}^{n} (f(v_i) - f(0)) X_i$$

und $\mathcal{N}(f)$ eine Hyperebene in V. Mit Satz 150 kann eine bijektive affine Abbildung $h : V \longrightarrow V$ so gewählt werden, dass $f \circ h = X_1$. Dann ist $\mathcal{N}(f)$

das Bild der „Koordinatenhyperebene“ $\mathcal{N}(X_1)$ unter der affinen Abbildung h.

§2. Quadratische Formen

Es seien $\underline{v}$ eine Basis von V und $\underline{X}$ die dazu duale Basis von V^*.

Definition 154 : Für Abbildungen f und g von V nach K heißt die Abbildung

$$f \cdot g : V \longrightarrow K \quad , \quad w \longmapsto f(w) \cdot g(w) \ ,$$

das *punktweise Produkt* von f und g.
Oft wird statt $f \cdot g$ nur fg geschrieben und f^2 statt $f \cdot f$.

Satz 155 : *Die Menge $\mathcal{F}(V,K)$ ist mit der punktweisen Addition und Multiplikation ein kommutativer Ring mit Einselement*

$$\mathbf{1} : V \longrightarrow K \quad , \quad w \longmapsto 1 \ .$$

Mit der punktweisen Addition und Skalarmultiplikation ist $\mathcal{F}(V,K)$ ein Vektorraum. Für $c \in K$ und $f,g \in \mathcal{F}(V,K)$ ist

$$c(f \cdot g) = (cf) \cdot g = f \cdot (cg) \ .$$

Beweis: Übung.

Definition 156 : Eine Abbildung $q : V \longrightarrow K$ heißt *quadratische Form* auf V, wenn sie Summe von Produkten von Linearformen auf V ist, dh. es gibt Linearformen $f_1, \ldots, f_m, g_1, \ldots, g_m$ auf V so, dass

$$q = \sum_{i=1}^{m} f_i \cdot g_i$$

ist. Die Menge aller quadratischen Formen ist ein Untervektorraum von $\mathcal{F}(V,K)$ und wird mit $Q(V)$ bezeichnet.

Satz 157 :

(1) *Die Familie $(X_i \cdot X_j)_{1 \leq i \leq j \leq n}$ ist eine K-Basis von $Q(V)$.*
Insbesondere ist

$$\dim_K(Q(V)) = \frac{n(n+1)}{2} \ .$$

(2) *Für $c \in K, w \in V$ und $q \in Q(V)$ ist*

$$q(cw) = c^2 q(w) \ .$$

Beweis:

(1) Es seien $f = \sum_{i=1}^n c_i X_i$ und $g = \sum_{i=1}^n d_j X_j$. Nach Satz 155 ist dann

$$f \cdot g = \sum_{1 \le i,j \le n} c_i d_j X_i X_j = \sum_{i=1}^n c_i d_i X_i^2 + \sum_{1 \le i < j \le n} (c_i d_j + c_j d_i) X_i X_j \ ,$$

also ist $(X_i \cdot X_j)_{1 \le i \le j \le n}$ ein Erzeugendensystem von $Q(V)$.
Wenn eine Linearkombination

$$\sum_{1 \le i \le j \le n} c_{ij} X_i X_j = 0$$

ist, dann ist für alle k

$$0 = \left(\sum_{1 \le i \le j \le n} c_{ij} X_i X_j \right) (v_k) = c_{kk}$$

und für alle k, l mit $k < l$

$$0 = \left(\sum_{1 \le i \le j \le n} c_{ij} X_i X_j \right) (v_k + v_l) = c_{kl} \ ,$$

daher ist $(X_i \cdot X_j)_{1 \le i \le j \le n}$ auch linear unabhängig.

