Lineare Algebra 2

Ein Skriptum zur Vorlesung
im Sommersemester 2010

Arne Dür und Franz Pauer

3.Auflage

Vorwort

Das vorliegende Skriptum soll den Hörerinnen und Hörern der Vorlesung „Lineare Algebra 2“ im Sommersemester 2010 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 „Einführung in die Mathematik 1“ von Franz Pauer wird als bekannt vorausgesetzt.

In dieser Vorlesung werden einerseits Ergebnisse der Vorlesung „Einführung in die Mathematik 1“ über Vektorräume, lineare Funktionen, Eigenwertprobleme, Skalarprodukte und multilineare Funktionen ergänzt und erweitert (Kapitel 1, 2, 4 und 6), andererseits eine Einführung in die *Analytische Geometrie* gegeben (Kapitel 3 und 5). In der Analytischen Geometrie werden geometrische Fragestellungen mit Methoden der linearen Algebra behandelt. Im Kapitel 5 werden der Bewegungen starrer Körper in der Ebene und im Raum beschrieben, dieses Thema ist für die Mechanik von großer Bedeutung .

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.

Die dritte Auflage unterscheidet sich von der zweiten (Februar 2009) durch eine Umstellung der Reihenfolge der Kapitel und einige Korrekturen und Ergänzungen.
Martin Huber danken wir für die Anregung, den Abschnitt 5 in Kapitel 5 (Zeigerrechnung) in das Skriptum aufzunehmen, sowie für die Erlaubnis, dafür seine Materialien in www.tech4math.com zu verwenden. Die Zeichnungen dazu hat Simone Graml angefertigt.
Hubert Herdlinger verdanken wir eine Verbesserung des Abschnittes über Systeme linearer Ungleichungen.

Innsbruck, Februar 2010

Inhaltsverzeichnis

KAPITEL 1

Rechnen mit Koordinaten, koordinatenfreie Modellierung

In diesem Kapitel sei K ein Körper.

§1. Erzeugendensysteme, lineare Unabhängigkeit und Basen

In der Vorlesung „Einführung in die Mathematik 1“ wurde der Begriff *Linearkombination* eines n-Tupels von Vektoren eingeführt. Wir erweitern diesen Begriff nun auf beliebige (auch unendliche) Familien $(v_i)_{i\in I}$ von Vektoren. Die Definition im Wintersemester entspricht dann dem Spezialfall $I := \{1,\ldots,n\}$.

Definition 1 : Sei V ein Vektorraum über K und I eine (beliebige) Menge.
Eine Familie $(c_i)_{i\in I}$ von Elementen in K heißt *Koeffizientenfamilie*, wenn $c_i \neq 0$ für nur endlich viele $i \in I$ ist.
Sei $(v_i)_{i\in I}$ eine Familie von Vektoren in V.
Ein Vektor $w \in V$ heißt eine *Linearkombination* von $(v_i)_{i\in I}$, wenn es eine Koeffizientenfamilie $(c_i)_{i\in I}$ gibt, sodass

$$w = \sum_{i\in I, c_i\neq 0} c_i v_i$$

ist. Wir schreiben im weiteren einfach

$$\sum_{i\in I} c_i v_i \qquad \text{anstatt} \qquad \sum_{i\in I, c_i\neq 0} c_i v_i\,.$$

Die Menge aller Linearkombinationen von $(v_i)_{i\in I}$ ist ein Untervektorraum von V und enthält alle Vektoren v_i, $i \in I$. Er heißt der *von v_i, $i \in I$, erzeugte Untervektorraum von V* und wird mit

$$_K\langle v_i \mid i \in I\rangle \qquad \text{oder} \qquad \sum_{i\in I} K v_i$$

bezeichnet.

Definition 2 : Sei $V \neq \{0\}$ ein Vektorraum über K. Eine Familie $(v_i)_{i\in I}$ von Vektoren in V heißt genau dann ein *Erzeugendensystem* von V bzw. *linear unabhängig* in V bzw. eine *Basis* von V, wenn jeder Vektor in V auf mindestens eine bzw. höchstens eine bzw. genau eine Weise als Linearkombination von $(v_i)_{i\in I}$ geschrieben werden kann.
Wir schreiben *linear abhängig* anstatt *nicht linear unabhängig*.
Die leere Menge ist eine *Basis von* $\{0\}$.

Satz 3: *Sei $V \neq \{0\}$ ein Vektorraum über K und $(v_i)_{i\in I}$ eine Familie von Vektoren in V.*

(1) *Eine Basis von V ist ein linear unabhängiges Erzeugendensystem.*
(2) *Die Familie $(v_i)_{i\in I}$ von Vektoren ist genau dann ein Erzeugendensystem von V, wenn*
$$_K\langle v_i \mid i \in I\rangle = V$$
ist.
(3) *Die Familie $(v_i)_{i\in I}$ von Vektoren ist genau dann linear unabhängig, wenn für jede Koeffizientenfamilie $(c_i)_{i\in I}$ aus*
$$\sum_{i\in I} c_i v_i = 0$$
folgt, dass
$$c_i = 0 \quad \textit{für alle } i \in I$$
ist.

Beweis:

(1) und (2) folgen aus der Definition der Begriffe Erzeugendensystem, linear unabhängig und Basis.
(3) Wenn sich jeder Vektor aus V auf höchstens eine Weise als Linearkombination von $(v_i)_{i\in I}$ schreiben lässt, dann folgt aus
$$\sum_{i\in I} c_i v_i = 0_V = \sum_{i\in I} 0_K v_i$$
auf Grund der Eindeutigkeit $c_i = 0_K$ für $1 \leq i \leq n$.
Sei umgekehrt $(v_i)_{i\in I}$ linear unabhängig in V und $w \in V$ so, dass es eine Koeffizientenfamilie $(c_i)_{i\in I}$ mit
$$w = \sum_{i\in I} c_i v_i$$
gibt. Falls $(d_i)_{i\in I}$ eine weitere Koeffizientenfamilie mit
$$w = \sum_{i\in I} d_i v_i$$
ist, erhält man
$$0_V = w - w = \sum_{i\in I} c_i v_i - \sum_{i\in I} d_i v_i = \sum_{i\in I} (c_i - d_i) v_i\,.$$
Nach Annahme folgt $c_i - d_i = 0_K$ für alle $i \in I$, also $c_i = d_i$ für alle $i \in I$.

Beispiel 4: Die Familie $(x^i)_{i\in\mathbb{N}}$ ist eine K-Basis des Polynomringes $K[x]$ in der Variablen x mit Koeffizienten in K.

Beispiel 5 : Die Familie

$$(E_{k\ell})_{\substack{1\le k\le m\\ 1\le \ell\le n}} = (E_{k\ell})_{(k,\ell)\in\{1,\dots,m\}\times\{1,\dots,n\}}$$

der Standard-Matrizen ist eine Basis von $K^{m\times n}$ und heißt die *Standardbasis* von $K^{m\times n}$.

Für $A \in K^{m\times n}$ ist

$$A = \sum_{\substack{1\le k\le m\\ 1\le \ell\le n}} A_{k\ell}E_{k\ell}\,,$$

also ist A die Koordinatenfamilie von A bezüglich der Standardbasis $(E_{k\ell})_{\substack{1\le k\le m\\ 1\le \ell\le n}}$ (und $A_{k\ell}$ die Koordinate von A bei $E_{k\ell}$).

Beispiel 6 : Es seien I eine endliche Menge und V die Menge aller Funktionen von I in einen Körper K. Für $i \in I$ sei δ_i die durch

$$\delta_i(j) := \delta_{ij},\ j \in I,$$

definierte Funktion von I nach K. Dann ist die Familie

$$(\delta_i)_{i\in I}$$

eine K-Basis von V. Für $f \in V$ ist

$$f = \sum_{i\in I} f(i)\delta_i\,.$$

§2. Rechnen mit Koordinaten

In diesem Abschnitt seien V und W Vektorräume über K, $n \in \mathbb{N}$ die Dimension von V und $\underline{v} := (v_1,\dots,v_n)$ eine Basis von V.

Definition 7 : Seien p und q positive ganze Zahlen und

$$\underline{u} := (u_1,\dots,u_p) \in V^p$$

ein p-Tupel von Vektoren in V. Für $T \in K^{p\times q}$ sei

$$\underline{u}T := \big((\underline{u}T)_1,\dots,(\underline{u}T)_q\big) := \left(\sum_{i=1}^{p} T_{i1}u_i,\dots,\sum_{i=1}^{p} T_{iq}u_i\right) \in V^q.$$

Die Bezeichnung $\underline{u}T$ ist eine Merkhilfe, weil man analog zur Matrizenrechnung über K den Vektor $(\underline{u}T)_j \in V$ nach der Regel

„Zeile $\underline{u}$ mal Spalte T_{-j}“

berechnet.

Satz 8 : *Seien p,q,r positive ganze Zahlen und $\underline{u} := (u_1,\dots,u_p) \in V^p$. Dann gilt:*

(1) $\underline{u}I_p = \underline{u}$

(2) *Für* $T \in K^{p\times q}$ *und* $U \in K^{q\times r}$ *ist* $\underline{u}(TU) = (\underline{u}T)U$.
(3) *Für* $T \in \mathrm{GL}_p(K)$ *und* $\underline{u}' := \underline{u}T$ *gilt* $\underline{u} = \underline{u}'T^{-1}$.

Beweis: (1) gilt nach Definition, (2) rechnet man nach und (3) folgt aus $\underline{u} = \underline{u}I_p = \underline{u}(TT^{-1}) = (\underline{u}T)T^{-1} = \underline{u}'T^{-1}$.

Definition 9 : Seien $w \in V$ und $c_1,\dots,c_n \in K$ die Koordinaten von w bezüglich $\underline{v}$, also

$$w = \sum_{i=1}^{n} c_i v_i = \underline{v}c \in V\,,$$

Dann heißt die Spalte

$$\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in K^{n\times 1}$$

die *Koordinatenspalte* von w bezüglich $\underline{v}$.

Satz 10 : *Sei* $\underline{u} := (u_1,\dots,u_q) \in V^q$. *Dann gilt:*

(1) *Es gibt genau eine Matrix* $T \in K^{n\times q}$ *mit*

$$\underline{u} = \underline{v}T\,.$$

Diese Matrix T *heißt die* Transformationsmatrix *von* $\underline{v}$ *zu* $\underline{u}$, *und die Spalten von* T *sind die Koordinatenspalten von* $u_1,\dots,u_q$ *bezüglich* $\underline{v}$.
(2) *Die Familie* $\underline{u}$ *ist genau dann eine Basis von* V, *wenn* T *invertierbar ist.*

Beweis:

(1) Sei $T \in K^{n\times k}$ jene Matrix, die sich durch Nebeneinanderschreiben der Koordinatenspalten von $u_1,\dots,u_k$ bezüglich $\underline{v}$ ergibt. Dann ist $u_j = \underline{v}T_{-j}$ für $1 \le j \le k$, also $\underline{u} = \underline{v}T$. Wenn umgekehrt $\underline{u} = \underline{v}S$ mit $S \in K^{n\times k}$ ist, d.h. $u_j = \underline{v}S_{-j}$ für $1 \le j \le k$, dann ist S_{-j} die Koordinatenspalte von u_j bezüglich $\underline{v}$ für $1 \le j \le k$.

(2) Wenn $T \in \mathrm{GL}_n(K)$ ist, dann ist $\underline{v} = \underline{u}T^{-1}$ nach Satz 8, also sind die Vektoren $v_1,\dots,v_n$ Linearkombinationen von $u_1,\dots,u_n$, somit ist ${}_K\langle u_1,\dots,u_n\rangle = V$ und, weil V n-dimensional ist, $(u_1,\dots,u_n)$ eine Basis von V. Wenn umgekehrt $\underline{u}$ eine Basis von V ist, dann gibt es eine Matrix $U \in K^{n\times n}$ mit $\underline{v} = \underline{u}U$. Weil $\underline{u}$ eine Basis ist, folgt aus $\underline{u}I_n = \underline{v}T = \underline{u}(UT)$, dass $UT = I_n$ ist und analog $TU = I_n$, also ist $T \in \mathrm{GL}_n(K)$.

Satz 11 : *Sei $\underline{u}$ eine Basis von V und $T \in \mathrm{GL}_n(K)$ die Transformationsmatrix von $\underline{v}$ zu $\underline{u}$. Ist c die Koordinatenspalte von $w \in V$ bezüglich $\underline{v}$, dann ist*

$$T^{-1}c$$

die Koordinatenspalte von w bezüglich $\underline{u}$, d.h. bei Basiswechsel mit der Matrix T „transformieren sich die Koordinaten" mit der Matrix T^{-1}.

Beweis: Es ist $w = \underline{v}c = (\underline{u}T^{-1})c = \underline{u}(T^{-1}c)$.

§3. Der Graph einer linearen Funktion

Satz 12 : *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

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

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

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

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 $c_{11}=\cdots=c_{\ell n_\ell}=0$, was zu zeigen war.

Satz 13 : *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 $\mathrm{Graph}(f)$.

§4. Basiswechsel

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

Satz 14: *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*

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

Beweis: Es ist

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

und

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

Satz 15: *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 14 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 16:

(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 17:

(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 $B \in K^{m\times n}$ 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 15 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 15 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 18: Sei $A \in K^{n\times n}$. Dann heißt
$$\mathrm{spur}(A) := \sum_{i=1}^{n} A_{ii}$$
die *Spur* von A.

Satz 19 :

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

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

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 20 :

(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).

§5. 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 21 : 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 22 : $\mathrm{Bild}(f)$ *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 23: *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*

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

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 24: *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 $(f(v_1), \ldots, f(v_r))$ von $\mathrm{Bild}(f)$ *zu einer Basis $(w_1, \ldots, w_m)$ 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 $(1,1), \ldots, (r,r)$ stehen Einsen und sonst Nullen) die Matrix von f bezüglich der Basen $(v_1, \ldots, v_n)$ und $(w_1, \ldots, w_m)$.

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

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

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

$$f(y) = \sum_{i=1}^r a_i w_i = \sum_{i=1}^r a_i f(v_i) = f(\sum_{i=1}^r a_i v_i)$$

folgt, dass $z := y - \sum_{i=1}^r a_i v_i \in \mathrm{Kern}(f)$ ist. Daher ist $y = z + \sum_{i=1}^r a_i v_i$ 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

$$\sum_{i=1}^r c_i v_i + \sum_{j=1}^s d_j u_j = 0 \,.$$

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

Satz 25 : *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 26 :

(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 17, Satz 24 und Satz 25.
(2) folgt aus (1), Satz 17 und Satz 25.

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

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

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

Beweis: Nach Satz 26,(2) gibt es Matrizen $P \in \mathrm{GL}_m(K)$ und $Q \in \mathrm{GL}_n(K)$ so, dass $PAQ = D_r$ ist, wobei $r := \mathrm{rg}(A)$ 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 26 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 $n-\mathrm{rg}(A)$ 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 27 ist $\mathrm{rg}(A) = \mathrm{rg}(A^T)$, 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.

§6. Systeme linearer Gleichungen in koordinatenfreier Form

Definition 28 : 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

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

Satz 29 : *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*

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

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

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

Beweis: Sei $v \in \mathrm{Kern}(f)$. 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 $x - z \in \mathrm{Kern}(f)$ und $x = z + (x-z) \in \{z + v \mid v \in \mathrm{Kern}(f)\}$.
Nach Satz 24 ist $\dim_K(\mathrm{Kern}(f)) = \dim_K(V) - \mathrm{rg}(f)$.

Beispiel 30 : 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 31 : 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 $\mathrm{Kern}(D)$ 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

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

Satz 32 : *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 $\mathrm{Kern}(f)$ auf $L(A,0)$.

Beweis: Nach Satz 14 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 32 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 $(u_1,\dots,u_s)$ eine Basis von $L(A,0)$ ist, dann ist $\underline{v}z \in L(f,y)$ und $(\underline{v}u_1,\dots,\underline{v}u_s)$ eine Basis von $\mathrm{Kern}(f)$.

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

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

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

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

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 34 : *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 $(u_1,\ldots,u_k)$ *zu einer Basis* $(u_1,\ldots,u_k,u_{k+1},\ldots,u_n)$ *von* V. *Setze*

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

und $y := f(p)$.

Beweis: Seien f und y wie im Satz definiert. Dann ist $\mathrm{Kern}(f) = \, =_K < u_1,\ldots,u_k >$ und $p \in L(f,y)$. Nach Satz 29 ist $Z = L(f,y)$.

Hilfssatz 35 : *Seien* V *ein Vektorraum über* K *und* U,U' *Untervektorräume von* V. *Dann ist auch* $U \cap U'$ *ein Untervektorraum.*

Beweis: Übung.

Satz 36 : *Seien* V *ein Vektorraum über* K *und* Z,Z' *affine Unterräume von* V *mit* $Z \cap Z' \neq \emptyset$. *Dann ist auch* $Z \cap Z'$ *ein affiner Unterraum, und der parallele Untervektorraum von* $Z \cap Z'$ *ist der Durchschnitt der parallelen Untervektorräume von* Z *und* Z'.

Wenn Z *und* Z' *in impliziter Form gegeben sind, d.h.* $Z = L(f,y)$ *und* $Z' = L(f',y')$, *dann ist* $Z \cap Z' = L(g,(y,y'))$ *mit*

$$g : V \to W \times W' \, , \, v \mapsto (f(v), f'(v)) \, ,$$

d.h. man erhält die Gleichungen für $Z \cap Z'$ *durch „Zusammenschreiben" der Gleichungen für* Z *und für* Z'.

Beweis: Aus $Z = L(f,y)$ und $Z' = L(f',y')$ folgt
$Z \cap Z' = \{v \in V \mid f(v) = y \text{ und } f'(v) = y'\} = L(g,(y,y'))$.
Man prüft leicht nach, dass g K-linear ist. Daher ist $Z \cap Z'$ ein affiner Unterraum von V mit $\mathrm{Kern}(g) = \mathrm{Kern}(f) \cap \mathrm{Kern}(f')$ als parallelem Untervektorraum.

Hilfssatz 37 : *Seien* V *ein Vektorraum über* K *und* $U_1,\ldots,U_\ell$ *Untervektorräume von* V. *Dann ist*

$$U_1 + \cdots + U_\ell := \{u_1 + \cdots + u_\ell \mid u_1 \in U_1,\ldots,u_\ell \in U_\ell\}$$

ein Untervektorraum von V, *enthält* $U_1,\ldots,U_\ell$ *und heißt die* Summe *von* $U_1,\ldots,U_\ell$.

Beweis: Übung.

Satz 38 : *Sei V ein Vektorraum über K und seien U,U' Untervektorräume von V. Dann gilt*

$$\dim_K(U+U') = \dim_K(U) + \dim_K(U') - \dim_K(U \cap U').$$

Beweis: Die Funktion

$$f : U \times U' \to V\,,\ (u,u') \mapsto u+u'\,,$$

ist linear mit $\mathrm{Bild}(f) = U+U'$ und
$\mathrm{Kern}(f) = \{(u,u') \in U \times U' \mid u+u' = 0\} = \{(u,-u) \mid u \in U \cap U'\}$.
Da die Funktion

$$U \cap U' \to \mathrm{Kern}(f)\,,\ u \mapsto (u,-u)\,,$$

ein Isomorphismus ist, gilt $\dim_K(\mathrm{Kern}(f)) = \dim_K(U \cap U')$. Aus Satz 24 und Satz 12 folgt

$$\begin{aligned}\dim_K(U+U') &= \dim_K(U \times U') - \dim_K(\mathrm{Kern}(f))\\ &= \dim_K(U) + \dim_K(U') - \dim_K(U \cap U').\end{aligned}$$

§7. Injektive und surjektive lineare Funktionen

Satz 39 :

(1) *Die Menge*

$$\mathrm{Lin}_K(V,W) := \{f \mid f : V \to W \ K\text{-}linear\}$$

aller linearen Funktionen von V nach W ist ein Untervektorraum des Vektorraums $\mathcal{F}(V,W)$ aller Funktionen von V nach W mit der punktweisen Addition und Skalarmultiplikation.

(2) *Die Menge* $\mathrm{Lin}_K(V,V)$ *ist mit der punktweisen Addition und der Hintereinanderausführung* $\circ$ *ein Ring mit Einselement* Id_V*, und für $a \in K$ und $f,g \in \mathrm{Lin}_K(V,V)$ gilt*

$$a(f \circ g) = (af) \circ g = f \circ (ag).$$

(3) *Die Menge*

$$\mathrm{GL}_K(V) := \{f \mid f : V \to V \ K\text{-}linear\ und\ bijektiv\}$$

ist mit der Hintereinanderausführung eine Gruppe mit Einselement Id_V *und heißt die* allgemeine lineare Gruppe *von V.*

Beweis: Übung.

Definition 40 : M und N seien Mengen. Eine Funktion $f : M \longrightarrow N$ ist *injektiv* bzw. *surjektiv* genau dann, wenn jedes Element von N höchstens bzw. mindestens ein Urbild hat.

Beispiel 41 : M und N seien Mengen. Eine Funktion $f : M \longrightarrow N$ ist genau dann bijektiv, wenn sie sowohl injektiv als auch surjektiv ist.