(2) Für $c \in K$, $w \in V$ und $f, g \in V^*$ ist

$$(f \cdot g)(cw) = f(cw) g(cw) = c^2 f(w) g(w) = c^2 (f \cdot g)(w) \ .$$

Satz 158 :

(1) *Ist $q : V \longrightarrow K$ eine quadratische Form, dann ist*

$$b_q : V \times V \longrightarrow K \quad , \quad (u, w) \longmapsto \frac{1}{2} [q(u + w) - q(u) - q(w)] \ ,$$

eine symmetrische Bilinearform und heißt die durch q definierte Bilinearform. *Für $w \in V$ ist dann*

$$q(w) = b_q(w, w) \ .$$

Die Matrix von q *bezüglich $\underline{v}$ (Schreibweise: $M(q, \underline{v})$) ist die Matrix von b_q bezüglich $\underline{v}$.*
Der Rang von q *bzw. die* Signatur von q *ist dann der Rang bzw. die Signatur dieser Matrix.*

(2) *Ist $b : V \times V \longrightarrow K$ eine symmetrische Bilinearform auf V, dann ist*

$$q_b : V \longrightarrow K \quad , \quad w \longmapsto b(w, w) \ ,$$

eine quadratische Form und heißt die durch b definierte quadratische Form.
Für $u, w \in V$ ist dann

$$b(u, w) = \frac{1}{2} (q_b(u + w) - q_b(u) - q_b(w)) \ .$$

Wenn $A \in K^{n\times n}$ die Matrix von b bezüglich $\underline{v}$ ist, dann ist

$$q_b = \sum_{i=1}^{n} A_{ii}X_i^2 + \sum_{1\leq i<j\leq n} 2A_{ij}X_iX_j \ .$$

Für jede n-Spalte $y \in K^{n\times 1}$ ist

$$q_b(\underline{v}\cdot y) = y^\top \cdot A \cdot y \ .$$

(„Jeder quadratischen Form entspricht genau eine symmetrische Bilinearform und, nach Wahl einer Basis von V, genau eine symmetrische Matrix, und umgekehrt.“)

Beweis:

(1) Die Abbildung b_q ist offensichtlich symmetrisch, also genügt es zu zeigen, dass für alle $c \in K$, $u, u', w \in V$

$$b_q(c(u+u'), w) = cb_q(u,w) + cb_q(u',w)$$

ist. Mit $q = \sum_{1\leq i\leq j\leq n} a_{ij}X_iX_j$ kann das leicht nachgerechnet werden.

(2) Sei A die Matrix von b bezüglich $\underline{v}$. Für alle n-Spalten y ist dann

$$q_b(\underline{v}\cdot y) = b(\underline{v}\cdot y, \underline{v}\cdot y) = y^\top \cdot A \cdot y = \sum_{1\leq i\leq j\leq n} A_{ij}y_iy_j =$$

$$= \sum_{1\leq i\leq j\leq n} A_{ij}X_i(\underline{v}\cdot y)X_j(\underline{v}\cdot y) = (\sum_{1\leq i\leq j\leq n} A_{ij}X_iX_j)(\underline{v}\cdot y) \ .$$

Daher ist $q_b = \sum_{1\leq i\leq j\leq n} A_{ij}X_iX_j \in Q(V)$.

Beispiel 159: Sei $K = \mathbb{R}$ und $\langle -,- \rangle : V \times V \longrightarrow \mathbb{R}$ ein Skalarprodukt. Die durch $\langle -,- \rangle$ definierte quadratische Form ist

$$\|.\|^2 : V \longrightarrow \mathbb{R} \quad , \quad w \longmapsto \langle w, w \rangle = \|w\|^2 \ ,$$

und ihre Matrix bezüglich $\underline{v}$ ist

$$\big(\langle v_i, v_j \rangle\big)_{1\leq i,j\leq n} \ .$$

Satz 160: *Es seien q eine quadratische Form auf V, A ihre Matrix bezüglich $\underline{v}$ und r der Rang von A. Seien $P \in GL_n(K)$ und $d_1, \ldots, d_r \in K$ so, dass*

$$P^\top \cdot A \cdot P = Diag(d_1, \ldots, d_r, 0, \ldots, 0)$$

(siehe Satz 138).
Sei g die bijektive lineare Abbildung von V nach V, deren Matrix bezüglich $\underline{v}$ die Matrix P ist. Dann ist

$$q \circ g = d_1X_1^2 + \ldots + d_rX_r^2 \ .$$

Beweis: Für alle n-Spalten $y \in K^{n\times 1}$ ist $(q\circ g)(\underline{v}\cdot y) =$

$$= q(g(\underline{v}\cdot y)) = q(\underline{v}\cdot(Py)) = (Py)^\top \cdot A\cdot(Py) = y^\top\cdot P^\top\cdot A\cdot P\cdot y =$$

$$= y^\top\cdot Diag(d_1,\ldots,d_r,0,\ldots,0)\cdot y = \sum_{i=1}^{r} d_i y_i^2 = \left(\sum_{i=1}^{r} d_i X_i^2\right)(\underline{v}\cdot y)\ .$$

§3. Quadratische Funktionen

Es seien $\underline{v}$ eine Basis von V und $\underline{X}$ die dazu duale Basis von V^*.