Satz 42 : *Sei $f : V \to W$ eine lineare Funktion und $\underline{v} := (v_i)_{i\in I}$ eine Familie in V. Es sei*

$$f(\underline{v}) := (f(v_i))_{i\in I}\,.$$

(1) *Sei $\underline{v}$ ein Erzeugendensystem von V. Dann ist f genau dann surjektiv, wenn $f(\underline{v})$ ein Erzeugendensystem von W ist.*
Insbesondere: Wenn $(v_i)_{i\in I}$ ein Erzeugendensystem von V ist, dann ist $(f(v_i))_{i\in I}$ ein Erzeugendensystem von $\mathrm{Bild}(f)$.
(2) *Sei $\underline{v}$ eine Basis von V. Dann ist f genau dann injektiv, wenn $f(\underline{v})$ linear unabhängig ist.*
(3) *Sei $\underline{v}$ eine Basis von V. Dann ist f genau dann bijektiv, wenn $f(\underline{v})$ eine Basis von W ist.*

Beweis:

(1) Wenn $f(\underline{v})$ ein Erzeugendensystem von W ist, dann gibt es zu einem beliebigen $w \in W$ eine Koeffizientenfamilie $(c_i)_{i\in I}$ so, dass $w = \sum_{i\in I} c_i f(v_i)$ ist. Weil f linear ist, ist $w = f(\sum_{i\in I} c_i v_i) \in \mathrm{Bild}(f)$, also f surjektiv.
Wenn umgekehrt f surjektiv ist, dann existiert zu einem beliebigen $w \in W$ ein Urbild $y \in V$. Da $\underline{v}$ ein Erzeugendensystem von V ist, gibt es eine Koeffizientenfamilie $(c_i)_{i\in I}$ so, dass $y = \sum_{i\in I} c_i v_i$ ist. Weil f linear ist, ist $w = f(y) = \sum_{i\in I} c_i f(v_i)$ eine Linearkombination von $f(\underline{v})$, also $f(\underline{v})$ ein Erzeugendensystem von W.

(2) Sei $f(\underline{v})$ linear unabhängig und $y, z \in V$ mit

$$f(y) = f(z)\,.$$

Seien $y = \sum_{i\in I} c_i v_i$ und $z = \sum_{i\in I} d_i v_i$, dann ist

$$0 = f(y) - f(z) = \sum_{i\in I} c_i f(v_i) - \sum_{i\in I} d_i f(v_i) = \sum_{i\in I} (c_i - d_i) f(v_i),$$

somit $c_i = d_i$ für alle $i \in I$. Daher muss $y = z$ sein und f ist injektiv.
Wenn umgekehrt f injektiv ist, dann ist für eine Koeffizientenfamilie $(c_i)_{i\in I}$ mit $0 = \sum_{i\in I} c_i f(v_i) = f(\sum_{i\in I} c_i v_i)$ auch $\sum_{i\in I} c_i v_i = 0$. Da $\underline{v}$ linear unabhängig ist, folgt $c_i = 0$ für alle $i \in I$, also ist $f(\underline{v})$ linear unabhängig.

(3) folgt aus (1) und (2).

Satz 43 : *Sei $f : V \to W$ eine lineare Funktion. Die Funktion f ist genau dann injektiv, wenn* $\mathrm{Kern}(f) = \{0_V\}$ *ist.*

Beweis: Wenn f injektiv ist, dann gilt für $u \in V$ mit $f(u) = 0$ wegen $f(0_V) = 0_W$ auch $u = 0_V$, also ist $\mathrm{Kern}(f) = \{0_V\}$. Wenn umgekehrt $\mathrm{Kern}(f) = \{0_V\}$ ist, dann folgt für $u, v \in V$ aus $f(u) = f(v)$ wegen $f(u-v) = f(u) - f(v) = 0_W$ auch $u = v$.

Satz 44 : *Seien V, W endlich-dimensionale Vektorräume über K mit $\dim_K(V) = \dim_K(W)$ und $f : V \to W$ eine lineare Funktion. Dann sind die folgenden Aussagen äquivalent:*

(1) *f ist bijektiv.*
(2) *f ist injektiv.*
(3) *f ist surjektiv.*

Beweis: Nach Voraussetzung und Satz 24 gilt

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

Wenn f injektiv ist, dann ist $\mathrm{Kern}(f) = \{0\}$, also $\dim_K(W) = \dim_K(\mathrm{Bild}(f))$, $W = \mathrm{Bild}(f)$ und f surjektiv. Wenn hingegen f surjektiv ist, dann ist $\mathrm{Bild}(f) = W$, also $\dim_K(\mathrm{Kern}(f)) = 0$, $\mathrm{Kern}(f) = \{0\}$ und f injektiv.

§8. Eigenwerte und Eigenvektoren

In diesem Abschnitt sei K ein Körper.

Definition 45 : Sei V ein Vektorraum über K und $f : V \to V$ eine lineare Funktion.

(1) Ein Vektor $u \in V$ heißt *Eigenvektor* von f, wenn $u \neq 0$ ist und eine Zahl $c \in K$ existiert mit

$$f(u) = cu.$$

In diesem Fall ist c eindeutig bestimmt und heißt der *Eigenwert* von f zum Eigenvektor u.

(2) Die Menge aller Eigenwerte von f

$$\mathrm{Sp}_K(f) := \{c \in K \mid \text{ es gibt ein } u \in V \text{ mit } u \neq 0$$
$$\text{und } f(u) = cu\} \subset K$$

nennt man das *Spektrum* von f. Für $c \in \mathrm{Sp}_K(f)$ ist

$$E(f,c) := \{x \in V \mid f(x) = cx\} = \mathrm{Kern}(c\,\mathrm{Id}_V - f)$$

ein Untervektorraum von V, heißt der *Eigenraum* von f zum Eigenwert c, und besteht aus dem Nullvektor sowie allen Eigenvektoren von f zum Eigenwert c.

Satz 46: *Sei V ein n-dimensionaler Vektorraum über K, $f : V \longrightarrow V$ eine lineare Funktion, $\underline{v}$ eine Basis von V und $A := M(f, \underline{v})$ die Matrix von f bezüglich $\underline{v}$.*
Die Spalte $y \in K^{n\times 1}$ ist genau dann ein Eigenvektor von A zum Eigenwert $c \in K$, wenn der Vektor $\underline{v}y \in V$ ein Eigenvektor von f zum Eigenwert $c \in K$ ist.

Beweis: $f(\underline{v}y) = \underline{v}Ay$.

Beispiel 47: Sei

$$\mathcal{C}^\infty(\mathbb{R}, \mathbb{R}) := \{g \mid g : \mathbb{R} \to \mathbb{R} \text{ beliebig oft differenzierbar}\}$$

und

$$(-)' : \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}) \to \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}),\, g \mapsto g',$$

wobei g' die Ableitung der Funktion g bezeichnet. Dann ist $\mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ mit der punktweisen Addition und Skalarmultiplikation ein Vektorraum über $\mathbb{R}$ und die Funktion $(-)'$ ist $\mathbb{R}$-linear. Für $c \in \mathbb{R}$ sei

$$g_c : \mathbb{R} \to \mathbb{R},\, t \mapsto e^{ct}.$$

Dann gilt

$$(-)'(g_c) = cg_c,$$

also ist g_c ein Eigenvektor von $(-)'$ zum Eigenwert c. Da $c \in \mathbb{R}$ beliebig war, folgt

$$\mathrm{Sp}_{\mathbb{R}}((-)') = \mathbb{R}.$$

Da Eigenvektoren zu paarweise verschiedenen Eigenwerten linear unabhängig sind, ist die Familie $(g_c)_{c\in\mathbb{R}}$ linear unabhängig.

Satz 48: *Sei V ein endlich-dimensionaler Vektorraum über K und $f : V \to V$ eine lineare Funktion. Dann ist $c \in K$ genau dann Eigenwert von f, wenn*

$$\det(c\,\mathrm{Id}_V - f) = 0$$

ist. Die Funktion

$$\chi_f : K \to K,\, z \mapsto \det(z\,\mathrm{Id}_V - f),$$

heißt das charakteristische Polynom *von f.*

Beweis: Es ist c Eigenwert von f genau dann, wenn ein Vektor $u \in V$ mit $u \neq 0$ existiert, sodass $f(u) = cu$ ist. Dies ist äquivalent dazu, dass die lineare Funktion

$$c\,\mathrm{Id}_V - f : V \to V,\, x \mapsto cx - f(x),$$

nicht injektiv ist. Nach Satz 44 ist $c\,\mathrm{Id}_V - f$ nicht injektiv genau dann, wenn $c\,\mathrm{Id}_V - f$ nicht bijektiv ist. Das ist genau dann der Fall, wenn $\det(c\,\mathrm{Id}_V - f) = 0$ ist.

Satz 48 legt folgendes Verfahren nahe, die Eigenwerte und Eigenvektoren einer linearen Funktion $f : V \to V$ zu berechnen, falls V endlichdimensional ist:

(1) Wähle eine Basis $\underline{v} = (v_1, \dots, v_n)$ von V und bestimme die Matrix A von f bezüglich $\underline{v}$. Dann ist $cI_n - A$ die Matrix von $c\,\mathrm{Id}_V - f$ bezüglich $\underline{v}$.
(2) Finde alle $c \in K$ mit $\det(cI_n - A) = 0$.
(3) Bestimme für jeden Eigenwert c von f den Eigenraum $E(f,c) = L(c\,\mathrm{Id}_V - f, 0)$ durch Lösen des durch $cI_n - A$ gegebenen homogenen Systems linearer Gleichungen.

Satz 49: *Sei K ein Körper mit unendlich vielen Elementen, V ein endlichdimensionaler Vektorraum über K der Dimension n, $f : V \to V$ eine lineare Funktion und χ_f das charakteristische Polynom von f. Dann gilt:*

(1) *χ_f ist eine Polynomfunktion mit $\mathrm{gr}(\chi_f) = n$.*
(2) *χ_f ist normiert.*
(3) *Der Koeffizient von χ_f bei X^{n-1} ist $\mathrm{spur}(f)$.*
(4) *Der Koeffizient von χ_f bei 1 ist $(-1)^n \det(f)$.*

Somit hat χ_f die Gestalt

$$\chi_f = X^n - \mathrm{spur}(f) X^{n-1} + \cdots + (-1)^n \det(f).$$

Für ($K = \mathbb{R}$ und n ungerade) oder ($K = \mathbb{C}$ und $n > 0$) besitzt f mindestens einen Eigenvektor.

Beweis: Folgt aus der Definition des charakteristischen Polynoms und der Determinante.

Satz 50: *Sei V ein endlichdimensionaler Vektorraum über K und seien $U_1, \dots, U_\ell$ Untervektorräume von V.*
Die Summe $U_1 + \cdots + U_\ell$ heißt direkt, *wenn für alle $u_1 \in U_1, \dots, u_\ell \in U_\ell$ aus $u_1 + \cdots + u_\ell = 0$ auch $u_1 = \cdots = u_\ell = 0$ folgt. In diesem Fall schreibt man*

$$U = U_1 \oplus \cdots \oplus U_\ell$$

und es gilt:

(1) *Für alle $u \in U_1 + \cdots + U_\ell$ sind die Summanden $u_1 \in U_1, \dots, u_\ell \in U_\ell$ in der Darstellung*

$$u = u_1 + \cdots + u_\ell$$

eindeutig bestimmt.

(2) *Sind* $(u_{11},\dots,u_{1n_1}),\dots,(u_{\ell 1},\dots,u_{\ell n_\ell})$ *Basen von* $U_1,\dots,U_\ell$, *dann ist* $(u_{11},\dots,u_{1n_1},\dots,u_{\ell 1},\dots,u_{\ell n_\ell})$ *eine Basis von* $U_1\oplus\cdots\oplus U_\ell$.

(3) $\dim_K(U_1\oplus\cdots\oplus U_\ell)=\dim_K(U_1)+\cdots+\dim_K(U_\ell)$.

Beweis: Die Funktion

$$F: U_1\times\cdots\times U_\ell\to U_1+\cdots+U_\ell\,,\,(u_1,\dots,u_\ell)\mapsto u_1+\dots u_\ell\,,$$

ist linear und surjektiv. Wenn die Summe $U_1+\cdots+U_\ell$ direkt ist, dann ist $\mathrm{Kern}(f)=\{0\}$, also F injektiv und F ein Isomorphismus. Damit folgen (1) aus der Injektivität von F und (2),(3) aus Satz 12.

Satz 51: *Sei* V *ein Vektorraum über* K, $f:V\to V$ *eine lineare Funktion und* $\mathrm{Sp}_K(f)=\{c_1,\dots,c_\ell\}$. *Dann gilt:*

(1) *Die Summe der Eigenräume von* f *ist direkt, d.h.*

$$E(f,c_1)+\cdots+E(f,c_\ell)=E(f,c_1)\oplus\cdots\oplus E(f,c_\ell)\,.$$

(Die Summe der Eigenräume ist im Allgemeinen aber nicht der ganze Vektorraum V*).*

(2) *Die Zahl der Eigenwerte ist durch die Dimension beschränkt, d.h.*

$$\#(\mathrm{Sp}_K(f))\le\dim_K(V)\,.$$

Beweis:

(1) Eigenvektoren zu paarweise verschiedenen Eigenwerten sind linear unabhängig.

(2) Nach Definition ist ein Eigenraum nicht der Nullraum und hat daher mindestens Dimension 1. Nach Satz 50 folgt

$$\begin{aligned}\#(\mathrm{Sp}_K(f))=\ell&\le\dim_K(E(f,c_1))+\cdots+\dim_K(E(f,c_\ell))\\&=\dim_K(E(f,c_1)\oplus\cdots\oplus E(f,c_\ell))\le\dim_K(V)\,,\end{aligned}$$

was zu zeigen war.

KAPITEL 2

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 52: Seien $V_1, \dots, V_\ell$ und W Vektorräume über K. Eine Funktion

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

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

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

und

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

Sei

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

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

Beispiel 53: Die Funktion

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

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

$$\begin{aligned} x_1 \dots x_{k-1}(x_k + y)x_{k+1} \dots x_\ell = x_1 \dots x_{k-1} x_k x_{k+1} \dots x_\ell \\ + x_1 \dots x_{k-1} y x_{k+1} \dots x_\ell \end{aligned}$$

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

Beweis: Übung.

Satz 55 : *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

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

genügt es nach Satz 54 zu zeigen, dass für jedes $\sigma \in S_n$ die Funktion

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

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 56 : *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 $x_1 = \sum_{i_1\in I_1} c_{1i_1} v_{1i_1} \in V_1,\dots,x_\ell = \sum_{i_\ell\in I_\ell} c_{\ell i_\ell} v_{\ell i_\ell} \in V_\ell$

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

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 57 : 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 56 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 58 : Seien V und W Vektorräume über K. Eine Funktion

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

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 $(x_1,\ldots,x_\ell) \in V^\ell$ mit $x_i = x_k$ für irgendwelche $i \neq k$ ist $f(x_1,\ldots,x_\ell) = 0$. Sei

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

Satz 59 : $\mathrm{Alt}_K(V^\ell,W)$ *ist mit der punktweisen Addition und Skalarmultiplikation ein Untervektorraum von* $\mathrm{Mult}(V^\ell,W)$.

Beweis: Übung.

Satz 60 : *Die Funktion*

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

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

Beweis: Sei $(x_1,\ldots,x_\ell) \in (K^{n\times 1})^\ell$ 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

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

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

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

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

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

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

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

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

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

also

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

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

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

Hilfssatz 62: *Seien V,W Vektorräume über K und $f : V^\ell \to W$ alternierend. Sei $\underline{v}=(v_1,\dots,v_n)\in V^n$, $T\in K^{n\times\ell}$ und $\underline{w}=\underline{v}T\in V^\ell$. Dann ist*

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

wobei

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

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 61 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 63 : *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 62

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

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 60 und Satz 59 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 64: 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 65 : Sei $A \in K^{n\times n}$. Für $1 \le i,j \le n$ 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 66 : *Sei $A \in K^{n\times n}$. Dann gilt:*

(1) *Für $1 \le j \le n$ 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 $1 \le i \le n$ 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 60 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

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

Satz 67 : *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 $n \geq 2$ entwickeln wir die Determinante von A nach der ersten Spalte und erhalten nach Satz 66

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

Da die Streichungsmatrix $A^{(i,1)}$ wieder Blockdreiecksform besitzt, können wir die Induktionsannahme anwenden und erhalten nach Satz 66

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

§4. Die adjungierte Matrix

Satz 68: *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 $1 \le i, j \le n$, 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 $1 \le i, j \le n$ ist wegen Satz 66,(2)

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

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

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

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 69: *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 $\det(A) \ne 0$ und $\det(A^{-1}) = \det(A)^{-1}$. Wenn umgekehrt $\det(A) \ne 0$ ist, dann ist nach Satz 68

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

und

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

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

Beispiel 70 : 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 71 : *(*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 69 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 $1 \le i \le n$

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

woraus die Behauptung folgt.

§5. Symmetrische Bilinearformen

Definition 72 : 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

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

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

$$A^\top = A$$

ist.

Beispiel 73: 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 74: 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 75: *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*

$$b(\underline{v}x,\underline{v}y) = x^\top \cdot A \cdot y \ .$$

(2) *Die Matrix A ist genau dann symmetrisch, wenn b symmetrisch ist.*
(3) *Für alle $S \in \mathrm{GL}_n(K)$ ist*

$$M(b,\underline{v}S) = S^\top \cdot A \cdot S \ .$$

Beweis:

(1)

$$\begin{aligned} b(\underline{v}x,\underline{v}y) &= b\left(\sum_{i=1}^n x_iv_i, \sum_{j=1}^n y_jv_j\right) = \sum_{i,j=1}^n x_iy_ib\left(v_i,v_j\right) \\ &= \sum_{i,j=1}^n x_iA_{ij}y_j = x^\top \cdot A \cdot y. \end{aligned}$$

(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 $1 \leq i, j \leq n$ 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 76 : 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 77 : Für $d_1, ..., d_n \in K$ sei

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

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

Satz 78 : *Es seien $A \in K^{n\times n}$ eine symmetrische Matrix und r der Rang von A. Dann gibt es $d_1, ..., d_r \in K\setminus\{0\}$ so, dass A und*

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

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

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

ist:

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

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

und das Verfahren zu Ende.
(3) *Falls $C_{jj} = 0$ ist für alle $j \in \{i, ..., n\}$ und $C_{k\ell} \neq 0$ ist für Indizes $k, \ell \in \{i, ..., n\}$, 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 $j \in \{i+1, ..., n\}$, 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 $j = i+1, ..., n$ 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 $r = \operatorname{rg}(A) = \operatorname{rg}\left(P^\top \cdot A \cdot P\right)$ folgt daraus die Behauptung.

Satz 79 : *„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 $s+t = \operatorname{rg}(A)$. 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 78 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 $s+t = \operatorname{rg}(I_n^{s,t})$ 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 $U \cap W \neq \{0\}$.
Für $0 \neq x \in U \cap W$ würde dann

$$\begin{aligned} 0 \geq -\sum_{i=s+1}^{s+t} x_i^2 &= x^T \cdot I_n^{s,t} \cdot x = \\ &= (Sx)^T \cdot I_n^{s',t'} \cdot (Sx) = \sum_{i=1}^{s'} (Sx)_i^2 > 0 \end{aligned}$$

sein. Widerspruch.

§6. Positiv definite Matrizen

Definition 80 : Eine Bilinearform b auf einem reellen Vektorraum V heißt *positiv definit* bzw. *positiv semidefinit*, wenn für alle $v \in V \setminus \{0\}$

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

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

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

ist.

Beispiel 81 : 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 82 :

(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 83 : *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 79 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 82 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 79 ist A kongruent zu I_n. Nach Satz 82 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 84 : *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 83.

Satz 85 : *Eine symmetrische reelle Matrix A ist genau dann positiv definit, wenn alle ihre* Hauptminoren

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

positive Zahlen sind.

Beweis: Für $1 \leq k \leq n$ 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 83, (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 83, $(1) \Rightarrow (2)$, 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 83 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 83 ist C positiv definit, daher auch die zu C kongruente Matrix A.

Satz 86 : *Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch und positiv definit. Dann gibt es eine eindeutig bestimmte untere Dreiecksmatrix C mit positiven Einträgen in der Diagonale so, dass*

$$C \cdot C^\top = A$$

*ist (*Cholesky-Zerlegung *von A).*

Beweis: Wir zeigen die Existenz durch Induktion über n. Weil A positiv definit ist, muss nach Satz 85 der Eintrag $A_{11} > 0$ sein.
Für $n = 1$ ist die Aussage richtig, weil jede positive reelle Zahl genau eine positive reelle Wurzel hat.

Sei $n > 1$. Mit Schritt (5) des Verfahrens in Satz 78 und anschließender Multiplikation der ersten Zeile und ersten Spalte mit $1/\sqrt{A_{11}}$ erhält man

$$P^\top \cdot A \cdot P = \begin{pmatrix} 1 & 0 \\ 0 & A' \end{pmatrix},$$

wobei P eine obere Dreiecksmatrix mit positiven Einträgen in der Diagonale und A' eine symmetrische und positiv definite $(n-1) \times (n-1)$-Matrix ist. Nach Induktionsannahme gibt es eine untere Dreiecksmatrix C' mit positiven Einträgen in der Diagonale so, dass $C' \cdot C'^\top = A'$ ist. Die Matrix

$$C := (P^\top)^{-1} \cdot \begin{pmatrix} 1 & 0 \\ 0 & C' \end{pmatrix}$$

hat dann die gewünschten Eigenschaften.
Zum Beweis der Eindeutigkeit nehmen wir an, dass B und C untere Dreiecksmatrizen mit positiven Einträgen in der Diagonale sind so, dass

$$C \cdot C^\top = B \cdot B^\top$$

ist. Dann ist

$$B^{-1} \cdot C = B^\top \cdot (C^\top)^{-1}.$$

Weil $B^{-1} \cdot C$ eine untere und $B^\top \cdot (C^\top)^{-1}$ eine obere Dreiecksmatrix ist, müssen beide Diagonalmatrizen sein und $B_{ii}^2 = C_{ii}^2$, $1 \le i \le n$, sein. Weil die Einträge von B und C in der Diagonale positiv sind, folgt daraus $B^{-1} \cdot C = I_n$, also $B = C$.

KAPITEL 3

Affine Geometrie

§1. Parallele affine Unterräume

Es seien K ein Körper, V ein Vektorraum über K und $M_1 = p_1 + U_1$, $M_2 = p_2 + U_2$ affine Unterrräume von V mit Aufpunkten p_1, $p_2 \in V$ und parallelen Untervektorräumen U_1, U_2 in V.

Definition 87 : Die affinen Unterräume M_1 und M_2 heißen *parallel*, wenn $U_1 \subseteq U_2$ oder $U_2 \subseteq U_1$ ist.

Beispiel 88 : Ein affiner Unterrraum und sein paralleler Untervektorraum sind parallel. Jeder Punkt von V ist zu jedem affinen Unterraum von V parallel.

Satz 89 : *Wenn M_1 und M_2 parallel sind, dann ist $M_1 \subseteq M_2$ oder $M_2 \subseteq M_1$ oder $M_1 \cap M_2 = \emptyset$.*

Beweis: Wir nehmen o.E.d.A. an, dass $U_1 \subseteq U_2$ ist. Wenn $M_1 \cap M_2$ nicht leer ist, dann gibt es ein $p \in M_1 \cap M_2$. Daher ist $M_1 = p + U_1 \subseteq p + U_2 = M_2$.

Definition 90 : Es seien I eine endliche Menge und $(v_i)_{i \in I}$ eine Familie in V. Eine Linearkombination $\sum_{i \in I} c_i v_i$ von $(v_i)_{i \in I}$ heißt *affine Kombination* von $(v_i)_{i \in I}$, wenn $\sum_{i \in I} c_i = 1$ ist. Die Menge aller affinen Linearkombinationen von $(v_i)_{i \in I}$ heißt *affine Hülle* von $(v_i)_{i \in I}$.

Beispiel 91 : Die affine Hülle von zwei Vektoren v_1 und v_2 ist ein Punkt, wenn $v_1 = v_2$ ist, bzw. die Gerade

$$\{c_1 v_1 + c_2 v_2 \mid c_1, c_2 \in K, c_1 + c_2 = 1\} = \{v_1 + c(v_2 - v_1) \mid c \in K\},$$

wenn $v_1 \neq v_2$ ist.

Satz 92 :

(1) *Es seien M ein affiner Unterraum von V und $(v_i)_{i \in I}$ eine endliche Familie in M. Dann ist die affine Hülle von $(v_i)_{i \in I}$ in M enthalten.*

(2) *Die affine Hülle einer Familie $(v_i)_{i\in I}$ in V ist ein affiner Unterraum von V. Der dazu parallele Untervektorraum wird von $(v_i - v_j)_{i\in I,\, i\neq j}$ erzeugt, wobei $j \in I$ beliebig gewählt werden kann.*
(3) *Die affine Hülle von $(v_i)_{i\in I}$ ist der (bezüglich Inklusion) kleinste affine Unterraum, der alle v_i, $i \in I$, enthält.*

Beweis:

(1) Sei $p \in M$, U der zu M parallele Untervektorraum und $(c_i)_{i\in I}$ eine Familie in K mit $\sum_{i\in I} c_i = 1$. Dann ist

$$\sum_{i\in I} c_i v_i = \left(\sum_{i\in I} c_i\right) p - \left(\sum_{i\in I} c_i\right) p + \sum_{i\in I} c_i v_i = \\ = p + \sum_{i\in I} c_i (v_i - p) \in p + U = M.$$

(2) Sei $j \in I$ und

$$M := v_j + {}_K\langle v_i - v_j;\, i \in I, i \neq j\rangle.$$

Dann ist $(v_i)_{i\in I}$ eine Familie in M und nach (1) ist ihre affine Hülle in M enthalten.
Sei umgekehrt $(d_i)_{i\in I}$ eine Familie in K.
Dann ist

$$v_j + \sum_{\substack{i\in I\\ i\neq j}} d_i (v_i - v_j) = \sum_{\substack{i\in I\\ i\neq j}} d_i v_i + \left(1 - \sum_{\substack{i\in I\\ i\neq j}} d_i\right) v_j$$

eine affine Linearkombination von $(v_i)_{i\in I}$. Daher ist jedes Element von M in der affinen Hülle von $(v_i)_{i\in I}$ enthalten.
(3) Folgt aus (1) und (2).

Definition 93 : Affine Unterräume von V heißen *kollinear* bzw. *koplanar*, wenn sie alle in einer Geraden bzw. Ebene in V enthalten sind.

Satz 94 :

(1) *Drei Punkte v_1, v_2, $v_3 \in V$ sind genau dann kollinear, wenn die Vektoren $v_2 - v_1$ und $v_3 - v_1$ linear abhängig sind.*
(2) *Vier Punkte v_1, v_2, v_3, $v_4 \in V$ sind genau dann koplanar, wenn die Vektoren $v_2 - v_1$, $v_3 - v_1$ und $v_4 - v_1$ linear abhängig sind.*
(3) *Zwei Geraden $p_1 + Kv_1$ und $p_2 + Kv_2$ sind genau dann koplanar, wenn die Vektoren $p_1 - p_2$, v_1 und v_2 linear abhängig sind.*

Beweis: Die ersten zwei Aussagen folgen aus Satz 92, (2). Der zur affinen Hülle von $(p_1, p_2, p_1+v_1, p_2+v_2)$ parallele Untervektorraum wird von $p_1 - p_2$, v_1 und v_2 erzeugt.

Satz 95 : *Zwei verschiedene koplanare Geraden schneiden einander in genau einem Punkt oder sie sind parallel.*

Beweis: Seien M_1 und M_2 verschiedene koplanare Geraden und E die Ebene, die beide enthält. Wenn M_1 und M_2 nicht parallel sind, dann ist $U_1 \cap U_2 = \{0\}$ und $U_1 + U_2 = U_1 \oplus U_2$ ist der zu E parallele Untervektorraum. Wegen $p_1,\ p_2 \in E$ ist $p_1 - p_2 \in U_1 \oplus U_2$, daher gibt es eindeutig bestimmte Vektoren $u_1 \in U_1$, $u_2 \in U_2$ so, dass $p_1 - p_2 = u_1 + u_2$ ist. Somit ist $M_1 \cap M_2 = \{p_1 - u_1\} = \{p_2 + u_2\}$.

§2. Affine Funktionen

Es seien V und W Vektorräume über einem Körper K.

Definition 96 : Eine Funktion $a\colon V \longrightarrow W$ 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 97 : Lineare Funktionen und Translationen sind affine Funktionen.

Beispiel 98 : 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 99 : *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 100 : *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 101 : *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
$f(c_1x_1 + c_2x_2) = a(c_1x_1 + c_2x_2) - w =$
$= a(c_1x_1 + c_2x_2 + (1 - c_1 - c_2)0) - w =$
$= c_1a(x_1) + c_2a(x_2) + (1 - c_1 - c_2) \cdot a(0) - w =$
$= c_1(a(x_1) - w) + c_2(a(x_2) - w) = c_1f(x_1) + c_2f(x_2)$.

§3. Polytope und Schwerpunkte

Es seien $K = \mathbb{Q}$ oder $\mathbb{R}$ und V ein Vektorraum über K.

Definition 102: Es seien I eine endliche Menge und $(v_i)_{i\in I}$ eine Familie in V.
Eine Linearkombination $\sum_{i\in I} c_iv_i$ von $(v_i)_{i\in I}$ heißt *konvexe Linearkombination* von $(v_i)_{i\in I}$, wenn $\sum_{i\in I} c_i = 1$ und $c_i \geq 0$ für alle $i \in I$ ist.
Die Menge der konvexen Linearkombinationen von $(v_i)_{i\in I}$ heißt *konvexe Hülle* von $(v_i)_{i\in I}$.
Die konvexe Hülle zweier Vektoren v_1, v_2 heißt *Strecke* zwischen v_1 und v_2.
Die konvexe Hülle dreier nicht kollinearer Punkte v_1, v_2, v_3 heißt *Dreieck* mit Eckpunkten v_1, v_2, v_3.
Eine Teilmenge von V heißt *Polytop*, wenn sie die konvexe Hülle einer endlichen Familie in V ist.

Es sei $I := \{1, \ldots, n\}$ und $c_1, \ldots, c_n \in \mathbb{R}_{\geq 0}$. Für $c_n \neq 1$ ist

$$\sum_{i=1}^{n} c_iv_i = (1 - c_n)\left(\sum_{i=1}^{n-1} \frac{c_i}{1 - c_n} v_i\right) + c_nv_n = (1 - c_n)w + c_nv_n\,,$$

wobei $w := \sum_{i=1}^{n-1} \frac{c_i}{1-c_n} v_i$ in der konvexen Hülle H von $(v_1, \ldots, v_{n-1})$ liegt.
Daraus folgt: Für $n \geq 3$ ist die konvexe Hülle von $(v_1, \ldots, v_n)$ die Vereinigung aller Strecken zwischen v_n und den Elementen von H.

Beispiel 103: Eine Teilmenge von $\mathbb{R}$ ist genau dann ein Polytop, wenn sie ein abgeschlossenes Intervall ist.

Satz 104: *Es seien P die konvexe Hülle einer Familie $\left(w_j\right)_{j\in J}$ in V und $(v_i)_{i\in I}$ eine Familie in P. Dann ist die konvexe Hülle von $(v_i)_{i\in I}$ in P enthalten.*

Beweis: Für alle $i \in I$ ist der Vektor v_i eine konvexe Linearkombination $\sum\limits_{j\in J} c_{ji} w_j$ von $\left(w_j\right)_{j\in J}$.
Sei $\sum\limits_{i\in I} d_i v_i$ eine konvexe Linearkombination von $(v_i)_{i\in I}$. Dann ist

$$\sum_{i\in I} d_i v_i = \sum_{i\in I}\sum_{j\in J} d_i c_{ji} w_j = \sum_{j\in J}\left(\sum_{i\in I} d_i c_{ji}\right) w_j$$

mit $\sum\limits_{i\in I} d_i c_{ji} \geq 0,$ für alle $j \in J$, und

$$\sum_{j\in J}\left(\sum_{i\in I} d_i c_{ji}\right) = \sum_{i\in I} d_i \left(\sum_{j\in J} c_{ji}\right) = \sum_{i\in I} d_i = 1 \ .$$

Daher ist $\sum\limits_{i\in I} d_i v_i \in P$.

Definition 105 : 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 106 : *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 101.

Beispiel 107 : Es seien P ein Polytop in V und $a : V \to \mathbb{R}$ eine affine Funktion. Dann ist $a(P)$ ein abgeschlossenes Intervall in $\mathbb{R}$.

Satz 108 : *Es seien u, v, w drei nicht kollineare Punkte in V . Die Gerade durch u bzw. v bzw. w und den Mittelpunkt der Strecke zwischen den anderen zwei Punkten heißt* Schwerlinie *des Dreiecks mit Eckpunkten u, v, w durch u bzw. v bzw. w.*
Die drei Schwerlinien sind paarweise verschieden und schneiden einander im Schwerpunkt $\frac{1}{3}(u + v + w)$ von (u, v, w) .

Beweis: Da u, v, w nicht kollinear sind, sind nach Satz 94 die Vektoren $v-u$ und $w-u$ linear unabhängig. Also sind auch

$$v-u \quad \text{und} \quad \frac{1}{2}(v-u)+\frac{1}{2}(w-u)=\frac{1}{2}(v+w)-u$$

linear unabhängig, nach Satz 94 sind daher $u, v, \frac{1}{2}(v+w)$ nicht kollinear. Somit liegt v nicht auf der Schwerlinie durch u . Daher sind die Schwerlinien durch u und durch v verschieden und die drei Schwerlinien haben höchstens einen Schnittpunkt. Wegen

$$\frac{1}{3}(u+v+w)=\frac{1}{3}u+\frac{2}{3}\left(\frac{1}{2}(v+w)\right)=\frac{1}{3}v+\frac{2}{3}\left(\frac{1}{2}(u+w)\right)= \\ =\frac{1}{3}w+\frac{2}{3}\left(\frac{1}{2}(u+v)\right)$$

liegt der Schwerpunkt auf allen Schwerlinien.

§4. Affine Räume

Definition 109 : Es seien $(G,\star)$ eine Gruppe mit neutralem Element e und M eine Menge. Eine Funktion $G\times M \longrightarrow M$, $(s,m)\longmapsto s\cdot m$, ist eine *Operation der Gruppe G auf der Menge M*, wenn gilt:
für alle $m\in M$ ist $e\cdot m=m$ und
für alle $s,t\in G$ und alle $m\in M$ ist $(s\star t)\cdot m=s\cdot(t\cdot m)$.

Beispiel 110 : Die Funktion

$$S_n\times\{1,2,\dots,n\}\to\{1,2,\dots,n\},\ (\sigma,i)\mapsto\sigma(i),$$

ist eine Operation der Permutationsgruppe S_n auf der Menge $\{1,2,\dots,n\}$.

Definition 111 : 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 112 : *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.

Beispiel 113 : Sei V ein Vektorraum, $p \in V$ und U ein Untervektorraum von V. Dann ist der affine Unterraum $p+U$ mit

$$U \times (p+U) \longrightarrow p+U \ \ (v, p+u) \longmapsto p+(u+v)\,,$$

ein affiner Raum über U. Insbesondere ist jeder Vektorraum ein affiner Raum (über sich selbst).

Beispiel 114 : Sie E die Zeichenebene oder der Anschauungsraum und $T(E)$ 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.

§5. 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 115 : 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 116 : 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 $L(f, \leq b) = \{x \in \mathbb{R}^n \mid \sum_{i=0}^{n} a_i x_i \leq b\}$.

Satz 117 : *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 118 : 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 $\mathrm{Kern}(f)$. Für

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

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

$$L(f, \leq 1) = \{c_1v_1 + c_2v_2 + c_3v_3 \mid c_1, c_2, c_3 \in \mathbb{R}, c_3 \leq 1\} =$$
$$= \big\{g \in V \mid g(t) = c_2t^2 + c_1t + (c_3 - \frac{1}{3}c_2 - \frac{1}{2}c_1),\ c_1, c_2, c_3 \in \mathbb{R},\ c_3 \leq 1\big\}.$$

Beispiel 119 : 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 120 : 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 (siehe Definition 102).

Definition 121 : 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.

Satz 122 : *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 st.*

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.

§6. Systeme linearer Ungleichungen und Polyeder

In diesem Abschnitt sei n eine positive ganze Zahl und V ein n-dimensionaler reeller Vektorraum.

Definition 123: 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

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

Es ist

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

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

Definition 124: Ein *Polyeder in V* ist der Durchschnitt von endlich vielen Halbräumen in V.

Nach Satz 122 ist jedes Polyeder in V die Lösungsmenge eines Systems von endlich vielen linearen Ungleichungen.

Wenn $\underline{v} = (v_1, \ldots, v_n)$ eine Basis von V ist, dann werden die linearen Funktionen $f_1, \ldots, f_k \in V^*$ eindeutig durch die n-Tupel $\big(f_i(v_1), \ldots, f_i(v_n)\big) \in \mathbb{R}^n$ 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 $L(f_1, \ldots, f_k, \leq b_1, \ldots, b_k) = \{\underline{v}x \mid x \in L(A, \leq b)\}$.

Beispiel 125 : Die Aufgabe “Finde alle $x \in \mathbb{R}^4$ mit

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

kann als System linearer Ungleichungen aufgefasst werden, weil die Menge dieser $x \in \mathbb{R}^4$ gleich $L(A, \le (2,-1,1,-1)^T)$ 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 $Ax \le b$ umgeschrieben werden”).

Definition 126 : 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 $\ell \in \mathbb{N}$, $v_0, \dots, v_\ell \in L$, $c_0, \dots, c_\ell \in \mathbb{R}_{\ge 0}$ ist $c_0 v_0 + \cdots + c_\ell v_\ell \in L$.

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

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

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

Beispiel 127 : Sei $\dim_{\mathbb{R}}(V) = 1$ und $0 \ne v \in V$. Dann gibt es in V genau vier Kegel und zwar $\{0\}$, $\mathcal{K}(v)$, $\mathcal{K}(-v)$ und V.

Beispiel 128 : Sei $\dim_{\mathbb{R}}(V) = 2$ und L ein Kegel in V mit $\{0\} \ne L \ne V$. 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 129 : Die Menge

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

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

Satz 130: *Es seien $u_0, u_1, \dots, u_\ell \in V$, $L := \mathcal{K}(u_0, \dots, u_\ell)$ und $0 \neq g \in V^*$. Dann wird $L \cap Kern(g)$ von der endlichen Menge*

$$\begin{aligned} E := &\big\{u_j \mid g(u_j) = 0,\ 0 \le j \le \ell\big\} \cup \\ &\cup \big\{g(u_i)u_j - g(u_j)u_i \mid 0 \le i, j \le \ell,\ g(u_i) > 0,\ g(u_j) < 0\big\} \end{aligned}$$

erzeugt.

Beweis: Wegen $g\big(g(u_i)u_j - g(u_j)u_i\big) = g(u_i)g(u_j) - g(u_j)g(u_i) = 0$ ist der von E erzeugte Kegel $\mathcal{K}(E)$ in $\mathrm{Kern}(g) \cap L$ enthalten.

Seien $c_0, \dots, c_\ell \in \mathbb{R}_{\ge 0}$ so, dass $w := \sum_{i=0}^{\ell} c_i u_i \in L$ und $g(w) = 0$ ist. Wir zeigen, dass w als nichtnegative Linearkombination von E geschrieben werden kann. Seien

$$\begin{aligned} \mathcal{N} &:= \{i \mid 0 \le i \le \ell,\ g(u_i) < 0\}, \\ \mathcal{O} &:= \{i \mid 0 \le i \le \ell,\ g(u_i) = 0\}, \\ \mathcal{P} &:= \{i \mid 0 \le i \le \ell,\ g(u_i) > 0\} \end{aligned}$$

und

$$d := \sum_{i \in \mathcal{P}} c_i g(u_i) \ge 0.$$

Wegen

$$0 = g(w) = \sum_{i=0}^{\ell} c_i g(u_i) = \sum_{i \in \mathcal{N}} c_i g(u_i) + \sum_{i \in \mathcal{P}} c_i g(u_i)$$

ist

$$\sum_{i \in \mathcal{N}} c_i g(u_i) = -d.$$

Wenn $d = 0$ ist, dann ist $w = \sum_{i \in \mathcal{O}} c_i u_i \in \mathcal{K}(E)$. Wir nehmen daher an, dass $d > 0$ ist. Dann ist

$$\begin{aligned} w &= \sum_{j \in \mathcal{O}} c_j u_j + \sum_{j \in \mathcal{N}} c_j u_j + \sum_{i \in \mathcal{P}} c_i u_i = \\ &= \sum_{j \in \mathcal{O}} c_j u_j + \sum_{j \in \mathcal{N}} \frac{1}{d} d c_j u_j + \sum_{i \in \mathcal{P}} \frac{1}{d} d c_i u_i = \\ &= \sum_{j \in \mathcal{O}} c_j u_j + \sum_{\substack{j \in \mathcal{N} \\ i \in \mathcal{P}}} \frac{1}{d} c_i c_j g(u_i) u_j + \sum_{\substack{i \in \mathcal{P} \\ j \in \mathcal{N}}} \frac{1}{d} c_i c_j \big(-g(u_j)u_i\big) = \\ &= \sum_{j \in \mathcal{O}} c_j u_j + \sum_{\substack{i \in \mathcal{P} \\ j \in \mathcal{N}}} \frac{1}{d} c_i c_j \big(g(u_i)u_j - g(u_j)u_i\big) \end{aligned}$$

und $\frac{1}{d} c_i c_j \ge 0$.