Definition 161 : Es seien $c\in K$, $\ell\in V^*$, $q\in Q(V)$ und $q\neq 0$.
Dann ist die Abbildung $f := q+\ell+c\cdot\mathbf{1}$ von V nach K eine *quadratische Funktion auf* V. Oft wird statt $c\cdot\mathbf{1}$ nur c geschrieben.
Die Abbildungen q bzw. ℓ bzw. c heißen *quadratischer* bzw. *linearer* bzw. *konstanter Anteil* von f.
Die Nullstellenmenge einer quadratischen Funktion auf V heißt *Quadrik in* V.

Beispiel 162 : Es seien V ein zweidimensionaler euklidischer Raum, $\underline{v}$ eine ON-Basis, $u\in V$ und r eine positive reelle Zahl. Der Kreis um u mit Radius r ist eine Quadrik, denn

$$\{w\in V \mid \|w-u\| = r\} = \{w\in V\mid \langle w-u, w-u\rangle = r^2\} =$$

$$= \{w\in V\mid X_1^2(w-u)+X_2^2(w-u) = r^2\} =$$

$$= \mathcal{N}(X_1^2+X_2^2-2X_1(u)X_1-2X_2(u)X_2+X_1(u)^2+X_2(u)^2-r^2)\ .$$

Satz 163 : *Der quadratische bzw. lineare bzw. konstante Anteil einer quadratischen Funktion ist durch diese eindeutig bestimmt.*

Beweis: Seien $q,q'\in Q(V)$, $\ell,\ell'\in V^*$ und $c,c'\in K$ so, dass $q+\ell+c = q'+\ell'+c'$. Dann ist

$$c = (q+\ell+c)(0) = (q'+\ell'+c')(0) = c'\ ,$$

also $c = c'$ und

$$q+\ell = q'+\ell'\ .$$

Für alle $w\in V$ ist

$$2^2q(w)+2\ell(w) = (q+\ell)(2w) = (q'+\ell')(2w) = 2^2q'(w)+2\ell'(w)\ .$$

Wegen $2 = 1+1\neq 0$ ist daher

$$2q+\ell = 2q'+\ell'\ .$$

Somit ist $\ell = \ell'$ und $q = q'$.

Satz 164 : *(Affine Normalformen quadratischer Funktionen)*
Es seien f eine quadratische Funktion auf V, q ihr quadratischer Anteil und r der Rang von q. Dann gibt es eine bijektive affine Abbildung $h : V \longrightarrow V$ und $d_0 \in K$, $d_1, \ldots, d_r \in K \setminus \{0\}$ so, dass $f \circ h$ eine der folgenden quadratischen Funktionen ist:

$$d_1X_1^2 + \ldots + d_rX_r^2 + d_0 \ ,$$

$$d_1X_1^2 + \ldots + d_rX_r^2 - 2X_n \quad (\textit{nur wenn } r < n).$$

Wenn $K = \mathbb{C}$ ist, kann $d_1 = \ldots = d_r = 1$ gewählt werden.
Wenn $K = \mathbb{R}$ und $(s, r-s)$ die Signatur von q ist, kann
$d_1 = \ldots = d_s = 1$ und $d_{s+1} = \ldots = d_r = -1$ gewählt werden.
Diese quadratischen Funktionen heißen quadratische Funktionen in affiner Normalform bezüglich $\underline{v}$.