Beispiel 131 : Sei $V := \mathbb{R}^3$, $g : \mathbb{R}^3 \to \mathbb{R}, z \mapsto z_1 + z_2 + z_3$,

$$L_1 := \mathcal{K}\big((1,2,3),(1,1,1),(1,2,1)\big) \quad \text{und}$$
$$L_2 := \mathcal{K}\big((1,-2,-3),(1,-1,1),(1,2,2)\big).$$

Dann ist $L_1 \cap \mathrm{Kern}(g) = \mathcal{K}(\emptyset) = \{0\}$ und

$$L_2 \cap \mathrm{Kern}(g) = \mathcal{K}\big((1,-2,-3)+4(1,-1,1), 5(1,-2,-3)+4(1,2,2)\big) =$$
$$= \mathcal{K}\big((5,-6,1),(9,-2,-7)\big).$$

Satz 132 : *Es seien $\underline{v} = (v_1,\dots,v_n)$ eine Basis von V, $(e_1,\dots,e_k)$ die Standardbasis von $\mathbb{R}^k$, $f_1,\dots,f_k \in V^*$ und*

$$f : V \to \mathbb{R}^k, \quad u \mapsto \big(f_1(u),\dots,f_k(u)\big).$$

Die Menge

$$M := \{(u,y) \in V \times \mathbb{R}^k \mid f_i(u) \le y_i,\ 1 \le i \le k\}$$

ist ein Kegel und wird von der endlichen Menge

$$\big\{\pm\big(v_i, f(v_i)\big) \in V \times \mathbb{R}^k \mid 1 \le i \le n\big\} \cup \big\{(0,e_j) \in V \times \mathbb{R}^k \mid 1 \le j \le k\big\}$$

erzeugt.

Beweis: Man prüft leicht nach, dass M ein Kegel ist und die Elemente $\pm\big(v_i, f(v_i)\big)$, $1 \le i \le n$, sowie $(0,e_j)$, $1 \le j \le k$, enthält. Seien $(u,y) \in M$ und $c_1,\dots,c_n$ die Koordinaten von u bezüglich $\underline{v}$. Für $1 \le i \le n$ sei $\bar{c}_i := 1$, wenn $c_i \ge 0$, und $\bar{c}_i := -1$, wenn $c_i < 0$.

Dann ist

$$(u,y) = \big(u, f(u)\big) + \big(0, y - f(u)\big) =$$
$$= \sum_{i=1}^{n} |c_i| \Big(\bar{c}_i \big(v_i, f(v_i)\big)\Big) + \sum_{j=1}^{k} \big(y_j - f_j(u)\big)(0,e_j)$$

und $|c_i| \ge 0$, $1 \le i \le n$, $y_j - f_j(u) \ge 0$, $1 \le j \le k$.

Satz 133 : *Die Lösungsmenge eines Systems von endlich vielen homogenen linearen Ungleichungen ist ein endlich erzeugter Kegel. Mit dem folgenden Verfahren kann eine endliche Menge berechnet werden, die diesen Kegel erzeugt:*

Seien $f_1,\dots,f_k \in V^*$ und

$$p_j : V \times \mathbb{R}^k \to \mathbb{R}, \quad \big(u,(y_1,\dots,y_k)\big) \mapsto y_j, \quad 1 \le j \le k.$$

- *Berechne mit Satz 132 eine endliche Menge die den Kegel*

$$M := \{(u,y) \in V \times \mathbb{R}^k \mid f_i(u) \le y_i,\ 1 \le i \le k\}$$

erzeugt.

- *Für* $1 \leq j \leq k$ *berechne mit Satz 130 eine endliche Menge* E_j*, die den Kegel*
$$\big(M \cap \bigcap_{i=1}^{j-1} \mathrm{Kern}(p_i)\big) \cap \mathrm{Kern}(p_j)$$
erzeugt.
- *Dann erzeugt die Menge* $\{u \in V \mid (u,0) \in E_k\}$ *den Kegel* $L(f_1,\dots,f_k,\leq 0)$.

Beweis: Folgt aus Satz 130 und Satz 132.

Das Verfahren in Satz 133 kann wie folgt verbessert werden:

Sei $A \in \mathbb{R}^{k\times n}$.
Falls der Rang von A gleich k ist, können wir $G \in \mathbb{R}^{n\times k}$ so berechnen, dass $A\cdot G = I_k$ ist. (Die Spalte G_{-i} ist ein Element von $L(A,e_i)$, wobei e_i die i-te Standardspalte ist). Dann ist
$$L(A,\leq 0) = L(A,0) - \mathcal{K}(G_{-1},G_{-2},\dots,G_{-k})\,.$$

Denn: Für $z \in L(A,0)$ und $y \geq 0$ ist $A(z-Gy) = Az - AGy = 0 - y \leq 0$. Also ist $L(A,0) - \mathcal{K}(G_{-1},G_{-2},\dots,G_{-k})$ in $L(A,\leq 0)$ enthalten.
Sei umgekehrt $u \in L(A,\leq 0)$. Dann ist $Au \leq 0$, also $y := -Au \geq 0$ und $Au + y = 0$. Wegen
$$0 = Au + y = Au + AGy = A(u+Gy)$$
ist $u + Gy \in L(A,0)$ und
$$u = (u+Gy) - Gy = (u+Gy) - \sum_{i=1}^{k} y_i G_{-i} \in L(A,0) - \mathcal{K}(G_{-1},G_{-2},\dots,G_{-k})\,.$$

Falls der Rang von A kleiner als k ist, ergänzen wir die Spalten von A durch Standardspalten von $\mathbb{R}^{k\times 1}$ zu einer Matrix $(A|C) =: B \in \mathbb{R}^{k\times(n+s)}$, deren Spalten ein Erzeugendensystem von $\mathbb{R}^{k\times 1}$ bilden. Dann ist der Rang von B gleich k und
$$L(B,\leq 0) = L(B,0) - \mathcal{K}(H_{-1},H_{-2},\dots,H_{-k})\,,$$
wobei $H \in \mathbb{R}^{n\times k}$ so gewählt wird, dass $B\cdot H = I_k$ ist. Dann ist
$$L(A,\leq 0) = \{u \in \mathbb{R}^{n\times 1} \mid \begin{pmatrix} u \\ 0 \end{pmatrix} \in L(B,\leq 0)\} = L(B,\leq 0) \cap \bigcap_{i=1}^{s} \mathrm{Kern}(q_{n+i})\,,$$
wobei q_{n+i} die Projektion
$$q_{n+i} : \mathbb{R}^{(n+s)\times 1} \longrightarrow \mathbb{R}\,,\ x \longmapsto x_{n+i}$$
ist. Der Durchschnitt $L(B,\leq 0) \cap \bigcap_{i=1}^{s} \mathrm{Kern}(q_{n+i})$ wird mit Satz 130 berechnet.

Der Vorteil dieser Methode gegenüber dem Verfahren in Satz 133 ist, dass nur s-mal (anstatt k-mal) ein Kegel mit dem Kern einer linearen Funktion durchschnitten werden muss.

Auch Satz 130 kann verbessert werden, falls der Kegel L als Lösungsmenge $L(f_1,\ldots,f_m,\leq 0)$ eines homogenen Systems linearer Ungleichungen gegeben ist. Dann wird $L\cap \mathrm{Kern}(g)$ sogar von

$$\bigcup_{\ell=1}^{m}(E\cap \mathrm{Kern}(f_\ell))$$

erzeugt.
Beweisidee dazu: Wir verwenden die Bezeichnungen von Satz 130. Seien u_i,u_j so, dass $g(u_i)>0$ und $g(u_j)<0$ ist. Falls für alle ℓ gilt: aus $f_\ell(u_i)=0$ folgt $f_\ell(u_j)<0$, dann ist für alle ℓ

$$f_\ell(g(u_i)u_j-g(u_j)u_i)=g(u_i)f_\ell(u_j)-g(u_j)f_\ell(u_i)<0\,,$$

also liegt $g(u_i)u_j-g(u_j)u_i$ „im Inneren"von L.

Definition 134 : 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 135 : Polyeder sind konvex.

Beispiel 136 : Die konvexe Hülle einer Menge N ist konvex.
Denn: Seien $\sum_{i=0}^{\ell}c_iv_i$ und $\sum_{j=0}^{m}d_jw_j$ konvexe Linearkombinationen von Elementen in N und $s,t\in\mathbb{R}_{\geq 0}$ mit $s+t=1$. Wegen

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

ist $s(\sum_{i=0}^{\ell}c_iv_i)+t(\sum_{j=0}^{m}d_jw_j)\in \mathrm{conv}(N)$.

Satz 137 : *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 $\mathrm{conv}(N)$ einer Teilmenge N von V ist die (bezüglich Inklusion) kleinste konvexe Teilmenge von V, die N enthält.

Beweis: Induktion über ℓ.

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

$\ell>1$: Sei $\sum_{i=0}^{\ell}c_iv_i$ 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 138 : 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*.

Satz 139 : *Es seien $f_1,\dots,f_k \in V^*$ und $b_1,\dots,b_k \in \mathbb{R}$. Dann gibt es endliche Teilmengen E und S von V so, dass*

$$L(f_1,\dots,f_k,\leq,b_1,\dots,b_k) = conv(E) + \mathcal{K}(S)$$

ist. Also: Jedes Polyeder ist die Summe der konvexen Hülle einer endlichen Menge und eines endlich erzeugten Kegels.

Die Mengen E und S können wie folgt berechnet werden:

- *Sei*

$$\begin{array}{rrcll} \bar{f}_i & : V \times \mathbb{R} & \to & \mathbb{R} & , \quad 1 \leq i \leq k\ , \\ & (u,t) & \mapsto & f_i(u) - b_i t & \end{array}$$

  und

$$\begin{array}{rrcll} p & : V \times \mathbb{R} & \to & \mathbb{R} & . \\ & (u,t) & \mapsto & t & \end{array}$$

- *Berechne mit Satz 133 eine endliche Teilmenge M von $V \times \mathbb{R}$, die den Kegel $L(\bar{f}_1,\dots,\bar{f}_k,-p,\leq 0)$ erzeugt.*
- *Dann ist*

$$S := \big\{u \in V \mid (u,0) \in M\big\} \quad \text{und}$$
$$E := \big\{\frac{1}{t}u \mid (u,t) \in M,\ t \neq 0\big\}.$$

Beweis: Es seien E und S die mit dem angegebenen Verfahren berechneten Mengen. Es ist zu zeigen, dass

$$L(f_1,\dots,f_k,\leq b_1,\dots,b_k) = \mathrm{conv}(E) + \mathcal{K}(S)$$

ist.

$\subseteq$: Sei $x \in L(f_1,\dots,f_k,\leq b_1,\dots,b_k)$.
Dann ist $(x,1) \in L(\bar{f}_1,\dots,\bar{f}_k,-p,\leq 0)$, also ist $(x,1)$ eine nicht-negative Linearkombination der Elemente in M:

$$(x,1) = \sum_{(u,t)\in M} c_{u,t}(u,t) \text{ mit } c_{u,t} \in \mathbb{R}_{\geq 0}$$

Wegen

$$(x,1) = \sum_{\substack{(u,t)\in M\\ t\neq 0}} tc_{u,t}\big(\frac{1}{t}u,1\big) + \sum_{(u,0)\in M} c_{u,0}(u,0)$$

ist

$$x = \sum_{\substack{(u,t)\in M\\ t\neq 0}} tc_{u,t}(\frac{1}{t}u) + \sum_{u\in S} c_{u,0}u$$

und

$$\sum_{\substack{(u,t)\in M\\ t\neq 0}} tc_{u,t} = 1.$$

Somit ist $x \in \operatorname{conv}(E) + \mathcal{K}(S)$.

$\supseteq$: Sei $v := \sum_{e\in E} c_e e + \sum_{s\in S} d_s s$ mit $c_e, d_s \in \mathbb{R}_{\geq 0}$ und $\sum_{e\in E} c_e = 1$, also $v \in \operatorname{conv}(E) + \mathcal{K}(S)$. Seien $e \in E$ und $s \in S$. Wegen

$$(s,0) \in L(\bar{f}_1,\dots,\bar{f}_k,-p,\leq 0)$$

ist $f_i(s) \leq 0$, $1 \leq i \leq k$, und wegen

$$(e,1) \in L(\bar{f}_1,\dots,\bar{f}_k,-p,\leq 0) = \mathcal{K}(M)$$

ist $f_i(e) \leq b_i$, $1 \leq i \leq k$. Daher ist

$$\begin{aligned} f_i(v) = {} & \sum_{e\in E} c_e f_i(e) + \sum_{s\in S} d_s f_i(s) \leq \\ & \leq \Big(\sum_{e\in E} c_e\Big) b_i + \sum_{s\in S} d_s f_i(s) \leq b_i, \quad 1 \leq i \leq k\ , \end{aligned}$$

also $v \in L(f_1,\dots,f_k,\leq b_1,\dots,b_k)$.

KAPITEL 4

Vektorräume mit Skalarprodukt

In diesem Kapitel seien K der Körper der reellen oder komplexen Zahlen und V ein Vektorraum über K.

Wir werden die folgende Eigenschaft der reellen Zahlen verwenden: Zu jeder reellen Zahl $a \geq 0$ gibt es genau eine reelle Zahl $d \geq 0$ mit $d^2 = a$. Schreibweise: $d =: \sqrt{a}$. Sind a und b reelle Zahlen mit $0 \leq a < b$, dann ist auch $\sqrt{a} < \sqrt{b}$.

§1. Skalarprodukte

Definition 140 : Für eine komplexe Zahl $z = \mathrm{Re}(z) + \mathrm{Im}(z)i$ heißt

$$\overline{z} := \mathrm{Re}(z) - \mathrm{Im}(z)i$$

die zu z *konjugierte komplexe Zahl*.
Für eine Matrix $A \in \mathbb{C}^{m \times n}$ heißt

$$\overline{A} := \left(\overline{A_{ij}}\right)_{i,j}$$

die zu A *konjugierte Matrix*.

Hilfssatz 141 : *Für komplexe Zahlen w und z ist*

$$\overline{w+z} = \overline{w} + \overline{z} \quad \textit{und} \quad \overline{w \cdot z} = \overline{w} \cdot \overline{z} \; .$$

Für Matrizen $A, B \in \mathbb{C}^{m \times n}$ und $C \in \mathbb{C}^{n \times p}$ ist

$$\overline{A+B} = \overline{A} + \overline{B} \quad \textit{und} \quad \overline{B \cdot C} = \overline{B} \cdot \overline{C} \; .$$

Beweis: Übung.

Definition 142 : Ein *Skalarprodukt* auf V ist eine Funktion

$$\langle -,- \rangle : V \times V \longrightarrow K \; , \quad (v,w) \longmapsto \langle v,w \rangle \; ,$$

mit den folgenden Eigenschaften:

Für alle $c \in K$, $u,v,w \in V$ gilt

(1) $\langle v,w \rangle = \overline{\langle w,v \rangle}$
   („$\langle -,- \rangle$ ist *hermitesch*“)
(2) $\langle u, c(v+w) \rangle = c\langle u,v \rangle + c\langle u,w \rangle$
   („$\langle -,- \rangle$ ist in der zweiten Komponente linear“)

(3) Für $v \neq 0$ ist $\langle v, v\rangle$ eine positive reelle Zahl, („$\langle -,-\rangle$ ist *positiv definit*").

Aus (1) und (2) folgt:

$$\langle c(u+v), w\rangle = \overline{c}\langle u, w\rangle + \overline{c}\langle v, w\rangle \ .$$

Wenn $K = \mathbb{R}$ ist, dann ist

$$\langle v, w\rangle = \langle w, v\rangle$$

(„$\langle -,-\rangle$ ist *symmetrisch*") und $\langle -,-\rangle$ ist in beiden Komponenten linear („$\langle -,-\rangle$ ist bilinear").

Definition 143 : Ein reeller bzw. komplexer Vektorraum mit einem Skalarprodukt heißt *reeller* bzw. *komplexer Prähilbertraum*. Ein endlich-dimensionaler reeller bzw. komplexer Prähilbertraum heißt *euklidischer* bzw. *unitärer Raum*.

Definition 144 : Ist $\langle -,-\rangle$ ein Skalarprodukt auf V, dann heißen die Funktionen

$$||\cdot|| : V \longrightarrow \mathbb{R}\ , \quad v \longmapsto ||v|| := \sqrt{\langle v, v\rangle}\ ,$$

bzw.

$$d : V \times V \longrightarrow \mathbb{R}\ , \quad (v, w) \longmapsto ||v-w|| := \sqrt{\langle v-w, v-w\rangle}\ ,$$

die von $\langle -,-\rangle$ induzierte *Norm* bzw. *Metrik* auf V. Die Zahl $d(v,w)$ heißt *Abstand* zwischen v und w. Die Zahl $||v|| = d(v,0)$ heißt *Abstand zwischen v und* 0, *Norm*, *Betrag* oder *Länge* von v.

Zwei Vektoren v, w stehen *zueinander senkrecht* oder *orthogonal*, wenn $\langle v, w\rangle = 0$ ist.

Beispiel 145 : Die Funktion

$$\langle -,-\rangle : K^n \times K^n \longrightarrow K\ , \quad ((a_1, \ldots, a_n), (b_1, \ldots, b_n)) \longmapsto \sum_{i=1}^{n} \overline{a_i} b_i\ ,$$

ist ein Skalarprodukt auf K^n und heißt *Standardskalarprodukt* auf K^n. Für die Standardbasis $(e_1, \ldots, e_n)$ von K^n gilt

$$\langle e_i, e_j\rangle = \delta_{ij},\ \|e_i\| = 1 \text{ und } \|e_i - e_j\| = \sqrt{2}(1-\delta_{ij}),\ \text{für } 1 \leq i, j \leq n\,.$$

Im Spezialfall $n = 1$ und $K = \mathbb{C}$ ist

$$\langle -,-\rangle : \mathbb{C} \times \mathbb{C} \longrightarrow \mathbb{C}\ , \quad (a, b) \longmapsto \overline{a}b\ ,$$

das Standardskalarprodukt und

$$|z| := \|z\| = \sqrt{\overline{z}z} = \sqrt{\mathrm{Re}(z)^2 + \mathrm{Im}(z)^2}$$

heißt *Betrag* der komplexen Zahl z. Für $w, z \in \mathbb{C}$ ist

$$|w \cdot z| = |w| \cdot |z| \ \text{ und } |z| = |\overline{z}|\ .$$

Beispiel 146 : Es seien a, b reelle Zahlen mit $a < b$ und V der Vektorraum aller stetigen Funktionen vom Intervall $[a, b]$ nach $\mathbb{C}$. Die Abbildung

$$\langle -, - \rangle : V \times V \longrightarrow \mathbb{C} \ , \quad (f, g) \longmapsto \int_a^b \overline{f(x)} g(x) dx \ ,$$

ist ein Skalarprodukt auf V. Die Norm von $f \in V$ ist

$$\|f\| = \sqrt{\int_a^b |f(x)|^2 dx} \quad .$$

Satz 147 : *Es seien V mit $\langle -, - \rangle$ ein Prähilbertraum und v, w Vektoren in V. Dann ist*

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

(„Cauchy-Schwarz'sche Ungleichung").
Weiters sind die Zahlen $|\langle v, w \rangle|$ und $\|v\| \cdot \|w\|$ genau dann gleich, wenn v und w linear abhängig sind.

Beweis: Wenn $v = 0$ oder $w = 0$ ist, dann ist $|\langle v, w \rangle| = 0 = \|v\| \cdot \|w\|$.
Seien nun $v \neq 0$ und $w \neq 0$. Wenn v und w linear abhängig sind, gibt es ein $c \in K$ mit $w = c \cdot v$. Daher ist

$$|\langle v, w \rangle| = |c| \langle v, v \rangle = |c| \cdot \|v\|^2 = \|v\| \cdot \|w\| \ .$$

Wenn v und w linear unabhängig sind, dann ist

$$0 \neq w - (\langle v, w \rangle / \langle v, v \rangle) v$$

und

$$0 < \langle w - (\langle v, w \rangle / \langle v, v \rangle) v, w - (\langle v, w \rangle / \langle v, v \rangle) v \rangle = \langle w, w \rangle - (|\langle v, w \rangle|^2 / \langle v, v \rangle) \ .$$

Daher ist

$$|\langle v, w \rangle|^2 < \|v\|^2 \cdot \|w\|^2 \ .$$

Beispiel 148 : Für $V = \mathbb{C}^n$ mit dem Standardskalarprodukt und $a, b \in \mathbb{C}^n$ ergibt sich

$$\left| \sum_{i=1}^n \overline{a_i} b_i \right| \leq \sqrt{\sum_{i=1}^n |a_i|^2} \sqrt{\sum_{i=1}^n |b_i|^2}$$

als Cauchy-Schwarz'sche Ungleichung.

Satz 149 : *Ist V mit $\langle -, - \rangle$ ein unitärer Raum, dann ist V mit*

$$\mathrm{Re}(\langle -, - \rangle) : V \times V \to \mathbb{R} \quad , \quad (v, w) \longmapsto \mathrm{Re}(\langle v, w \rangle) \ ,$$

ein euklidischer Raum (mit $\dim_{\mathbb{R}}(V) = 2 \cdot \dim_{\mathbb{C}}(V)$) und die von $\langle -, - \rangle$ und $\mathrm{Re}(\langle -, - \rangle)$ induzierten Normen sind gleich.

Beweis: Übung.

§2. Orthonormalbasen

Sei $\langle -,- \rangle$ ein Skalarprodukt auf V.

Definition 150 : Eine Familie $(v_i)_{i\in I}$ in V heißt *orthonormal* bezüglich $\langle -,- \rangle$, wenn für alle $i,j \in I$ gilt:

$$\langle v_i, v_j \rangle = \delta_{ij}.$$

Eine Familie $(v_i)_{i\in I}$ in V heißt *Orthonormalbasis* (kurz: *ON-Basis*) von V bezüglich $\langle -,- \rangle$, wenn sie eine Basis von V und orthonormal bezüglich $\langle -,- \rangle$ ist.

Beispiel 151 : Die Standardbasis von K^n ist eine Orthonormalbasis bezüglich des Standardskalarproduktes.

Satz 152 : *Eine orthonormale Familie ist linear unabhängig.*
Insbesondere: Wenn V endlich-dimensional ist, dann ist jede orthonormale Familie mit $\dim_K(V)$ *Elementen eine ON-Basis von V.*

Beweis: Sei $(v_i)_{i\in I}$ eine orthonormale Familie und $(c_i)_{i\in I}$ eine Koeffizienten-Familie in K. Wenn $\sum_{i\in I} c_i v_i = 0$ ist, dann ist für alle $j \in I$ auch

$$0 = \langle v_j, \sum_{i\in I} c_i v_i \rangle = \sum_{i\in I} c_i \langle v_j, v_i \rangle = c_j \ .$$

Satz 153 : *Sei $w \in V$ und $(v_i)_{i\in I}$ eine ON-Basis von V. Dann ist*

$$w = \sum_{i\in I} \langle v_i, w \rangle v_i \ .$$

(„Die Koordinate von w bei v_i ist das Skalarprodukt von v_i mit w“*.)*

Beweis: Sei $w = \sum_{i\in I} c_i v_i$. Dann ist

$$\langle v_j, w \rangle = \langle v_j, \sum_{i\in I} c_i v_i \rangle = \sum_{i\in I} c_i \langle v_j, v_i \rangle = \sum_{i\in I} c_i \delta_{ji} = c_j \ .$$

Satz 154 : *Es sei $(w_1,\dots,w_n)$ eine Basis von V. Mit dem folgenden Verfahren („Schmidt'sches Orthonormalisierungsverfahren“) kann eine ON-Basis $(v_1,\dots,v_n)$ von V berechnet werden:*

- $u_1 := w_1$

- *Für* $2 \leq j \leq n$ *sei*

$$u_j := w_j - \sum_{i=1}^{j-1} (\langle u_i, w_j \rangle / \langle u_i, u_i \rangle) u_i$$

- *Für* $1 \leq j \leq n$ *sei*

$$v_j := \|u_j\|^{-1} u_j \,.$$

- $v_1 := \|u_1\|^{-1} u_1$

Insbesondere: Jeder endlichdimensionale Prähilbertraum hat eine ON-Basis. Für alle j *ist* ${}_K\langle v_1, \ldots, v_j \rangle = {}_K\langle w_1, \ldots, w_j \rangle$.

Beweis: Nach Definition ist $\langle v_i, v_i \rangle = 1$, $1 \leq i \leq n$. Es genügt also zu zeigen, dass für alle $1 \leq k < \ell \leq n$ die Vektoren u_k und u_ℓ zueinander senkrecht stehen. Das kann einfach nachgerechnet werden:

$$\langle u_k, u_\ell \rangle = \langle u_k, w_\ell - \sum_{i=1}^{\ell-1} (\langle u_i, w_\ell \rangle / \langle u_i, u_i \rangle) u_i \rangle =$$

$$= \langle u_k, w_\ell \rangle - \sum_{i=1}^{\ell-1} (\langle u_i, w_\ell \rangle / \langle u_i, u_i \rangle) \delta_{ki} \langle u_k, u_k \rangle = \langle u_k, w_\ell \rangle - \langle u_k, w_\ell \rangle = 0 \,.$$

Nach Satz 152 ist dann $(v_1, \ldots, v_n)$ linear unabhängig, daher wegen $\dim_K(V) = n$ auch eine Basis.

Definition 155 : Eine Matrix $A \in \mathbb{R}^{n \times n}$ heißt *orthogonal*, wenn

$$A \cdot A^T = I_n$$

ist. Eine Matrix $A \in \mathbb{C}^{n \times n}$ heißt *unitär*, wenn

$$A \cdot \overline{A}^T = I_n$$

ist.
Wir schreiben $\mathcal{O}_n$ bzw. $\mathcal{U}_n$ für die Menge aller orthogonalen bzw. unitären $n \times n$-Matrizen.

Definition 156 : G mit $G \times G \longrightarrow G$, $(a,b) \longmapsto a \cdot b$, sei eine Gruppe. Eine *Untergruppe* von G ist eine nicht-leere Teilmenge H von G mit den Eigenschaften: Wenn $a \in H$ ist, dann auch a^{-1}. Wenn $a, b \in H$ sind, dann auch $a \cdot b$. Schreibweise: $H \leq G$.

Satz 157 : *Die Mengen* $\mathcal{O}_n$ *bzw.* $\mathcal{U}_n$ *sind Untergruppen von* $GL_n(\mathbb{R})$ *bzw.* $GL_n(\mathbb{C})$ *und heißen* orthogonale *bzw.* unitäre Gruppe.
Es ist $\mathcal{O}_n = GL_n(\mathbb{R}) \cap \mathcal{U}_n$.

Beweis: Übung.

Satz 158: *Es seien $\underline{v} := (v_1, \ldots, v_n)$ eine ON-Basis von V, $S \in K^{n\times n}$ und $\underline{u} := (u_1, \ldots, u_n) = \underline{v}S$. Dann ist*

$$S = (\langle v_i, u_j \rangle)_{1\leq i,j\leq n}$$

und die folgenden Aussagen sind äquivalent:

(1) *$\underline{u}$ ist eine ON-Basis von V;*
(2) *die Matrix S ist orthogonal (wenn $K = \mathbb{R}$ ist) bzw. unitär (wenn $K = \mathbb{C}$ ist);*
(3) *die Spalten von S bilden eine ON-Basis von $K^{n\times 1}$ mit dem Standardskalarprodukt;*
(4) *die Zeilen von S bilden eine ON-Basis von $K^{1\times n}$ mit dem Standardskalarprodukt.*

Beweis: Mit Satz 153 folgt

$$\langle u_i, u_j \rangle = \langle \underline{v}S_{-i}, \underline{v}S_{-j} \rangle = \sum_{k,\ell} \overline{S}_{ki} S_{\ell j} \langle v_k, v_\ell \rangle =$$

$$= \sum_{k,\ell} \overline{S}_{ki} S_{\ell j} \delta_{k\ell} = \sum_{k} \overline{S}_{ki} S_{kj} = \langle S_{-i}, S_{-j} \rangle = (\overline{S}^T \cdot S)_{ij} ,$$

damit ist die Behauptung leicht nachzuprüfen.

Definition 159: Es seien U ein endlich-dimensionaler Untervektorraum von V. Dann ist

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

ein Untervektorraum von V und heißt *das orthogonale Komplement von U in V*.

Satz 160: *Es seien U ein endlich-dimensionaler Untervektorraum von V, $v \in V$ und $(u_1, \ldots, u_n)$ eine ON-Basis von U.*

(1) *Der Vektor*

$$p_U(v) := \sum_{i=1}^{n} \langle u_i, v \rangle u_i \in U$$

hängt nicht von der Wahl der ON-Basis $(u_1, \ldots, u_n)$ ab und heißt Fußpunkt des Lotes von v auf U. *Die Funktion*

$$p_U : V \longrightarrow V \,,\; v \longmapsto p_U(v) \,,$$

heißt orthogonale Projektion *von V auf U.*

(2) *Jeder Vektor $v \in V$ lässt sich eindeutig als Summe eines Vektors in U und eines Vektors in $U^\perp$ schreiben, und zwar*

$$v = p_U(v) + (v - p_U(v)) \,,$$

wobei $p_U(v) \in U$ und $(v - p_U(v)) \in U^\perp$ ist. Insbesondere ist $U \cap U^\perp = \{0\}$ und $V = U + U^\perp := \{y + y' \,|\, y \in U,\, y' \in U^\perp\}$.

(3) *Für $v \in V$ und $y \in U$ mit $p_U(v) \neq y$ ist*

$$\|v - y\| > \|v - p_U(v)\| \,,$$

das heißt: $p_U(v)$ ist der eindeutig bestimmte Vektor in U, der von v den kleinsten Abstand hat.

Beweis:

(1) Es sei $(w_1, \ldots, w_n)$ eine ON-Basis von U. Es ist zu zeigen, dass $p_U(v) = \sum_{k=1}^n \langle w_k, v\rangle w_k \in U$ ist. (Dann hängt $p_U(v)$ nicht von der Wahl der ON-Basis in U ab). Nach Satz 153 ist

$$u_i = \sum_{j=1}^n \langle w_j, u_i\rangle w_j \,,\ 1 \leq i \leq n \,.$$

Daher ist

$$p_U(v) = \sum_{i=1}^n \langle \sum_{j=1}^n \langle w_j, u_i\rangle w_j, v\rangle \sum_{k=1}^n \langle w_k, u_i\rangle w_k =$$

$$= \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \langle w_j, u_i\rangle \langle w_j, v\rangle \langle w_k, u_i\rangle w_k =$$

$$= \sum_{k=1}^n (\sum_{j=1}^n \langle w_j, v\rangle \sum_{i=1}^n (\langle w_j, u_i\rangle \langle w_k, u_i\rangle)) w_k = (\text{ cf. Satz 153 })$$

$$= \sum_{k=1}^n (\sum_{j=1}^n \langle w_j, v\rangle \langle w_j, w_k\rangle) w_k = \sum_{k=1}^n \langle w_k, v\rangle) w_k \,.$$

(2) Für $y \in U$ ist $\langle v, y\rangle = \langle p_U(v), y\rangle$, daher

$$\langle v - p_U(v), y\rangle = \langle v, y\rangle - \langle p_U(v), y\rangle = 0 \,,$$

daher ist $v - p_U(v) \in U^\perp$ und $v = p_U(v) + (v - p_U(v)) \in U + U^\perp$. Wenn $y \neq 0$ ist, dann ist $0 < \langle y, y\rangle$, also $y \notin U^\perp$. Daher ist $U \cap U^\perp = \{0\}$.

(3) Für $y \in U$ mit $p_U(v) \neq y$ ist $0 \neq p_U(v) - y \in U$ und $v - p_U(v) \in U^\perp$. Deshalb ist

$$\|v - y\|^2 = \|(v - p_U(v)) + (p_U(v) - y)\|^2 =$$

$$= \|v - p_U(v)\|^2 + \|p_U(v) - y\|^2 > \|v - p_U(v)\|^2 \,.$$

Beispiel 161: Es seien $0 \neq u \in V$ und U die Gerade Ku. Dann ist $\|u\|^{-1}u$ eine ON-Basis von U. Der Fußpunkt des Lotes von $v \in V$ auf die Gerade U ist

$$p_{Ku}(v) = \langle \|u\|^{-1}u, v\rangle \|u\|^{-1}u = (\langle u,v\rangle / \langle u,u\rangle)u\,.$$

Definition 162: Es seien Z ein endlichdimensionaler affiner Unterraum von V mit Aufpunkt z und parallelem Untervektorraum U. Der Vektor

$$p_Z(v) := z + p_U(v-z)$$

heißt *Fußpunkt des Lotes* von v auf den affinen Unterraum Z. Die Funktion

$$p_Z : V \longrightarrow V\,,\ v \longmapsto p_Z(v)\,,$$

heißt *orthogonale Projektion* von V auf Z. Die Zahl $\|v - p_Z(v)\|$ heißt *Abstand des Punktes v vom affinen Unterraum Z.*

Satz 163: *Es seien Z und Z' endlichdimensionale affine Unterräume von V mit Aufpunkten z, z' und parallelen Untervektorräumen U, U'. Dann gibt es Elemente $v \in Z$, $v' \in Z$ so, dass für alle $w \in Z$, $w' \in Z$ gilt:*

$$\|w - w'\| \geq \|v - v'\|\ .$$

Die Zahl $\|v - v'\|$ heißt Abstand der affinen Unterräume Z und Z' *und ist gleich dem Abstand des Punktes $z - z'$ vom Untervektorraum $U + U'$.*

Beweis: Es seien $u - u' \in U + U'$ der Fußpunkt des Lotes von $z - z'$ auf $U + U'$, $v := z - u$ und $v' := z' - u'$. Für alle $x \in U$, $x' \in U'$ ist dann

$$\|(z-x)-(z'-x')\| = \|(z-z')-(x-x')\| \geq \|(z-z')-(u-u')\| = \|v-v'\|\ .$$

§3. Lineare Gleichungen mit ungenau bestimmten Daten

Es sei $\langle -,- \rangle$ ein Skalarprodukt auf V, $y \in V$, U ein K-Vektorraum und $f : U \longrightarrow V$ eine lineare Funktion.

Wenn $y \notin \mathrm{Bild}(f) =: W$ ist, dann hat das durch f und y gegebene System linearer Gleichungen keine Lösung. Wenn etwa y durch Messungen (mit Fehlern) bestimmt wurde, kann $L(f,y)$ leer sein, obwohl bei exakter Messung eine Lösung existiert. In diesem Fall legt Satz 160 nahe, y durch den Fußpunkt des Lotes von y auf W zu ersetzen und dann $L(f, p_W(y))$ zu berechnen.

Beispiel 164: „Methode der kleinsten Quadrate“

Gegeben sind ein endlichdimensionaler Untervektorraum U des Vektorraums $\mathcal{F}(\mathbb{R},\mathbb{R})$, paarweise verschiedene reelle Zahlen $x_1, \ldots, x_n$ und reelle Zahlen $y_1, \ldots, y_n$. Gesucht ist eine Funktion $g \in U$ so, dass die Zahlen $g(x_i)$ „möglichst nahe“ bei y_i liegen, $1 \leq i \leq n$.

Wir betrachten dazu die lineare Funktion

$$f : U \longrightarrow \mathbb{R}^n \quad , \quad h \longmapsto (h(x_1), \ldots, h(x_n)) \ .$$

Die gesuchte Funktion ist ein Urbild von $y := (y_1, \ldots, y_n)$ bezüglich f. Sei $\underline{u} := (u_1, \ldots, u_m)$ eine Basis von U, dann wird das Bild W von f von den n-Tupeln $f(u_1), \ldots, f(u_m)$ erzeugt. Wir wählen auf $\mathbb{R}^n$ das Standardskalarprodukt. Dann ist der Abstand von y zu jedem Punkt von W größer oder gleich dem Abstand von y zu $p_W(y)$, also ist die gesuchte Funktion jenes Element $g \in U$ mit

$$\sqrt{\sum_{i=1}^{n} (g(x_i) - y_i)^2} \leq \sqrt{\sum_{i=1}^{n} (h(x_i) - y_i)^2} \ ,$$

für alle $h \in U$. Die Summe $\sum_{i=1}^{n} (g(x_i) - y_i)^2$ der Quadrate der „Fehler in x_i" ist daher für g am kleinsten.
Wir betrachten zwei Spezialfälle:

(1) Sei $U := \mathrm{Lin}_{\mathbb{R}}(\mathbb{R}, \mathbb{R})$.

Eine Basis von U ist $(Id_{\mathbb{R}})$. Das Bild von f ist die Gerade durch Null und $f(\mathrm{Id}_{\mathbb{R}}) = (x_1, \ldots, x_n) =: x$. Eine ON-Basis von W ist $(\frac{1}{\|x\|} x)$. Daher:

$$p_W(y) = \langle \frac{1}{\|x\|} x, y \rangle \frac{1}{\|x\|} x = \frac{\langle x, y \rangle}{\langle x, x \rangle} x$$

und das Urbild von $p_W(y)$ bezüglich f ist

$$g : \mathbb{R} \longrightarrow \mathbb{R} \quad , \quad \alpha \longmapsto \frac{\langle x, y \rangle}{\langle x, x \rangle} \alpha \ .$$

(2) Sei U der Vektorraum aller affinen Funktionen von $\mathbb{R}$ nach $\mathbb{R}$. Eine Basis von U ist $(E, Id_{\mathbb{R}})$, wobei E die konstante Funktion $\mathbb{R} \longrightarrow \mathbb{R}$, $x \longmapsto 1$, bezeichnet. Das Bild von f ist die von $(1, \ldots, 1) =: \mathbf{1}$ und $(x_1, \ldots, x_n) =: x$ erzeugte Ebene. Eine ON-Basis von W ist $(\frac{\sqrt{n}}{n} \mathbf{1}, \frac{1}{\|z\|} z)$, wobei $z := x - \frac{\langle x, \mathbf{1} \rangle}{n} \mathbf{1}$. Daher ist

$$p_W(y) = \frac{\langle \mathbf{1}, y \rangle}{n} \mathbf{1} + \frac{\langle z, y \rangle}{\langle z, z \rangle} z = \left(\frac{\langle \mathbf{1}, y \rangle}{n} - \frac{\langle \mathbf{1}, x \rangle}{n} \frac{\langle z, y \rangle}{\langle z, z \rangle} \right) \mathbf{1} + \frac{\langle z, y \rangle}{\langle z, z \rangle} x$$

und das Urbild von $p_W(y)$ bezüglich f ist

$$g : \mathbb{R} \longrightarrow \mathbb{R} \quad , \quad \alpha \longmapsto \left(\frac{\langle \mathbf{1}, y \rangle}{n} - \frac{\langle \mathbf{1}, x \rangle}{n} \frac{\langle z, y \rangle}{\langle z, z \rangle} \right) + \frac{\langle z, y \rangle}{\langle z, z \rangle} \alpha \ .$$

Der Graph von g heißt *Regressionsgerade* der Punkte $(x_1, y_1), \ldots, (x_n, y_n)$.

§4. Parallelprojektion

Es sei V ein orientierter dreidimensionaler euklidischer Raum.

Satz 165 : *Es seien E eine Ebene in V und G eine Gerade in V, die nicht parallel sind. Für $v \in V$ sei $\Pi(v)$ der Schnittpunkt der zu G parallelen Geraden durch v mit der Ebene E. Dann ist die Funktion*

$$\Pi : V \longrightarrow V \quad , \quad v \longmapsto \Pi(v) \ ,$$

affin und heißt die Parallelprojektion auf E längs G.
Wenn $0 \in E$ ist, dann ist Π eine lineare Funktion, deren Kern der zu G parallele Untervektorraum ist.

Beweis: Es seien U und W die zu G und E parallelen Untervektorräume von V. Weil G und E nicht parallel sind, ist V die direkte Summe von U und W. Die Funktion

$$f : V = U \oplus W \longrightarrow V \quad , \quad u + w \longmapsto w \ ,$$

ist linear und für $v \in V$ und $p \in E$ ist $t_p \circ f \circ t_{-p}(v) = p + f(v-p) \in E$. Wegen $p + f(v-p) = v - ((v-p) - f(v-p)) \in v + U$ ist $t_p \circ f \circ t_{-p}(v)$ der Schnittpunkt der zu G parallelen Geraden durch v mit E, also ist Π die affine Funktion $t_p \circ f \circ t_{-p}$.

Beispiel 166 : Die orthogonale Projektion p_W von V auf einen zweidimensionalen Untervektorraum W ist die Parallelprojektion auf W längs $W^{\perp}$.

Beispiel 167 : Seien $a, b \in \mathbb{R}$. Dann ist die Funktion

$$\mathbb{R}^{3\times 1} \longrightarrow \mathbb{R}^{3\times 1} \quad , \quad x \longmapsto \begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 0 \end{pmatrix} \cdot x \ ,$$

die Parallelprojektion auf $\{x \in \mathbb{R}^{3\times 1} \mid x_3 = 0\}$ längs $\mathbb{R} \begin{pmatrix} a \\ b \\ -1 \end{pmatrix}$.

Satz 168 : *Eine affine Funktion von V nach V ist genau dann eine Parallelprojektion auf eine Ebene E, wenn E sowohl ihr Bild als auch ihre Fixmenge ist.*

Beweis: Übung.

Hilfssatz 169 : *Es seien x, y linear unabhängige Vektoren in V. Dann gibt es genau eine positive reelle Zahl c und mindestens ein, aber höchstens zwei Paare von Vektoren $(v, w) \in V^2$ so, dass die folgenden Bedingungen erfüllt sind:*

$$\langle v, x\rangle = 0 , \quad \langle w, y\rangle = 0 , \quad \langle v, w\rangle = 0 ,$$
$$\|v\| = 1 = \|w\| \quad \textit{und} \quad \langle v, y\rangle = c = \langle w, x\rangle .$$

Beweis: Es seien

$$u_1 := \frac{x}{\|x\|} , \quad d := \|y - \frac{\langle x, y\rangle}{\|x\|} \cdot u_1\| , \quad u_2 := \frac{1}{d}\left(y - \frac{\langle x, y\rangle}{\|x\|} \cdot u_1\right)$$

und u_3 ein Vektor der Länge 1, der zu u_1 und u_2 orthogonal ist. Dann ist (u_1, u_2, u_3) eine ON-Basis von V,

$$d^2 = \langle y, y\rangle - \frac{\langle x, y\rangle^2}{\langle x, x\rangle} , \quad x = \|x\| \cdot u_1 \quad \text{und} \quad y = \frac{\langle x, y\rangle}{\|x\|} u_1 + du_2 .$$

Es seien $v = \sum_{i=1}^3 a_i u_i$ und $w = \sum_{i=1}^3 b_i u_i$ die gesuchten Vektoren.
Aus $\langle v, x\rangle = 0$, $\langle v, y\rangle = c$, und $\|v\| = 1$ folgt

$$a_1 = 0 , \quad a_2 = \frac{c}{d} , \quad a_3^2 = 1 - \frac{c^2}{d^2} ,$$

insbesondere muss $c \leq d$ sein.
Aus $\langle w, y\rangle = 0$, $\langle w, x\rangle = c$, und $\|w\| = 1$ folgt

$$b_1 = \frac{c}{\|x\|} , \quad b_2 = -\frac{\langle x, y\rangle}{\langle x, x\rangle} \cdot \frac{c}{d} , \quad b_3^2 = \frac{d^2\langle x, x\rangle - c^2\langle y, y\rangle}{d^2 \cdot \langle x, x\rangle} ,$$

insbesondere muss $c \cdot \|y\| \leq d \cdot \|x\|$ sein.
Die Bedingung $\langle v, w\rangle = 0$ bedeutet dann, dass

$$c^4 - (\langle x, x\rangle + \langle y, y\rangle)c^2 + \langle x, x\rangle d^2 = 0$$

ist und dass die Zahlen $\langle x, y\rangle$ und $a_3 \cdot b_3$ entweder beide Null sind oder gleiches Vorzeichen haben. Man kann nun nachprüfen, dass es genau eine positive reelle Zahl c mit $c \cdot \|y\| \leq d \cdot \|x\|$ und $c \leq d$ gibt, die diese Bedingung erfüllt.
Ist $(a_1u_1 + a_2u_2 + a_3u_3, b_1u_1 + b_2u_2 + b_3u_3)$ ein Paar von Vektoren, das die angegebenen Bedingungen erfüllt, dann ist $(a_1u_1 + a_2u_2 - a_3u_3, b_1u_1 + b_2u_2 - b_3u_3)$ das einzige andere.