Beweis: Die Abbildung h kann wie folgt berechnet werden:
Wähle $g \in GL_K(V)$ so, dass

$$q \circ g = d_1X_1^2 + \ldots + d_rX_r^2$$

(siehe Satz 160).
Es seien $\ell = \sum_{i=1}^n a_iX_i$ der lineare und c der konstante Anteil von $f \circ g$. Mit $u := \sum_{i=1}^r (-\frac{1}{2}a_id_i^{-1})v_i$ ist

$$f \circ g \circ t_u = \sum_{i=1}^r d_i(X_i - \frac{1}{2}a_id_i^{-1})^2 + \sum_{i=1}^r a_i(X_i - \frac{1}{2}a_id_i^{-1}) + \sum_{i=r+1}^n a_iX_i + c =$$

$$= \sum_{i=1}^r d_iX_i^2 + \sum_{i=r+1}^n a_iX_i + (c - \sum_{i=1}^r \frac{1}{4}a_i^2d_i^{-1}) \ .$$

Falls $\sum_{i=r+1}^n a_iX_i = 0$ ist, dann ist $h := g \circ t_u$ die gesuchte Abbildung.
Falls $\sum_{i=r+1}^n a_iX_i \neq 0$ ist, dann ist $r < n$ und wir können $g' \in GL_K(V)$ so wählen, dass

$$X_1 \circ g' = X_1, \ldots, X_r \circ g' = X_r \quad \text{und} \quad \left(\sum_{i=r+1}^n a_iX_i\right) \circ g' = -2X_n \ .$$

Dann ist

$$f \circ g \circ t_u \circ g' = \sum_{i=1}^r d_iX_i^2 - 2X_n + c' \ ,$$

wobei $c' := c - \sum_{i=1}^r \frac{1}{4}a_i^2d_i^{-1}$.
Sei $u' := -\frac{1}{2}c'v_n$, dann ist

$$f \circ g \circ t_u \circ g' \circ t_{u'} = \sum_{i=1}^r d_iX_i^2 - 2X_n$$

und $h := g \circ t_u \circ g' \circ t_{u'}$.

Beispiel 165 : Es seien $n=2$, $K=\mathbb{Q}$ und

$$f := X_1^2 + X_2^2 + 2X_1X_2 + 2X_1 + 3X_2 + 4 \ .$$

Es sei $g : V \longrightarrow V$ die durch $g(v_1) = v_1$ und $g(v_2) = -v_1 + v_2$ definierte lineare Abbildung. Dann ist

$$f \circ g = X_1^2 + 2X_1 + X_2 + 4$$

und

$$f \circ g \circ t_{-v_1} = X_1^2 + X_2 + 3 \ .$$

Es sei $g' : V \longrightarrow V$ die durch $g'(v_1) = v_1$ und $g'(v_2) = -2v_2$ definierte lineare Abbildung. Dann ist

$$f \circ g \circ t_{-v_1} \circ g' = X_1^2 - 2X_2 + 3$$

und

$$f \circ g \circ t_{-v_1} \circ g' \circ t_{\frac{3}{2}v_2} = X_1^2 - 2X_2 \ .$$

Sei $h := g \circ t_{-v_1} \circ g' \circ t_{\frac{3}{2}v_2}$. Dann ist

$$h(y_1v_1 + y_2v_2) = (y_1 + 2y_2 + 2)v_1 - (2y_2 + 3)v_2$$

und $h(\mathcal{N}(X_1^2 - 2X_2)) = \mathcal{N}(f)$. Daher ist

$$\mathcal{N}(f) = \{(y_1 + y_1^2 + 2)v_1 - (y_1^2 + 3)v_2 \mid y_1 \in \mathbb{Q}\} \ .$$

Für $y_1 = 0$ bzw. 1 bzw. -1 erhalten wir $2v_1 - 3v_2$ bzw. $4v_1 - 4v_2$ bzw. $2v_1 - 4v_2 \in \mathcal{N}(f)$.

Beispiel 166 : Es seien $n = 1$, $a, b \in K$ und

$$f := X_1^2 + aX_1 + b \ .$$

Dann ist

$$f \circ t_{-\frac{a}{2}} = X_1^2 - \frac{a^2}{4} + b$$

und

$$\mathcal{N}(f) = t_{-\frac{a}{2}}\left(\mathcal{N}(X_1^2 - (\frac{a^2}{4} - b))\right) \ .$$

Daher ist $\mathcal{N}(f)$ genau dann nicht leer, wenn es in K ein Element y gibt mit

$$y^2 = \frac{a^2}{4} - b \ .$$

Für $K = \mathbb{R}$ ist das genau dann der Fall, wenn $\frac{a^2}{4} - b \geq 0$.
Wenn ein solches $y \in K$ existiert, ist $\mathcal{N}(f \circ t_{-\frac{a}{2}}) = \{y, -y\}$ und

$$\mathcal{N}(f) = \{-\frac{a}{2} + y, -\frac{a}{2} - y\} \ .$$

§4. Quadriken

Es seien $\underline{v}$ eine Basis von V und $\underline{X}$ die dazu duale Basis von V^*.

Hilfssatz 167 : *Es sei W ein endlich-dimensionaler K-Vektorraum.*

(1) *Die Zusammensetzung $f \circ h$ einer quadratischen Funktion f auf V mit einer surjektiven affinen Abbildung $h : W \longrightarrow V$ ist eine quadratische Funktion auf W.*
(2) *Das Urbild $h^{-1}(Q)$ einer Quadrik Q in V unter einer surjektiven affinen Abbildung $h : W \longrightarrow V$ ist eine Quadrik.*
(3) *Das Bild $h(Q)$ einer Quadrik Q unter einer bijektiven affinen Abbildung $h : W \longrightarrow V$ ist eine Quadrik.*
(4) *Für eine quadratische Funktion f auf V und $c \in K \setminus \{0\}$ ist $\mathcal{N}(f) = \mathcal{N}(cf)$.*

Beweis:

(1) Es sei q der quadratische Anteil von f. Weil $q \neq 0$ und weil h surjektiv ist, ist auch $q \circ h \neq 0$ und $f \circ h \neq 0$. Es genügt also zu zeigen, dass $q \circ h$ eine quadratische Funktion ist. Wenn g_1, g_2 Linearformen auf V sind, dann ist $(g_1 g_2) \circ h = (g_1 \circ h)(g_2 \circ h)$. Die Abbildungen $g_1 \circ h$ und $g_2 \circ h$ sind affin, also ist $(g_1 \circ h)(g_2 \circ h)$ eine quadratische Funktion. Weil q eine Linearkombination von Produkten von Linearformen ist, folgt daraus die Behauptung.
(2) Es sei f eine quadratische Funktion mit $\mathcal{N}(f) = Q$. Dann ist $h^{-1}(Q) = \mathcal{N}(f \circ h)$, also folgt die Behauptung aus (1).
(3) Folgt aus (2).
(4) Wegen $c \neq 0$ ist $f(v) = 0$ genau dann, wenn $cf(v) = 0$.

Definition 168 : Zwei Teilmengen M und N von V heißen *affin kongruent*, wenn es eine bijektive affine Abbildung $h : V \longrightarrow V$ gibt, sodass $h(M) = N$ ist.
Zwei Teilmengen M und N eines euklidischen Raumes V heißen *euklidisch kongruent*, wenn es eine Isometrie $h : V \longrightarrow V$ gibt, sodass $h(M) = N$ ist.

Beispiel 169 : Je zwei Dreiecke in V sind affin kongruent.
Zwei Dreiecke in einem euklidischen Raum sind genau dann euklidisch kongruent, wenn die zwei Tripel der Seitenlängen der Dreiecke bis auf die Reihenfolge gleich sind.

Satz 170 : *(Affine Normalformen von Quadriken)*
Es seien $K = \mathbb{R}$, f eine quadratische Funktion auf V, r der Rang und $(s, r -$

s) die Signatur des quadratischen Anteils von f. Die Quadrik $\mathcal{N}(f)$ ist zu einer der folgenden Quadriken affin kongruent:

$$\mathcal{N}(X_1^2+\ldots+X_s^2-X_{s+1}^2-\ldots-X_r^2)\ ,$$
$$\mathcal{N}(X_1^2+\ldots+X_s^2-X_{s+1}^2-\ldots-X_r^2+1)\ ,$$
$$\mathcal{N}(X_1^2+\ldots+X_s^2-X_{s+1}^2-\ldots-X_r^2-1)\ ,$$
$$\mathcal{N}(X_1^2+\ldots+X_s^2-X_{s+1}^2-\ldots-X_r^2-2X_n)\ (\textit{nur wenn } r<n).$$

Beweis: Folgt aus Satz 164 und Lemma 167,(4).