Satz 170 : *(„Hauptsatz der Axonometrie“)*
Es seien W ein zweidimensionaler Untervektorraum von V und (w_1, w_2, w_3) ein Erzeugendensystem von W. Dann gibt es eine positive reelle Zahl c, eine positiv orientierte ON-Basis $\underline{v} := (v_1, v_2, v_3)$ und eine Parallelprojektion Π von V auf W so, dass für $1 \leq i \leq 3$

$$\Pi(cv_i) = w_i$$

ist. Die Zahl c ist eindeutig bestimmt, für die positiv orientierte ON-Basis $\underline{v}$ gibt es eine oder zwei Möglichkeiten und die Parallelprojektion ist durch $\underline{v}$ und c eindeutig bestimmt.

Beweis: Es sei $\underline{u}$ eine positiv orientierte ON-Basis von V so, dass (u_1, u_2) eine ON-Basis von W ist. Die Matrix einer Parallelprojektion von V auf W bezüglich $\underline{u}$ ist dann

$$\begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 0 \end{pmatrix} ,$$

wobei der Kern dieser Parallelprojektion die Gerade durch 0 und $au_1 + bu_2 - u_3$ ist.
Seien $x, y \in \mathbb{R}^{1\times 3}$ so, dass für $1 \leq i \leq 3$

$$w_i = x_i u_1 + y_i u_2$$

ist. Die Koordinatenspalten der gesuchten ON-Basis $\underline{v}$ bezüglich $\underline{u}$ bilden eine orthogonale Matrix mit Determinante 1. Wir suchen also eine orthogonale Matrix A und Zahlen $a, b \in \mathbb{R}$, $c \in \mathbb{R}_{>0}$ so, dass

$$c \cdot \begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \end{pmatrix} \cdot A = \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix}$$

ist. Nach Hilfssatz 169 gibt es ein, aber höchstens zwei Paare von zueinander orthogonalen 3−Spalten (s, t) der Länge 1 und genau eine positive reelle Zahl c so, dass

$$\begin{pmatrix} x \\ y \end{pmatrix} \cdot \begin{pmatrix} s & t \end{pmatrix} = \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix}$$

ist. Es gibt genau eine 3−Spalte z der Länge 1 so, dass die Matrix $\begin{pmatrix} s & t & z \end{pmatrix}$ orthogonal mit Determinante 1 ist.
Setze daher $A := \begin{pmatrix} s & t & z \end{pmatrix}^T = \begin{pmatrix} s & t & z \end{pmatrix}^{-1}$, $a := c^{-1} \sum_{i=1}^{3} x_i z_i$ und $b := c^{-1} \sum_{i=1}^{3} y_i z_i$.

Beispiel 171 : Es seien $\underline{u}$ eine positiv orientierte ON-Basis von V und W der von $w_1 := u_1$, $w_2 := u_2$, $w_3 := u_1 + u_2$ erzeugte Untervektorraum von V. Dann ist die gesuchte Zahl c gleich 1 und
entweder ist

$$(v_1, v_2, v_3) = (u_1, u_2, u_3)$$

und für alle $a_1, a_2, a_3 \in \mathbb{R}$ ist

$$\Pi(a_1u_1 + a_2u_2 + a_3u_3) = (a_1 + a_3)u_1 + (a_2 + a_3)u_2$$

oder

$$(v_1, v_2, v_3) = (u_1, u_2, u_3) \cdot \frac{1}{3} \begin{pmatrix} 1 & -2 & 2 \\ -2 & 1 & 2 \\ -2 & -2 & -1 \end{pmatrix}$$

und für alle $a_1, a_2, a_3 \in \mathbb{R}$ ist

$$\Pi(a_1u_1 + a_2u_2 + a_3u_3) = (a_1 - a_3)u_1 + (a_2 - a_3)u_2 \ .$$

KAPITEL 5

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 172 : 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 173 : *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 174 : *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, $w \in V$:

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

Satz 175 :

(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. *Insbesondere 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 173 und Satz 174 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 174, 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 176 :

(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 177 : *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 folgt die Behauptung aus Satz 152.

(2) $\Rightarrow$ (3) : Wenn $f(\underline{v}) = \underline{v}A$ eine *ON*-Basis ist, dann ist A orthogonal (Satz 158).

(3) $\Rightarrow$ (1) : Wenn A orthogonal ist, dann ist $\underline{v}A$ nach Satz 158 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 178 : *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 177 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 179 : Ein Untervektorraum W von V ist genau dann *unter f stabil*, wenn $f(W) \subseteq W$ ist.

Satz 180 : *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

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

Definition 181 : 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 $n_1 + \cdots + n_p = n$. 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

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

(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 $1 \leq k, \ell \leq p$ 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 $1 \leq k, \ell \leq p$ 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 $1 \leq k, \ell \leq p$ mit $k \neq \ell$ die Matrix $B_{k\ell} = 0$ ist.

Beispiel 182 : 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 183 : Es sei $M \in \mathbb{C}^{m\times n}$ eine komplexe Matrix. Dann heißt

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

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

Hilfssatz 184 :

(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 185 : *(„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 184 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 180). 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 186 : 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 187 : $Fix(-Id_V) = \{0\}, Fix(Id_V) = V$.
Die Fixmenge einer linearen Funktion ist ihr Eigenraum zum Eigenwert 1.
Die Fixmenge einer Translation t_u mit $u \neq 0$ ist leer.

Hilfssatz 188 : *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 Kern$(g - Id_V)$, also $Fix(g)$.

Definition 189 : 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 190 : 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 191 : *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 175 ist der lineare Anteil von s eine Isometrie, nach Hilfssatz 188 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 174 ist s orthogonal, daraus folgt mit Satz 177 , 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 192: *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 191.
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 193: *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 und s_U die Spiegelungen mit $Fix(s_W) = W$ und $Fix(s_U) = U$.

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 $s_1,...,s_k$ in U so, dass $(s_W \circ f)\,|_U = s_1 \circ ... \circ s_k$.

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 $s_W \circ f = s_1 \circ ... \circ s_k$ und $f = s_W \circ s_1 \circ ... \circ s_k$.

Definition 194 : 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 195 : *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 196:

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

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

(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 Einführung in die Mathematik 2.

Satz 197: *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 198 : 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 199 : *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 200 : Es seien $\alpha \in [0, 2\pi[$, $z \in V$ und $(-)^\perp$ die in Satz 197 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 201 :

(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*

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

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 197, (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 197, (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 202: 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 203: Drehungen in V sind orientierungserhaltend, Spiegelungen in V sind orientierungsumkehrend.

Satz 204: *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 188 ist $Fix(f)$ entweder leer oder ein affiner Unterraum mit parallelem Vektorraum $Fix(g)$.
Nach Satz 201, (2), und nach Satz 193 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 205: *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 206 : 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

$$d_\beta(d_\alpha(v)) = d_\beta(w) = (s_v \circ s_w)(w) = s_v(w) =$$

$$= 2\langle v, w\rangle v - w = 2\cos(\alpha)v - \Big(\cos(\alpha)v + \sin(\alpha)v^\perp\Big) =$$

$$= \cos(2\pi-\alpha)v + \sin(2\pi-\alpha)v^\perp = d_{2\pi-\alpha}(v) = d_\alpha^{-1}(v).$$

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 207 : 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 208 : 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 209 : 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

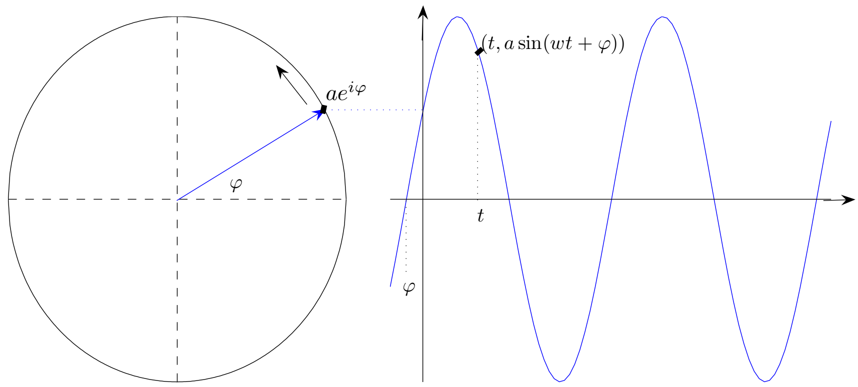
Ein Zeiger (zum Beispiel der Sekundenzeiger einer Uhr) der Länge a dreht sich in einer Ebene mit der Kreisfrequenz ω um einen Punkt $\mathcal{O}$. Wir identifizieren diese Ebene mit $\mathbb{C}=\mathbb{R}^2$ und $\mathcal{O}$ mit $0_\mathbb{C}=(0,0)$. Zur Zeit 0 sei die Spitze des Zeigers im Punkt $ae^{i\varphi}=(a\cos(\varphi),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

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

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 $ae^{i\varphi}$* 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

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

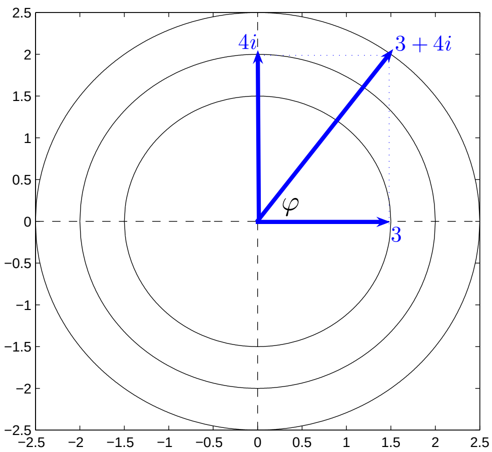
Wir erhalten daher die gesuchten Zahlen a und φ aus der Polardarstellung $z(0) = a(\cos(\varphi) + i\sin(\varphi)) = ae^{i\varphi}$ von $z(0)$, dabei ist a der Abstand zwischen $z(0)$ und 0 und φ der orientierte Winkel von $\mathbb{R}_{\geq 0} \cdot 1$ nach $\mathbb{R}_{\geq 0} \cdot z(0)$.

Beispiel 210 : 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 211 : Berechne reelle Zahlen a und φ so, dass für alle Zahlen $t \in \mathbb{R}$

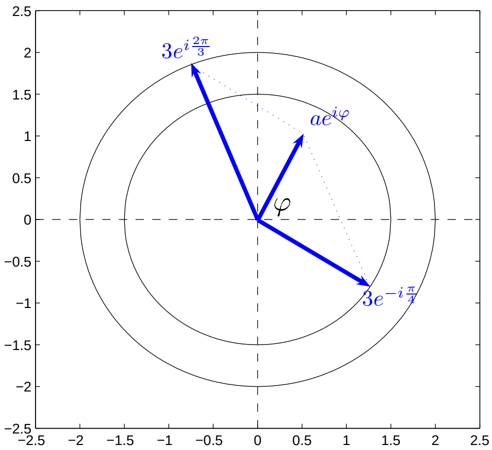
$$3 \cdot \sin(\omega t - \frac{\pi}{4}) + 4 \cdot \sin(\omega t + \frac{2\pi}{3}) = a \cdot \sin(\omega t + \varphi)$$

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

Definition 213: 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 214 :

(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 174 , dass g linear ist. Die Behauptung folgt daher aus Satz 201 .
(4) Folgt aus (3).

Satz 215 :

(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 185 gibt es einen 1-dimensionalen, unter g stabilen Untervektorraum W von V. Aus Satz 178 folgt, dass $g|_W = \mathrm{Id}_W$ oder $g|_W = -\mathrm{Id}_W$.

Nach Satz 180 ist auch $W^{\perp}$ stabil unter g. Nach Satz 204 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 195 eine Spiegelung, Gleitspiegelung oder Drehspiegelung.

Beispiel 216 : 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 217 : 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 218 : *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 $h_\beta(g_\gamma(w)) = v_1$. Die Funktion $h_\beta \circ g_\gamma \circ f^{-1}$ ist eine Drehung, wegen $(h_\beta \circ g_\gamma \circ f^{-1})(v_1) = v_1$ ist ihre Drehachse $\mathbb{R}v_1$. Setze $g_\alpha := (h_\beta \circ g_\gamma \circ f^{-1})^{-1}$.

§7. Symmetriegruppen

Definition 219: Es seien $M_1, ..., M_k$ Teilmengen von V. Dann heißt die Menge $Sym_V(M_1, ..., M_k) :=$

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

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

Satz 220:

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

Beweis:

(1) kann leicht nachgeprüft werden.
(2) Sei M_1 endlich. Nach Satz 106 ist für jede Symmetrieoperation f das Bild des Schwerpunktes von M_1 der Schwerpunkt des Bildes $f(M_1) = M_1$.

Satz 221 : *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 222 : 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 $Sym_V(M_1, M_2) =$

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

§8. Normale und selbstadjungierte Funktionen

Es seien K der Körper der reellen oder komplexen Zahlen, V ein euklidischer oder unitärer Raum und $\underline{v} = (v_1, \ldots, v_n)$ eine ON-Basis von V.

Definition 223 : Es seien f und g lineare Funktionen von V nach V. Die Funktion g heißt zu f *adjungiert*, wenn für alle $u, w \in V$

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

ist.

Beispiel 224 : Für $c \in \mathbb{C}$ ist $\overline{c} \cdot \mathrm{Id}_V$ zu $c \cdot \mathrm{Id}_V$ adjungiert.

Satz 225 : *Es seien f, g lineare Funktionen von V nach V und A, B ihre Matrizen bezüglich $\underline{v}$. Dann ist g genau dann zu f adjungiert, wenn*

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

ist. Insbesondere gibt es genau eine zu f adjungierte Funktion, wir bezeichnen sie mit f^*, und es ist $(f^*)^* = f$.

Beweis: Man prüft leicht nach, dass g genau dann zu f adjungiert ist, wenn für alle $1 \le i, j \le n$

$$\langle f(v_i), v_j \rangle = \langle v_i, g(v_j) \rangle$$

ist. Wegen

$$\langle f(v_i), v_j \rangle = \langle \sum_{k=1}^{n} A_{ki} v_k, v_j \rangle = \overline{A_{ji}} = (\overline{A}^\top)_{ij}$$

und

$$\langle v_i, g(v_j) \rangle = \langle v_i, \sum_{k=1}^{n} B_{kj} v_k \rangle = B_{ij}$$

folgt daraus die Behauptung.

Definition 226 : Eine lineare Funktion f von V nach V heißt *normal* bzw. *selbstadjungiert*, wenn

$$f \circ f^* = f^* \circ f \qquad \text{bzw.} \qquad f = f^*$$

ist. Eine Matrix $A \in K^{n \times n}$ heißt *symmetrisch* bzw. *normal* bzw. *selbstadjungiert*, wenn

$$A = A^\top \qquad \text{bzw.} \qquad A \cdot \overline{A}^\top = \overline{A}^\top \cdot A \qquad \text{bzw.} \qquad A = \overline{A}^\top$$

ist.

Eine reelle Matrix ist genau dann selbstadjungiert, wenn sie symmetrisch ist. Nach Satz 225 ist eine lineare Funktion genau dann normal bzw. selbstadjungiert, wenn ihre Matrix bezüglich einer ON-Basis von V normal bzw. selbstadjungiert ist.

Beispiel 227 : Jede orthogonale Funktion von V nach V ist normal.
Spiegelungen und orthogonale Projektionen auf Untervektorräume von V sind selbstadjungiert.
Diagonalmatrizen sind genau dann selbstadjungiert, wenn sie reell sind.

Hilfssatz 228: *Es seien f und g lineare Funktionen von V nach V und $c \in K$. Dann ist $(f+cg)^* = f^* + \overline{c}g^*$. Wenn f und g selbstadjungiert sind und c eine reelle Zahl ist, dann ist auch $f+cg$ selbstadjungiert.*

Beweis: Übung

Hilfssatz 229: *Es seien f und g vertauschbare lineare Funktionen von V nach V. Dann ist jeder Eigenraum von f unter g stabil.*

Beweis: Sei w ein Eigenvektor von f zum Eigenwert c. Dann ist

$$f(g(w)) = g(f(w)) = g(c \cdot w) = c \cdot g(w) \ ,$$

also ist $g(w)$ ein Element des Eigenraums von f zum Eigenwert c.

Satz 230: *Es sei $f : V \longrightarrow V$ eine normale Funktion, c ein beliebiger Eigenwert von f und $E(f,c)$ der zugehörige Eigenraum.*

(1) *Das orthogonale Komplement $E(f,c)^\perp$ von $E(f,c)$ ist unter f stabil.*
(2) $\mathrm{Kern}(f) = \mathrm{Kern}(f^*)$.
(3) $E(f,c) = E(f^*,\overline{c})$, *insbesondere sind alle Eigenwerte von selbstadjungierten Funktionen reell.*
(4) *Eigenvektoren von f zu verschiedenen Eigenwerten stehen zueinander senkrecht.*

Beweis:

(1) Sei $w \in E(f,c)^\perp$. Wir zeigen, dass auch $f(w) \in E(f,c)^\perp$ ist.
Für alle $u \in E(f,c)$ ist nach Hilfssatz 229 auch $f^*(u) \in E(f,c)$, also $\langle u, f(w)\rangle = \langle f^*(u), w\rangle = 0$. Daher ist $f(w) \in E(f,c)^\perp$.
(2) Für $w \in V$ ist
$$\langle f(w), f(w)\rangle = \langle f^*(f(w)), w\rangle = \langle f(f^*(w)), w\rangle = \langle f^*(w), f^*(w)\rangle \ ,$$
daher ist $f(w) = 0$ genau dann, wenn $f^*(w) = 0$.
(3) Aus Satz 225 folgt, dass $(f - c \cdot \mathrm{Id}_V)^* = f^* - \overline{c} \cdot \mathrm{Id}_V$. Mit (2) erhalten wir daher
$$E(f,c) = \mathrm{Kern}(f - c \cdot \mathrm{Id}_V) = \mathrm{Kern}(f^* - \overline{c} \cdot \mathrm{Id}_V) = E(f^*, \overline{c}) \ .$$
(4) Seien u, w Eigenvektoren von f zu zwei verschiedenen Eigenwerten c, d. Nach (3) ist
$$(c-d)\langle u,w\rangle = \langle \overline{c} \cdot u, w\rangle - \langle u, d \cdot w\rangle = \langle f^*(u), w\rangle - \langle u, f(w)\rangle = 0 \ .$$
Daher ist $\langle u, w\rangle = 0$.

Satz 231 : *(„Spektralsatz für normale Funktionen“)*
Es sei f eine lineare Funktion von V nach V. Im Fall $K = \mathbb{R}$ nehmen wir zusätzlich an, dass alle Nullstellen des charakteristischen Polynoms von f reell sind.
Dann sind die folgenden Aussagen äquivalent:

(1) *Die Funktion f ist normal.*
(2) *V hat eine ON-Basis aus Eigenvektoren von f.*

Für Matrizen formuliert:
Es sei $A \in K^{n\times n}$. Im Fall $K = \mathbb{R}$ nehmen wir zusätzlich an, dass alle Nullstellen des charakteristischen Polynoms von A reell sind.
Dann ist $\overline{A}^\top \cdot A = A \cdot \overline{A}^\top$ genau dann, wenn es eine Matrix $S \in \mathcal{O}_n$ (falls $K = \mathbb{R}$) bzw. $S \in \mathcal{U}_n$ (falls $K = \mathbb{C}$) gibt so, dass $S^{-1} \cdot A \cdot S$ eine Diagonalmatrix ist.

Beweis:

(1) $\Rightarrow$ (2) : Induktion über $n = \dim_K(V)$.
Für $n = 1$ ist nichts zu zeigen.
Sei $n > 1$. Nach Annahme gibt es eine Zahl $c \in K$ so, dass $U := E(f,c) \neq \{0\}$ ist. Nach Satz 230, (1) ist $U^\perp$ stabil unter f. Daher folgt aus der Induktionsannahme, dass $U^\perp$ eine ON-Basis aus Eigenvektoren von f hat. Ergänze diese durch eine ON-Basis von U zu einer ON-Basis von V.
(2) $\Rightarrow$ (1) : Sei $\underline{v}$ eine ON-Basis von V aus Eigenvektoren von f und A die Matrix von f bezüglich $\underline{v}$. Dann ist A eine Diagonalmatrix und somit normal.

Satz 232 : *(„Spektralsatz für selbstadjungierte Funktionen“)*
Es sei f eine lineare Funktion von V nach V.
Die folgenden Aussagen sind äquivalent:

(1) *Die Funktion f ist selbstadjungiert.*
(2) *V hat eine ON-Basis aus Eigenvektoren von f und alle komplexen Nullstellen des charakteristischen Polynoms von f sind reell.*

Für Matrizen formuliert:
Es sei $A \in K^{n\times n}$.
Dann ist $\overline{A}^\top = A$ genau dann, wenn es eine Matrix $S \in \mathcal{O}_n$ (falls $K = \mathbb{R}$) bzw. $S \in \mathcal{U}_n$ (falls $K = \mathbb{C}$) gibt so, dass $S^{-1} \cdot A \cdot S$ eine reelle Diagonalmatrix ist.

Beweis:

(1) $\Rightarrow$ (2) : Nach Satz 230 sind alle Eigenwerte von f reell. Daher folgt die Behauptung aus Satz 231.

(2) $\Rightarrow$ (1) : Die Matrix von f bezüglich der Basis aus Eigenvektoren ist eine reelle Diagonalmatrix. Daher ist f selbstadjungiert.

Satz 233 : *Zu jeder symmetrischen, positiv semidefiniten Matrix $A \in \mathbb{R}^{n\times n}$ gibt es genau eine symmetrische, positiv semidefinite Matrix $P \in \mathbb{R}^{n\times n}$ mit $P^2 = A$.*
Ist A positiv definit, dann auch P.

Beweis: Nach Satz 232 gibt es eine orthogonale Matrix Q so, dass

$$Q^\top \cdot A \cdot Q = Diag(c_1, \dots, c_r, 0, \dots, 0) =: D$$

ist, wobei r der Rang von A ist und $c_1, \dots, c_r$ die positiven Eigenwerte von A sind.
Sei $\sqrt{D} := Diag(\sqrt{c_1}, \dots, \sqrt{c_r}, 0, \dots, 0)$ und $P := Q \cdot \sqrt{D} \cdot Q^\top$. Dann ist

$$P^2 = Q \cdot \sqrt{D} \cdot \sqrt{D} \cdot Q^\top = Q \cdot D \cdot Q^\top = A \ .$$

Wir zeigen noch, dass P eindeutig bestimmt ist. Es seien $R \in \mathbb{R}^{n\times n}$ eine symmetrische, positiv semidefinite Matrix mit $R^2 = A$ und z ein Eigenvektor von A zum Eigenwert λ. Es sei $y := R \cdot z - \sqrt{\lambda} \cdot z$. Wegen

$$R \cdot y = A \cdot z - \sqrt{\lambda} \cdot R \cdot z = \lambda \cdot z - \sqrt{\lambda} \cdot R \cdot z = -\sqrt{\lambda} \cdot y$$

und weil R keine negativen Eigenwerte hat, ist $\lambda = 0$ oder $y = 0$.
Wenn $\lambda = 0$ ist, ist $R^2 \cdot z = A \cdot z = 0$ und, weil R invertierbar ist, auch $R \cdot z = 0$.
Wenn $y = 0$ ist, ist $R \cdot z = \sqrt{\lambda} z$.
Also ist in beiden Fällen z auch ein Eigenvektor von R zum Eigenwert $\sqrt{\lambda}$.
Daher ist jede Eigenbasis von A auch eine Eigenbasis von R und die Eigenwerte von R sind durch die von A eindeutig bestimmt. Daher ist auch R eindeutig.

Satz 234 : *Für eine symmetrische Matrix $A \in \mathbb{R}^{n\times n}$ sind die folgenden Aussagen äquivalent:*

(1) *A ist positiv definit bzw. positiv semidefinit.*
(2) *Alle Eigenwerte von A sind positiv bzw. nicht negativ.*

Beweis: Wir zeigen die Behauptung für „positiv definit“(für „positiv semidefinit“ analog).

(1) $\Rightarrow$ (2) : Es sei $x \in \mathbb{R}^{n\times 1}$ ein Eigenvektor von A zum Eigenwert λ. Dann ist $Ax = \lambda x$, also $0 < x^\top \cdot A \cdot x = \lambda x^\top \cdot x$.
Wegen $x^\top \cdot x > 0$ ist auch $\lambda > 0$.

(2) $\Rightarrow$ (1) : Weil A symmetrisch ist, gibt es nach Satz 232 eine orthogonale Matrix S so, dass $S^\top \cdot A \cdot S = Diag(c_1, \dots, c_n)$ ist, wobei $c_1, \dots, c_n$ die Eigenwerte von A sind. Nach (2) sind alle c_i positiv, $1 \le i \le n$,

daher gibt es positive reelle Zahlen d_i mit $d_i^2 = c_i$, $1 \leq i \leq n$.
Sei $B := Diag(d_1, \ldots, d_n) \cdot S$, dann ist $A = B^\top \cdot B$, somit nach Satz 83 positiv definit.

KAPITEL 6

Verallgemeinerte Eigenräume

Wenn V ein endlich-dimensionaler Vektorraum über einem Körper K ist und $f : V \to V$ eine lineare Funktion, dann stellt sich die Frage, wie „einfach“ die Matrix von f durch Wahl einer geeigneten Basis gemacht werden kann. In diesem Kapitel geben wir für $K = \mathbb{C}$ eine vollständige Antwort auf diese Frage. Da für nicht-diagonalisierbare Funktionen die Summe der Eigenräume nicht den ganzen Definitionsbereich ergibt, liegt es nahe, den Begriff des Eigenraums zu verallgemeinern.

§1. Diagonalisierbare Funktionen

Definition 235 : Sei V ein endlich-dimensionaler Vektorraum über K und $f : V \to V$ eine lineare Funktion.

(1) Die Funktion f heißt *diagonalisierbar*, wenn es eine Basis $\underline{u}$ von V gibt, sodass die Matrix von f bezüglich $\underline{u}$ Diagonalform hat.
(2) Eine Basis $(w_1, \ldots, w_n)$ von V heißt eine *Eigenbasis* von f, wenn $w_1, \ldots, w_n$ Eigenvektoren von f sind.

Satz 236 : *Sei V ein endlich-dimensionaler Vektorraum über K und $f : V \to V$ eine K-lineare Funktion. Dann sind die folgenden Aussagen äquivalent:*

(1) *f ist diagonalisierbar.*
(2) *Es gibt eine Eigenbasis von f.*
(3) *V ist die Summe der Eigenräume von f.*
(4) *Die Summe der Dimensionen der Eigenräume von f ist gleich der Dimension von V.*

In diesem Fall bekommt man eine Eigenbasis von f, indem man für jeden Eigenwert eine Basis des zugehörigen Eigenraums berechnet und diese Basen dann zusammensetzt.

Beweis: Seien $c_1, \ldots, c_\ell$ die verschiedenen Eigenwerte von f.
$(1) \Rightarrow (2)$: Sei $\underline{u}$ eine Basis von V so, dass die Matrix B von f bezüglich $\underline{u}$ Diagonalform hat. Dann ist für $j = 1, \ldots, n$

$$f(u_j) = \sum_{i=1}^{n} B_{ij} u_i = B_{jj} u_j \,,$$

also u_j ein Eigenvektor von f zum Eigenwert B_{jj}, und $\underline{u}$ eine Eigenbasis von f.
(2) $\Rightarrow$ (1): Sei $\underline{w} = (w_1, \ldots, w_n)$ eine Eigenbasis von f. Dann sind $w_1, \ldots, w_n$ Eigenvektoren von f. Seien $a_1, \ldots, a_n$ die zugehörigen Eigenwerte von f. Dann ist für $j = 1, \ldots, n$

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

also hat die Matrix von f bezüglich der Basis $\underline{w}$ Diagonalform.
(2) $\Rightarrow$ (3): Jeder Vektor der Eigenbasis ist in einem Eigenraum und damit in der Summe aller Eigenräume enthalten. Da die Eigenbasis ganz V erzeugt, ist auch V in der Summe aller Eigenräume enthalten. Somit ist V die Summe der Eigenräume von f.
(3) $\Rightarrow$ (2): Nach Satz 51 ist V die direkte Summe aller Eigenräume. Wählt man in jedem Eigenraum eine Basis, dann erhält man nach Satz 50 eine Eigenbasis.
(3) $\Rightarrow$ (4): Nach Satz 51 ist $V = E(f,c_1) \oplus \cdots \oplus E(f,c_\ell)$. Nach Satz 50 folgt daraus

$$dim_K(V) = \dim_K(E(f,c_1)) + \cdots + \dim_K(E(f,c_\ell)) \,.$$

(4) $\Rightarrow$ (3): Nach Satz 51 ist die Summe aller Eigenräume direkt und daher

$$\begin{aligned} \dim_K(V) &= \dim_K(E(f,c_1)) + \cdots + \dim_K(E(f,c_\ell)) \\ &= \dim_K(E(f,c_1) + \cdots + E(f,c_\ell)) \,, \end{aligned}$$

also $V = E(f,c_1) + \cdots + E(f,c_\ell)$.

Satz 237 : *Sei V ein endlich-dimensionaler Vektorraum über K der Dimension n und $f : V \to V$ eine lineare Funktion mit n verschiedenen Eigenwerten. Dann ist f diagonalisierbar und jeder Eigenraum von f ist eindimensional.*

Beweis: Seien $c_1, \ldots, c_n$ die Eigenwerte von f. Da jeder Eigenraum mindestens Dimension 1 hat, ist

$$\begin{aligned} \dim_K(V) = n &\leq \dim_K(E(f,c_1)) + \cdots + \dim_K(E(f,c_n)) \\ &= \dim_K(E(f,c_1) \oplus \cdots \oplus E(f,c_n)) \leq \dim_K(V) \,. \end{aligned}$$

Es folgt

$$\dim_K(V) = \dim_K(E(f,c_1)) + \cdots + \dim_K(E(f,c_n))$$

und $\dim_K(E(f,c_j)) = 1$ für alle $1 \leq j \leq n$. Nach Satz 236,(4) $\Rightarrow$ (1) ist f diagonalisierbar.

§2. Verallgemeinerte Eigenräume

In diesem Abschnitt seien V ein endlich-dimensionaler Vektorraum über einem Körper K, $n := \dim_K(V)$ und $f : V \to V$ eine K-lineare Funktion.

Definition 238 : Sei $c \in K$ ein Eigenwert von f. Dann heißt die Menge

$$V(f,c) := \{x \in V \mid \text{ es gibt ein } k \in \mathbb{N} \text{ mit } (f - c\,\mathrm{Id}_V)^k(x) = 0\}$$

der *verallgemeinerte Eigenraum* von f zum Eigenwert c. Die Vektoren in $V(f,c)$ ungleich 0 nennt man *verallgemeinerte Eigenvektoren* von f zum Eigenwert c.
Es ist

$$E(f,c) \subset V(f,c)\,,$$

d.h. der Eigenraum von f zum Eigenwert c ist im verallgemeinerten Eigenraum enthalten.

Satz 239 : *Sei $c \in K$ ein Eigenwert von f. Dann ist*

$$V(f,c) = \mathrm{Kern}((f - c\,\mathrm{Id}_V)^n)\,,$$

insbesondere ist $V(f,c)$ ein Untervektorraum von V.

Beweis: Offensichtlich ist $\mathrm{Kern}((f - c\,\mathrm{Id}_V)^n) \subset V(f,c)$, sodass nur $V(f,c) \subset \mathrm{Kern}((f - c\,\mathrm{Id}_V)^n)$ zu zeigen ist. Sei dazu $u \in V(f,c)$ mit $u \neq 0$ und bezeichne

$$h := f - c\,\mathrm{Id}_V\ .$$

Sei k die kleinste Zahl, sodass $h^k(u) = 0$ ist. Wegen $u \neq 0$ ist $k > 0$ und $h^{k-1}(u) \neq 0$. Unser Ziel ist, $k \leq n$ zu zeigen, woraus dann

$$h^n(u) = h^{n-k}(h^k(u)) = h^{n-k}(0) = 0$$

und $u \in \mathrm{Kern}(h^n)$ folgt. Um $k \leq n$ zu zeigen, genügt es zu zeigen, dass die Vektoren

$$u, h(u), h^2(u), \ldots, h^{k-1}(u)$$

linear unabhängig sind. Seien dazu $d_0, \ldots, d_{k-1} \in K$ mit

$$d_0 u + d_1 h(u) + \cdots + d_{k-1} h^{k-1}(u) = 0\,. \tag{1}$$

Anwenden von h^{k-1} auf (1) ergibt
$d_0 h^{k-1}(u) = 0$, also $d_0 = 0$.
Anwenden von h^{k-2} auf (1) ergibt dann $d_1 h^{k-1}(u) = 0$, also $d_1 = 0$.
Auf diese Weise folgt $0 = d_0 = d_1 = \cdots = d_{k-1}$.

Definition 240 : Seien $g, h : V \to V$ K-lineare Funktionen. Dann heißen g und h „*vertauschbar*“, wenn

$$g \circ h = h \circ g$$

ist, d.h. die Reihenfolge des Anwendens von g und h spielt keine Rolle.
Ein Untervektorraum U von V heißt *stabil* unter g, wenn $g(U) \subseteq U$ ist.

Hilfssatz 241 : *Sei V ein Vektorraum über einem Körper K, $f : V \to V$ eine lineare Funktion und $a, c \in K$. Dann sind $f - a\,\mathrm{Id}_V$ und $f - c\,\mathrm{Id}_V$ vertauschbar.*

Beweis: Es ist
$(f - c\,\mathrm{Id}_V)(f - a\,\mathrm{Id}_V) = f \circ f - cf - af + ca\,\mathrm{Id}_V = (f - a\,\mathrm{Id}_V)(f - c\,\mathrm{Id}_V)$.

Satz 242 : *Sei $g : V \longrightarrow V$ eine lineare Funktion, die mit f vertauschbar ist. Dann sind die Eigenräume und die verallgemeinerten Eigenräume von f stabil unter g.*

Beweis: Sei c ein Eigenwert bzw. verallgemeinerter Eigenwert von f und u ein entsprechender Eigenvektor bzw. verallgemeinerter Eigenvektor. Weil f und g vertauschbar sind, sind auch $f - cId_V$ und $(f - cId_V)^n$ mit g vertauschbar. Daher ist

$$(f - cId_V)(g(u)) = g((f - cId_V)(u)) = g(0) = 0$$

bzw.

$$(f - cId_V)^n(g(u)) = g((f - cId_V)^n(u)) = g(0) = 0\,,$$

also ist auch $g(u)$ ein Eigenvektor bzw. verallgemeinerter Eigenvektor von f.

Satz 243 : *Seien $c_1, \ldots, c_\ell$ die Eigenwerte von f. Dann ist*

$$V(f, c_1) + \cdots + V(f, c_\ell) = V(f, c_1) \oplus \cdots \oplus V(f, c_\ell)\,,$$

d.h. die Summe der verallgemeinerten Eigenräume von f ist direkt.

Beweis: Seien $u_1 \in V(f, c_1), \ldots, u_\ell \in V(f, c_\ell)$ mit

$$u_1 + \cdots + u_\ell = 0\,.$$

Sei k die kleinste Zahl, sodass $(f - c_1\,\mathrm{Id}_V)^k(u_1) = 0$ ist, und

$$h := (f - c_1\,\mathrm{Id}_V)^{k-1}(f - c_2\,\mathrm{Id}_V)^n \ldots (f - c_\ell\,\mathrm{Id}_V)^n\,.$$

Nach Hilfssatz 241 kommt es bei der Funktion h nicht auf die Reihenfolge der Faktoren an. Anwenden von h auf $u_1 + \ldots + u_\ell$ ergibt

$$0 = h(u_1 + \ldots + u_\ell) = h(u_1) =$$

$$= (f - c_2\,\mathrm{Id}_V)^n \ldots (f - c_\ell\,\mathrm{Id}_V)^n (f - c_1\,\mathrm{Id}_V)^{k-1}(u_1)$$

nach Satz 239. Der Vektor $(f - c_1 \operatorname{Id}_V)^{k-1}(u_1)$ ist ein Eigenvektor von f zum Eigenwert c_1, also ist

$$0 = h(u_1) = (c_1 - c_2)^n \dots (c_1 - c_\ell)^n (f - c_1 \operatorname{Id}_V)^{k-1}(u_1)\,,$$

woraus $u_1 = 0$ folgt. Analog zeigt man $u_j = 0$ für $2 \leq j \leq \ell$.

Hilfssatz 244 : *Seien U, U' Unterverktorräume von V mit*

$$U \cap U' = \{0\}\,.$$

Dann ist die Summe $U + U'$ direkt.

Beweis: Seien $u \in U$ und $u' \in U'$ mit $u + u' = 0$. Dann ist $u = -u' \in U \cap U'$, also $u = 0 = u'$.

Satz 245 : *Sei $K = \mathbb{C}$ und $c_1, \dots, c_\ell$ die Eigenwerte von f. Dann ist*

$$V = V(f, c_1) + \dots + V(f, c_\ell)\,.$$

Beweis: Wir zeigen durch Induktion nach n, dass die Behauptung für komplexe Vektorräume der Dimension $\leq n$ gilt. Für $n = 0$ ist V der Nullraum, f die Nullfunktion, $\ell = 0$ und die leere Summe der Nullraum.

Sei nun $n \geq 1$ und die Behauptung gelte für Vektorräume der Dimension $\leq n - 1$. Da $K = \mathbb{C}$ ist, hat f einen Eigenwert c_1. Sei

$$h := (f - c_1 \operatorname{Id}_V)^n \in \operatorname{Lin}_K(V, V)\,.$$

Nach Satz 239 ist

$$\operatorname{Kern}(h) = V(f, c_1)\,.$$

Sei $y \in \operatorname{Kern}(h) \cap \operatorname{Bild}(h)$. Dann gilt $h(y) = 0$ und es gibt es ein $x \in V$ mit $y = h(x)$. Daher ist $0 = h(y) = h^2(x) = (f - c_1 \operatorname{Id}_V)^{2n}(x)$, also $x \in V(f, c_1)$. Nach Satz 239 folgt $0 = h(x) = y$. Daher ist $\operatorname{Kern}(h) \cap \operatorname{Bild}(h) = \{0\}$ und nach Hilfssatz 244 die Summe von $\operatorname{Kern}(h)$ und $\operatorname{Bild}(h)$ direkt. Nach Satz 50 und Satz 24 folgt

$$\begin{aligned} \dim_{\mathbb{C}}(\operatorname{Kern}(h) \oplus \operatorname{Bild}(h)) &= \dim_{\mathbb{C}}(\operatorname{Kern}(h)) + \dim_{\mathbb{C}}(\operatorname{Bild}(h)) \\ &= \dim_{\mathbb{C}}(V)\,, \end{aligned}$$

also $\operatorname{Kern}(h) \oplus \operatorname{Bild}(h) = V$. Somit ist

$$V = V(f, c_1) \oplus W\,, \tag{2}$$

wobei $W := \operatorname{Bild}(h)$ eine kleinere Dimension als V hat. Nach Hilfssatz 241 gilt $fh = hf$. Für $y \in \operatorname{Bild}(h)$, d.h. $y = h(x)$ für ein $x \in V$, ist somit $f(y) = f(h(x)) = h(f(x))$, also $f(y) \in \operatorname{Bild}(h)$ und $f(W) \subset W$. Daher kann man f einschränken zur Funktion

$$g : W \to W\,, x \mapsto f(x)\,.$$

Nach Induktionsannahme gilt

$$W = V(g,a_1) + \cdots + V(g,a_m)\,, \tag{3}$$

wobei $a_1,\ldots,a_m$ die Eigenwerte von g sind. Jeder Eigenvektor u von g zum Eigenwert a ist wegen $f(u) = g(u) = au$ auch Eigenvektor von f zum Eigenwert a. Daher ist $\mathrm{Sp}_{\mathbb{C}}(g) \subset \mathrm{Sp}_{\mathbb{C}}(f)$ und, für $1 \le j \le m$,

$$V(g,a_j) = \{x \in W \mid \text{ es gibt ein } k \in \mathbb{N} \text{ mit}$$

$$(g - a_j \operatorname{Id}_W)^k(x) = 0\} \subset V(f,a_j)\ .$$

Zusammen mit (2) und (3) erhalten wir
$V \subset V(f,c_1) + \cdots + V(f,c_\ell)$, was zu zeigen war.