Beispiel 171 : Es seien d eine negative reelle Zahl und $h := \frac{\sqrt{|d|}}{|d|}\cdot \mathrm{Id}_V$. Dann ist

$$h\left(\mathcal{N}(X_1^2+\ldots+X_s^2-X_{s+1}^2-\ldots-X_r^2+d)\right)=$$
$$=\mathcal{N}(X_1^2+\ldots+X_s^2-X_{s+1}^2-\ldots-X_r^2-1)\ .$$

§5. Quadriken in der Ebene

Es seien V ein zweidimensionaler euklidischer Raum, $\underline{v}$ eine ON-Basis von V und $\underline{X}$ die dazu duale Basis von V^*.

Definition 172 : Es seien $u,w\in V$ und $r\in\mathbb{R}$ so, dass $\|u-w\|<2r$ ist. Dann heißt die Menge

$$E(u,w,r):=\{z\in V\mid \|z-u\|+\|z-w\|=2r\}$$

Ellipse mit den *Brennpunkten* u und w.

Fixiert man die Enden eines Fadens der Länge $2r$ in den Punkten u und w und durchläuft man dann mit einem Bleistift alle Punkte so, dass der Faden gespannt ist, dann erhält man die Ellipse $E(u,w,r)$ („Gärtner-Ellipse“ oder „Fadenkonstruktion“ der Ellipse).

Beispiel 173 : $E(u,u,r)$ ist der Kreis mit Mittelpunkt u und Radius r.

Satz 174 : *Es seien $u,w\in V$, $c:=\|\frac{u-w}{2}\|$ und $r\in\mathbb{R}$ mit $r>c$.*

(1) *Die Ellipse $E(u,w,r)$ ist euklidisch kongruent zu $E(cv_1,-cv_1,r)$. Eine Ellipse $E(u',w',r')$ ist genau dann euklidisch kongruent zu $E(u,w,r)$, wenn $r=r'$ und $\|u-w\|=\|u'-w'\|$ ist.*

(2)
$$E(cv_1,-cv_1,r)=\mathcal{N}(\frac{1}{r^2}X_1^2+\frac{1}{(r^2-c^2)}X_2^2-1)\ .$$

(3) *Die Ellipse $E(u,w,r)$ ist affin kongruent zu $\mathcal{N}(X_1^2+X_2^2-1)$. Insbesondere ist jede Ellipse eine Quadrik und je zwei Ellipsen sind zueinander affin kongruent.*

Beweis:

(1) Ist f eine Isometrie, dann ist $f(E(u,w,r))=E(f(u),f(w),r)$. Daher ist
$$t_{-\frac{u+w}{2}}(E(u,w,r))=E(\frac{u-w}{2},-\frac{u-w}{2},r)\ .$$
Es sei h die Drehung um 0, die den Punkt $\frac{u-w}{2}$ auf den Punkt cv_1 abbildet. Dann ist
$$(h\circ t_{-\frac{u+w}{2}})(E(u,w,r))=E(cv_1,-cv_1,r)\ ,$$
insbesondere ist $E(u,w,r)$ euklidisch kongruent zu $E(cv_1,-cv_1,r)$.

(2) Für $z\in V$ ist $z\in E(cv_1,-cv_1,r)$ genau dann, wenn
$$\|z-cv_1\|+\|z+cv_1\|=2r\ .$$
Man rechnet leicht nach, dass dies genau dann der Fall ist, wenn
$$(r^2-c^2)X_1^2(z)+r^2X_2^2(z)-r^2(r^2-c^2)=0$$
ist. Wegen $0\le c<r$ und $r^2>0$ ist $r^2-c^2>0$ und $r^2(r^2-c^2)>0$. Daher ist
$$E(cv_1,-cv_1,r)=\mathcal{N}((r^2-c^2)X_1^2+r^2X_2^2-r^2(r^2-c^2))=$$
$$=\mathcal{N}(\frac{1}{r^2}X_1^2+\frac{1}{r^2-c^2}X_2^2-1)\ .$$

(3) Es sei $g:V\longrightarrow V$ die lineare Abbildung mit $g(v_1)=\frac{1}{r}v_1$ und $g(v_2)=\frac{1}{\sqrt{r^2-c^2}}v_2$.
Dann ist
$$(g\circ h\circ t_{-\frac{u+w}{2}})(E(u,w,r))=g(E(cv_1,-cv_1,r))=$$
$$=\mathcal{N}(X_1^2+X_2^2-1)\ ,$$
somit ist $E(u,w,r)$ affin kongruent zu $\mathcal{N}(X_1^2+X_2^2-1)$.

Definition 175 : Es seien $u,w\in V$ und $r\in\mathbb{R}$ so, dass $0<2r<\|u-w\|$ ist. Dann heißt die Menge
$$H(u,w,r):=\{z\in V\mid 2r=|\,\|z-u\|-\|z-w\|\,|\}$$
Hyperbel mit den *Brennpunkten* u und w.

Satz 176 : *Es seien $u, w \in V$, $c := \|\frac{u-w}{2}\|$ und $r \in \mathbb{R}$ mit $0 < r < c$.*

(1) *Die Hyperbel $H(u,w,r)$ ist euklidisch kongruent zu $H(cv_1, -cv_1, r)$. Eine Hyperbel $H(u',w',r')$ ist genau dann euklidisch kongruent zu $H(u,w,r)$, wenn $r = r'$ und $\|u-w\| = \|u'-w'\|$ sind.*

(2)
$$H(cv_1, -cv_1, r) = \mathcal{N}(\frac{1}{r^2}X_1^2 - \frac{1}{(c^2-r^2)}X_2^2 - 1) .$$

(3) *Die Hyperbel $H(u,w,r)$ ist affin kongruent zu $\mathcal{N}(X_1^2 - X_2^2 - 1)$. Insbesondere ist jede Hyperbel eine Quadrik und je zwei Hyperbeln sind zueinander affin kongruent.*

Beweis: Übung (analog dem Beweis von Satz 174).

Definition 177 : Es seien $u \in V$ und G eine Gerade in V, die nicht u enthält. Dann heißt die Menge $P(u,G)$ aller Punkte $z \in V$, die zu u und zu G den gleichen Abstand haben, *Parabel* mit *Brennpunkt u* und *Leitlinie G*.

Satz 178 : *Es seien $u \in V$, G eine Gerade in V, die nicht u enthält, und c der Abstand des Punktes u von der Geraden G.*

(1) *Die Parabel $P(u,G)$ ist zu $P(\frac{c}{2}v_2, -\frac{c}{2}v_2 + \mathbb{R}v_1)$ euklidisch kongruent. Eine Parabel $P(u',G')$ ist genau dann euklidisch kongruent zu $P(u,G)$, wenn die Abstände von u' zu G' und von u zu G gleich sind.*

(2)
$$P(\frac{c}{2}v_2, -\frac{c}{2}v_2 + \mathbb{R}v_1) = \mathcal{N}(\frac{1}{c}X_1^2 - 2X_2) .$$

(3) *Die Parabel $P(u,G)$ ist affin kongruent zu $\mathcal{N}(X_1^2 - 2X_2)$. Insbesondere ist jede Parabel eine Quadrik und je zwei Parabeln sind zueinander affin kongruent.*

Beweis: Übung (analog dem Beweis von Satz 174).

Satz 179 :

(1) *Jede Quadrik in V ist euklidisch kongruent zu einer der folgenden Quadriken:*

Quadrik	*geometrische Beschreibung*
$\mathcal{N}(X_1^2)$	*eine Gerade*
$\mathcal{N}(X_1^2 - \frac{1}{b^2}X_2^2),\ b > 0$	*zwei einander schneidende Geraden*
$\mathcal{N}(X_1^2 + X_2^2)$	*ein Punkt*
$\mathcal{N}(\frac{1}{a^2}X_1^2 - 1),\ a > 0$	*zwei parallele Geraden mit Abstand $2a$*
$\mathcal{N}(X_1^2 + X_2^2 + 1)$	*die leere Menge*
$\mathcal{N}(\frac{1}{a^2}X_1^2 + \frac{1}{b^2}X_2^2 - 1)$, $0 < b \leq a$	*eine Ellipse mit Achsenabschnitten a und b*
$\mathcal{N}(\frac{1}{a^2}X_1^2 - \frac{1}{b^2}X_2^2 - 1)$, $a > 0, b > 0$	*eine Hyperbel mit Achsenabschnitten a und b*
$\mathcal{N}(\frac{1}{a^2}X_1^2 - 2X_2),\ a > 0$	*eine Parabel mit Parameter a^2*

(2) *Jede Quadrik ist affin kongruent zu einer der Quadriken in (1) mit $a = b = 1$.*

Beweis:

(1) Beweis wird weggelassen.
(2) Folgt aus Satz 170.