Satz 246 : *Sei V ein Vektorraum über $\mathbb{C}$ der Dimension $n \in \mathbb{N}$, $f : V \to V$ eine $\mathbb{C}$-lineare Funktion und χ_f das charakteristische Polynom von f. Es ist*

$$\chi_f = \prod_{j=1}^{\ell} (X - c_j)^{m_j}\,,$$

wobei $c_1,\ldots,c_\ell$ die Eigenwerte von f mit den Vielfachheiten $m_1,\ldots,m_\ell$ sind. Dann ist

$$V = V(f,c_1) \oplus \cdots \oplus V(f,c_\ell)\,,$$

d.h. V ist die direkte Summe der verallgemeinerten Eigenräume von f. Für $1 \le j \le \ell$ gilt

(1) $\dim_{\mathbb{C}}(V(f,c_j)) = m_j$
(2) $V(f,c_j) = \operatorname{Kern}((f - c_j \operatorname{Id}_V)^{m_j})$
(3) *$V(f,c_j)$ ist stabil unter f und die Einschränkung*

$$f_j : V(f,c_j) \to V(f,c_j)\,, x \mapsto f(x)\,,$$

hat nur einen Eigenwert, und zwar c_j.

Beweis: Für $1 \le j \le \ell$ sei $V_j := V(f,c_j)$ und $n_j := \dim_{\mathbb{C}}(V_j)$. Nach Satz 245 und Satz 243 ist V die direkte Summe der verallgemeinerten Eigenräume V_j . Nach Satz 242 ist V_j stabil unter f. Da V_j den Eigenraum $E(f,c_j)$ enthält, ist c_j ein Eigenwert von f_j . Wenn a ein beliebiger Eigenwert von f_j ist, dann gibt es einen Vektor $u \in V_j$, $u \neq 0$, mit $f(u) = au$, woraus

$$(f - c_j \operatorname{Id}_V)^k(u) = (a - c_j)^k u$$

für alle $k \in \mathbb{N}$ folgt, also $a = c_j$. Daher ist c_j der einzige Eigenwert von f_j und (3) bewiesen. Nach Satz 49 folgt

$$\chi_{f_j} = (X - c_j)^{n_j}\,.$$

Wählt man in jedem verallgemeinerten Eigenraum V_j eine Basis $(u_{j1},\ldots,u_{jn_j})$, dann kann man nach Satz 50 diese Basen zu einer Basis

$(u_{11},\dots,u_{\ell n_\ell})$ von V zusammensetzen. Da jeder verallgemeinerte Eigenraum stabil unter f ist, hat die Matrix von f bezüglich dieser Basis

$$B=\begin{pmatrix} B_1 & 0 & \dots & 0 \\ 0 & B_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & B_\ell \end{pmatrix},$$

Blockdiagonalform, wobei B_j die Matrix von f_j bezüglich der Basis $(u_{j1},\dots,u_{jn_j})$ ist. Damit hat auch $cI_n - B$ Blockdiagonalform, und nach Satz 67 folgt

$$\begin{aligned}\chi_f = \chi_B &= \det(cI_n - B) = \det(cI_{n_1} - B_1)\dots\det(cI_{n_\ell} - B_\ell) \\ &= \chi_{B_1}\dots\chi_{B_\ell} = \chi_{f_1}\dots\chi_{f_\ell} = (X-c_1)^{n_1}\dots(X-c_\ell)^{n_\ell}\,.\end{aligned}$$

Somit ist $n_j = m_j$ für $1 \le j \le \ell$ und (1) bewiesen. Anwenden von Satz 245 und Satz 239 auf f_j gibt

$$V_j = V(f_j,c_j) = \mathrm{Kern}((f_j - c_j\,\mathrm{Id}_{V_j})^{m_j}) \subset \mathrm{Kern}((f-c_j\,\mathrm{Id}_V)^{m_j}) \subset V_j\,.$$

Daher ist $V_j = \mathrm{Kern}((f-c_j\,\mathrm{Id}_V)^{m_j})$ und (2) bewiesen.

§3. Die Jordansche Normalform komplexer Matrizen

Satz 247 : *Sei V ein Vektorraum über $\mathbb{C}$ der Dimension $n \in \mathbb{N}$, $f: V \to V$ eine $\mathbb{C}$-lineare Funktion, $c_1,\dots,c_\ell$ die Eigenwerte von f mit den Vielfachheiten $m_1,\dots,m_\ell$ und*

$$V = V(f,c_1)\oplus\dots\oplus V(f,c_\ell)\,,$$

die Zerlegung von V in die direkte Summe der verallgemeinerten Eigenräume von f.
$D_f : V \to V$ sei jene lineare Funktion, deren Einschränkung auf $V(f,c_j)$ gleich $c_j Id_{V(f,c_j)}$ ist, $1 \le j \le \ell$. Weiters sei $N_f := f - D_f$.
Dann ist D_f diagonalisierbar, $N_f^n = 0$ und

$$f = D_f + N_f\,,\ D_f\circ N_f = N_f \circ D_f\,,\ D_f\circ f = f\circ D_f\,,\ \ N_f\circ f = f\circ N_f\,.$$

(„Jordan-Zerlegung von f“, D_f ist der diagonalisierbare Summand *von f, N_f ist der* nilpotente Summand *von f).*

Beweis: Nach Satz 246 und Satz 242 genügt es, den Satz für $\ell = 1$ zu beweisen. In diesem Fall sind die Aussagen aber leicht nachzuprüfen.

Hilfssatz 248 : *Sei $P : K \longrightarrow K$ eine Polynomfunktion mit Koeffizienten $a_0,a_1,\dots,a_n$. Dann sei*

$$P(f) := \sum_{i=0}^{n} a_i f^i \in \mathrm{Lin}_K(V,V)$$

(„f eingesetzt in P“). Dabei sind die Potenzen f^i bezüglich der Zusammensetzung von Funktionen zu verstehen, insbesondere ist $f^0 = \mathrm{Id}_V$. Wenn Q eine weitere Polynomfunktion ist, dann gilt

$$(P+Q)(f) = P(f) + Q(f) \quad \textit{und} \quad (P \cdot Q)(f) = P(f) \circ Q(f) \,.$$

Beweis: Wir beweisen nur die zweite Aussage, die erste beweist man analog. Sei Q durch die endliche Folge $b_0, b_1, \ldots, b_m$ gegeben. Dann ist $P \cdot Q$ durch die Koeffizienten

$$\left(\sum_{i=0}^{k} a_i b_{k-i} \right)_{0 \le k \le m+n}$$

gegeben. Andererseits ist

$$\begin{aligned} P(f) \circ Q(f) &= \left(\sum_{i=0}^{n} a_i f^i \right) \left(\sum_{j=0}^{m} b_j f^j \right) = \sum_{i=0}^{n} \sum_{j=0}^{m} a_i b_j f^{i+j} \\ &= \sum_{k=0}^{n+m} \left(\sum_{i=0}^{k} a_i b_{k-i} \right) f^k \,. \end{aligned}$$

Daher ist $(P \cdot Q)(f) = P(f) \circ Q(f)$.

Satz 249: *(Satz von Cayley-Hamilton)*
Sei V ein endlich-dimensionaler Vektorraum über $\mathbb{C}$ und $f : V \to V$ eine $\mathbb{C}$-lineare Funktion mit charakteristischem Polynom χ_f. Dann ist

$$\chi_f(f) = 0$$

in $\mathrm{Lin}_K(V,V)$.

Beweis: Seien $c_1, \ldots, c_\ell$ die Eigenwerte von f mit Vielfachheiten $m_1, \ldots, m_\ell$. Dann ist

$$\chi_f = \prod_{j=1}^{\ell} (X - c_j)^{m_j} \,,$$

nach Hilfssatz 248 somit

$$\chi_f(f) = \prod_{j=1}^{\ell} (f - c_j \,\mathrm{Id}_V)^{m_j} \,.$$

Für $x \in V$ gibt es nach Satz 246 Vektoren
$x_1 \in V(f, c_1), \ldots, x_\ell \in V(f, c_\ell)$ mit $x = \sum_{i=1}^{\ell} x_i$. Daraus folgt

$$\left(\prod_{j=1}^{\ell} (f - c_j \,\mathrm{Id}_V)^{m_j} \right)(x) = \sum_{i=1}^{\ell} \left(\prod_{j=1}^{\ell} (f - c_j \,\mathrm{Id}_V)^{m_j} \right)(x_i) = 0 \,,$$

weil nach Hilfssatz 241 die Faktoren des Produkts vertauschbar sind und nach Satz 246 ist $(f - c_j \,\mathrm{Id}_V)^{m_j}(x_j) = 0$.

Definition 250 : Sei $m \in \mathbb{N}$ mit $m \geq 1$ und $a \in K$. Dann heißt die Matrix

$$J(a) := \begin{pmatrix} a & 1 & 0 & \dots & 0 \\ 0 & a & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & a & 1 \\ 0 & \dots & \dots & 0 & a \end{pmatrix} \in K^{m \times m}$$

ein *Jordankästchen* der Größe m mit Diagonalelement a.

Definition 251 : Eine Matrix hat *Jordansche Normalform*, wenn sie Blockdiagonalform

$$D = \begin{pmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & J_p \end{pmatrix}$$

mit Jordankästchen $J_1, \dots, J_p$ in der Hauptdiagonalen hat.

Satz 252 : *Sei V ein endlich-dimensionaler Vektorraum über $\mathbb{C}$ und $f : V \to V$ eine $\mathbb{C}$-lineare Funktion. Dann gibt es eine Basis $\underline{u}$ von V, sodass die Matrix von f bezüglich $\underline{u}$ Jordansche Normalform mit Jordankästchen $J_1, \dots, J_p$ entlang der Hauptdiagonalen hat. Die Jordankästchen $J_1, \dots, J_p$ sind eindeutig bis auf die Reihenfolge.*

Die Funktion f ist genau dann diagonalisierbar, wenn jedes Jordankästchen die Größe 1 hat. Jede Matrix in $\mathbb{C}^{n \times n}$ ist ähnlich zu einer Matrix in Jordanscher Normalform.
Die Basis $\underline{u}$ heißt Jordanbasis *von f.*

Beispiel 253 : Für komplexe 3×3-Matrizen sind die folgenden Jordanschen Normalformen möglich:

$$\begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \quad , \quad \begin{pmatrix} a & 1 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{pmatrix} .$$

Dabei sind a, b, c (nicht notwendig verschiedene) komplexe Zahlen.

Beweis: Wir zeigen in Teil (1), dass die Jordankästchen eindeutig bis auf die Reihenfolge sind. In Teil (2) konstruieren wir dann eine Jordanbasis von f.

(1) Sei $n := \dim_{\mathbb{C}}(V)$ und $\underline{u}$ eine Basis von V so, dass die Matrix D von f bezüglich $\underline{u}$ Blockdiagonalform mit Blockgrößen $(n_1, \dots, n_p)$ hat, wobei in der Diagonale die Jordankästchen $J_1, \dots, J_p$ mit Diagonalelementen

$a_1,\dots,a_p$ stehen. Nach Satz 67 ist

$$\chi_f = \chi_D = \prod_{k=1}^{p} \chi_{J_k} = \prod_{k=1}^{p} (X-a_k)^{n_k}\,, \tag{4}$$

also $\{a_1,\dots,a_p\} = \mathrm{Sp}_{\mathbb{C}}(f)$. Für $c \in \mathrm{Sp}_{\mathbb{C}}(f)$ und $m \in \{1,\dots,n\}$ bezeichne

$$b_m(c)$$

die Anzahl der Jordankästchen mit Größe m und Diagonalelement c in der Matrix D. Wir zeigen nun, dass für jedes $c \in \mathrm{Sp}_{\mathbb{C}}(f)$ die Zahlen $b_1(c),\dots,b_n(c)$ durch f eindeutig bestimmt sind, woraus die Eindeutigkeit der Jordankästchen bis auf die Reihenfolge folgt. Dazu schreiben wir entsprechend der Blockdiagonalform von D

$$\underline{u} = (u_{kj})_{\substack{1\le k\le p\\ 1\le j\le n_k}}\,,$$

sodass für $1 \le k \le p$

$$(f(v_{k1}),\dots,f(v_{kn_k})) = (v_{k1},\dots,v_{kn_k}) \begin{pmatrix} a_k & 1 & 0 & \dots & 0 \\ 0 & a_k & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & a_k & 1 \\ 0 & \dots & \dots & 0 & a_k \end{pmatrix} \tag{5}$$

gilt. Im Folgenden wählen wir $c \in \mathrm{Sp}_{\mathbb{C}}(f)$ fest und setzen

$$g := f - c\,\mathrm{Id}_V$$

sowie $W_\ell := \mathrm{Kern}(g^\ell)$ für $0 \le \ell \le n$. Dann ist nach Satz 239

$$\{0\} = W_0 \subset W_1 \subset \dots \subset W_n = V(f,c)\,.$$

Für alle $k \in \{1,\dots,p\}$ mit $a_k = c$ erhält man aus (5)

$$\begin{aligned} g(u_{kn_k}) &= u_{k,n_k-1} \\ g(u_{k,n_k-1}) &= u_{k,n_k-2} \\ &\dots \\ g(u_{k2}) &= u_{k1} \\ g(u_{k1}) &= 0\,. \end{aligned} \tag{6}$$

Für $1 \le \ell \le n$ setzen wir

$$I_\ell := \{k \in \{1,\dots,p\} \mid a_k = c \text{ und } n_k \ge \ell\}\,,$$

sodass die J_k mit $k \in I_\ell$ alle Jordankästchen von D mit Diagonalelement c und Größe $\ge \ell$ sind und daher

$$\#(I_\ell) = \sum_{m=\ell}^{n} b_m(c) \tag{7}$$

ist. Wegen (6) ist

$$U_\ell := {}_K\langle v_{k\ell} \mid k \in I_\ell \rangle$$

ein Untervektorraum von W_ℓ. Weiters gilt $U_\ell \cap W_{\ell-1} = \{0\}$, weil für eine Koeffizientenfamilie $(d_k)_{k\in I_\ell}$ mit $\sum_{k\in I_\ell} d_k v_{k\ell} \in W_{\ell-1}$ wegen (6)

$$0 = \sum_{k\in I_\ell} d_k g^{\ell-1}(v_{k\ell}) = \sum_{k\in I_\ell} d_k v_{k1}$$

und $d_k = 0$ für alle $k \in I_\ell$ folgt. Nach Hilfssatz 244 ist die Summe von U_ℓ und $W_{\ell-1}$ direkt. Aus $U_\ell \oplus W_{\ell-1} \subset W_\ell$ folgt

$$\dim_{\mathbb{C}}(U_\ell) + \dim_{\mathbb{C}}(W_{\ell-1}) \leq \dim_{\mathbb{C}}(W_\ell)$$

und wegen (7) ist

$$\alpha_\ell := \sum_{m=\ell}^{n} b_m(c) \leq \dim_{\mathbb{C}}(W_\ell) - \dim_{\mathbb{C}}(W_{\ell-1}) =: \beta_\ell \tag{8}$$

für $1 \leq \ell \leq n$. Es ist

$$\sum_{\ell=1}^{n} \beta_\ell = \sum_{\ell=1}^{n} (\dim_{\mathbb{C}}(W_\ell) + \dim_{\mathbb{C}}(W_{\ell-1})) = \dim_{\mathbb{C}}(V(f,c)).$$

Weiters ist

$$\sum_{\ell=1}^{n} \alpha_\ell = \sum_{\ell=1}^{n}\sum_{m=\ell}^{n} b_m(c) = \sum_{m=1}^{n}\sum_{\ell=1}^{m} b_m(c) = \sum_{m=1}^{n} m b_m(c).$$

Nach Definition von $b_m(c)$ und Satz 246 ist $\sum_{m=1}^{n} m b_m(c)$ gleich der Dimension von $V(f,c)$. Also ist

$$\sum_{\ell=1}^{n} \alpha_\ell = \sum_{\ell=1}^{n} \beta_\ell$$

und wegen $0 \leq \alpha_\ell \leq \beta_\ell$ daher auch $\alpha_\ell = \beta_\ell$ für $1 \leq \ell \leq n$. Wegen

$$\begin{aligned} b_\ell(c) &= \alpha_\ell - \alpha_{\ell+1} = \beta_\ell - \beta_{\ell+1} = \\ &= 2\dim_{\mathbb{C}}(W_\ell) - \dim_{\mathbb{C}}(W_{\ell-1}) - \dim_{\mathbb{C}}(W_{\ell+1}) \end{aligned}$$

ist $b_\ell(c)$ durch f eindeutig bestimmt.

(2) Um eine Jordanbasis von f zu konstruieren, genügt es nach Satz 246, für $1 \leq j \leq \ell$ eine Jordanbasis für die Einschränkung von f auf den verallgemeinerten Eigenraum $V(f,c_j)$ zu konstruieren und diese Basen dann zusammenzusetzen. Sei im Folgenden $c \in \mathrm{Sp}_{\mathbb{C}}(f)$ fest gewählt, $W := V(f,c)$ und

$$g : W \to W\,, x \mapsto f(x) - cx\,.$$

Im Rest des Beweises konstruieren wir eine Jordanbasis von g. Diese Basis ist dann auch eine Jordanbasis der Einschränkung von f auf W. Sei m die Vielfachheit des Eigenwerts c, und für $0 \leq k \leq m$ sei $W_k := \mathrm{Kern}(g^k)$. Dann ist nach Satz 246

$$\{0\} = W_0 \subset W_1 \subset \cdots \subset W_m = W\,.$$

Für den weiteren Beweis brauchen wir folgendes Hilfssatz.

Hilfssatz 254 : *Für alle $k \in \{1,\dots,m\}$ ist*

$$g(W_k) \subset W_{k-1}\,.$$

Sei Z ein Untervektorraum von W_ℓ für ein $\ell \in \{2,\dots,m\}$ mit

$$Z \cap W_{\ell-1} = \{0\}\,.$$

Dann ist auch

$$g(Z) \cap W_{\ell-2} = \{0\}\,,$$

und die Funktion $Z \to g(Z)\,, x \mapsto g(x)\,,$ ist ein Isomorphismus.

Beweis: Für $x \in W_\ell$ ist $g^{\ell-1}(g(x)) = g^\ell(x) = 0$, also $g(x) \in W_{\ell-1}$ und damit $g(W_\ell) \subset W_{\ell-1}$.

Für $y \in g(Z) \cap W_{\ell-2}$ gibt es ein $x \in Z$, sodass $y = g(x)$ ist. Wegen $0 = g^{\ell-2}(y) = g^{\ell-1}(x)$ ist $x \in Z \cap W_{\ell-1}$, also $x = 0$ und $y = 0$. Daher ist $g(Z) \cap W_{\ell-2} = \{0\}$.

Die Funktion $h : Z \to g(Z)\,, x \mapsto g(z)\,,$ ist linear und surjektiv. Wegen $\mathrm{Kern}(h) = \mathrm{Kern}(g) \cap Z = W_1 \cap Z \subset W_{\ell-1} \cap Z = \{0\}$ ist h injektiv und ein Isomorphismus.

Wir setzen den Beweis von Satz 252 fort und wählen sukzessive Untervektorräume $U_m \subset W_m$, $U_{m-1} \subset W_{m-1}$,$\dots$,$U_1 \subset W_1$, sodass

$$\begin{aligned}
W_m &= W_{m-1} \oplus U_m \\
W_{m-1} &= W_{m-2} \oplus g(U_m) \oplus U_{m-1} \\
W_{m-2} &= W_{m-3} \oplus g^2(U_m) \oplus g(U_{m-1}) \oplus U_{m-2} \\
&\dots \\
W_1 &= W_0 \oplus g^{m-1}(U_m) \oplus g^{m-2}(U_{m-1}) \oplus \dots \oplus g(U_2) \oplus U_1
\end{aligned} \tag{9}$$

ist. Dies ist möglich, weil für $2 \le \ell \le m$ nach Wahl von $U_m,\dots,U_\ell$

$$W_\ell = W_{\ell-1} \oplus Z$$

mit

$$Z := g^{m-\ell}(U_m) \oplus g^{m-\ell-1}(U_{m-1}) \oplus \dots \oplus g(U_{\ell+1}) \oplus U_\ell$$

gilt und aus Hilfssatz 254

$$W_{\ell-1} \supset W_{\ell-2} \oplus g(Z)$$

sowie

$$g(Z) = g^{m-\ell+1}(U_m) \oplus g^{m-\ell}(U_{m-1}) \oplus \dots \oplus g(U_{\ell+2}) \oplus U_{\ell+1}$$

folgen. Weiters wählen wir für $1 \le \ell \le m$ eine Basis $(u_{\ell 1},\dots,u_{\ell n_\ell})$ von U_ℓ. Dann ist nach Hilfssatz 254 für $0 \le k < \ell$

$$(g^k(u_{\ell 1}),\dots,g^k(u_{\ell n_\ell}))$$

eine Basis von $g^k(U_\ell)$. Aus (9) folgt

$$W = U_1 \oplus (g(U_2) \oplus U_2) \oplus \cdots \oplus (g^{m-1}(U_m) \oplus \cdots \oplus g(U_m) \oplus U_m)\,.$$

Somit kann man eine Basis $\underline{w}$ von W aus den Familien

$$\underline{v_{\ell i}} := (g^{\ell-1}(u_{\ell i}), g^{\ell-2}(u_{\ell i}), \ldots, g(u_{\ell i}), u_{\ell i})$$

mit $1 \leq \ell \leq m$ und $1 \leq i \leq n_\ell$ zusammensetzen. Wegen

$$g(\underline{v_{\ell i}}) = \underline{v_{\ell i}} \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 0 & 1 \\ 0 & \ldots & \ldots & 0 & 0 \end{pmatrix}$$

ist $\underline{w}$ eine Jordanbasis von g und der Beweis zu Ende.