Satz 180 :

(1) *Es seien Q eine Ellipse in V mit Brennpunkten u und w. Dann gibt es für jeden Punkt $z \in Q$ genau eine Gerade T_z in V, deren Durchschnitt mit Q gleich $\{z\}$ ist. Diese Gerade heißt* Tangente *an Q im Punkt z. Die Winkel zwischen der Tangente T_z und der Geraden durch z und u sowie der Geraden durch z und w sind gleich.*

(2) *Es sei Q eine Parabel in V mit Brennpunkt u und Leitlinie L. Dann gibt es für jeden Punkt $z \in Q$ genau eine Gerade T_z in V, die nicht senkrecht zur Leitlinie von Q steht und deren Durchschnitt mit Q gleich $\{z\}$ ist. Diese Gerade heißt* Tangente *an Q im Punkt z. Die Winkel zwischen der Tangente T_z und der Geraden durch z und u sowie der zu L senkrecht stehenden Geraden durch z sind gleich.*

Beweis:

(1) Sei $f : V \longrightarrow V$ eine bijektive affine Abbildung. Wenn G eine Gerade in V ist, deren Durchschnitt mit $f(Q)$ gleich $\{f(z)\}$ ist, dann ist der Durchschnitt von $f^{-1}(Q)$ mit Q gleich $\{z\}$. Nach Satz 174 genügt es daher, die Aussage für

$$Q = \mathcal{N}(X_1^2 + X_2^2 - 1)$$

zu beweisen. In diesem Fall kann leicht nachgeprüft werden, dass die Gerade

$$T_z := \mathcal{N}(X_1(z)X_1 + X_2(z)X_2 - 1)$$

die angegebenen Eigenschaften hat.

Nach Satz 174 können wir annehmen, dass

$$Q = E(cv_1, -cv_1, r)$$

ist, wobei $r > c > 0$ ist. Sei

$$y := \frac{1}{(r^2 - c^2)} z_2 v_1 - \frac{1}{r^2} z_1 v_2 \ .$$

Dann ist $z + \mathbb{R}y$ die Tangente an Q im Punkt z. Seien α bzw. β die Winkel zwischen T_z und der Geraden durch z und cv_1 bzw. $-cv_1$. Dann ist

$$\cos(\alpha) = \frac{\langle z - cv_1, y\rangle}{\|z - cv_1\| \cdot \|y\|} \text{ und } \cos(\beta) = \frac{\langle z + cv_1, y\rangle}{\|z + cv_1\| \cdot \|-y\|} \ .$$

Da α und β im Intervall $[0, \pi]$ liegen, genügt es nachzurechnen, dass $\cos(\alpha) = \cos(\beta)$ ist. Mit $z_2^2 = \frac{1}{(r^2 - c^2)}(r^2 - z_1^2)$ erhält man

$$\|z \pm cv_1\| = \sqrt{(z_1 \pm c)^2 + z_2^2} = \frac{r^2 \pm cz_1}{r} \ .$$

Weiters ist

$$\langle z - cv_1, y\rangle = \frac{cz_2}{r^2(r^2 - c^2)}(cz_1 - r^2)$$

und
$$\langle z + cv_1, -y \rangle = \frac{-cz_2}{r^2(r^2 - c^2)}(cz_1 + r^2) \ ,$$
daraus folgt $\cos(\alpha) = \cos(\beta)$.

(2) Übung.

Aus diesem Satz folgt: Ein Lichtstrahl, der von einem Brennpunkt einer Ellipse ausgeht und an der Ellipse reflektiert wird, geht durch den anderen Brennpunkt. Lichtstrahlen, die vom Brennpunkt einer Parabel ausgehen und an der Parabel reflektiert werden, sind dann zueinander parallel.