Geometrie

Goethe–Universität Frankfurt — Sommersemester 2016

für Bachelor und L3

JAKOB STIX

Zusammenfassung. — Die Vorlesung behandelt die Theorie der Bilinearformen auf Vektorräumen. Thema sind insbesondere euklidische Vektorräume, Isometrien und Bewegungen, affine und projektive Geometrie. Als Anwendung klassifizieren wir Quadriken.

Das Skript wird fortlaufend aktualisiert und es werden weiterhin Fehler korrigiert. Sie lesen daher das Skript **auf eigene Gefahr!** Bitte teilen Sie mir Korrekturvorschläge per Email mit.

Inhaltsverzeichnis

Danksagung. Ich möchte mich gerne bei allen bedanken, insbesondere bei den Studierenden Adrian Baumann, Theresa Kumpitsch, Denise Melchin, und Julia Weber, die dazu beigetragen haben, das Skript von kleineren und größeren Eseleien zu befreien, auch wenn dies ein Kampf gegen die Windmühlen und die Rechtschreibreform ist. So mag ich beispielsweise beim besten Willen manches Mal nicht auf das “ß” verzichten.

Einführung

In der Linearen Algebra 1 entwickelt man die algebraische Theorie, um **lineare** (homogene) Polynome, wie zum Beispiel $3x + 5y$, und die daraus resultierenden Gleichungssysteme zu studieren. Dies ist die Theorie der Vektorräume und der linearen Abbildungen, in expliziter Form durch Matrizen gegeben.

In der Geometrie werden Vektorräume mit Begriffen für Abstand und Winkel versehen. Es stellt sich heraus, daß dazu bilineare Abbildungen nötig sind: in Koordinaten durch (homogene) **quadratische** Polynome, wie zum Beispiel $3x^2-7xy+19y^2$. Wir illustrieren den Zusammenhang mit Matrizen durch die folgende Beispielrechnung.

Eine quadratische Form, also homogen vom Grad 2, kann durch eine symmetrische Matrix beschrieben werden:

$$q(x,y) = 19x^2 - 4xy + 16y^2 = (x,y)\begin{pmatrix} 19 & -2 \\ -2 & 16 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

Die Matrix ist symmetrisch gewählt, aber die Nichtdiagonaleinträge tragen beide zum Monom

$$xy$$

bei. Diese symmetrische Aufteilung ist willkürlich, aber symmetrische Matrizen haben besondere Eigenschaften, die es hier auszunutzen gilt. Die quadratische Ergänzung

$$19x^2 - 4xy + 16y^2 = 3(x+2y)^2 + 4(y-2x)^2$$

zeigt, daß nach Koordinatenwechsel $u = x + 2y$ und $v = y - 2x$, also $\binom{u}{v} = S\binom{x}{y}$ mit

$$S = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix},$$

die quadratisch Form einfacher wird:

$$q(u,v) = 3u^2 + 4v^2.$$

In den neuen Koordinaten ist für $r > 0$ die Menge

$$E_r := \left\{ \begin{pmatrix} u \\ v \end{pmatrix} \in \mathbb{R}^2 \ ; \ q(u,v) = r \right\}$$

eine Ellipse. Die Achsen des neuen Koordinatensystems liegen in Richtung der Spalten von

$$S^{-1} = \frac{1}{5}\begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}$$

und diese enthalten gerade die Eigenvektoren (nachrechnen!) der symmetrischen Matrix

$$\begin{pmatrix} 19 & -2 \\ -2 & 16 \end{pmatrix},$$

welche für die quadratische Form verantwortlich ist. Die entsprechenden Eigenwerte 3 und 4 treten als Koeffizienten in $q(u,v)$ auf. Der Satz über die Hauptachsentransformation besagt insbesondere, daß die neuen Achsen wieder senkrecht aufeinander stehen und die Koordinatentransformation so gewählt werden kann, daß sie Winkel und Abstände erhält. Die Mengen E_r sind also auch in alten Koordinaten Ellipsen.

Die folgenden Lehrbücher werden für die Vorlesung empfohlen.

Literatur

[Ar93] Michael Artin, *Algebra*, Übersetzung des englischen Originals von 1991 durch Annette A'Campo, Birkhäuser Advanced Texts: Basler Lehrbücher, Birkhäuser Verlag, Basel, 1993, xiv+705 Seiten.
[Bo08] Siegfried Bosch, *Lineare Algebra*, Springer-Lehrbuch, 4. überarbeitete Auflage, 2008, x+297 Seiten.
[Br03] Theodor Bröcker, *Lineare Algebra und analytische Geometrie*, Birkhäuser, 2003, x+266 Seiten.
[Ko83] Max Koecher, *Lineare Algebra und analytische Geometrie*, Springer, 1983, xi+286 Seiten.
[Wo08] Jürgen Wolfart, *Geometrie*, Vorlesungsskript aus dem Sommersemester 2008, GU Frankfurt.

Teil 1. Affine und projektive Geometrie

Affine Geometrie und projektive Geometrie kommen noch ohne Längen und Winkel aus. Dies macht diese Geometrien flexibler, aber auch weniger geometrisch.

1. Ebene Inzidenz-Geometrie

1.1. Die Fano-Ebene. Zweidimensionale Geometrie, man spricht auch von der Geometrie der Ebene, besteht aus Punkten, Geraden und einer Inzidenzrelation, so daß die Axiome der ebenen Inzidenz-Geometrie gelten.

Beispiel 1.1. Wir beginnen mit dem Bild einer Ebene bestehend aus 7 Punkten und 7 Geraden. Da eine „Gerade“ die Form eines Kreises hat, ist bereits klar, daß wir die Begriffe Punkt und Gerade abstrahieren wollen.

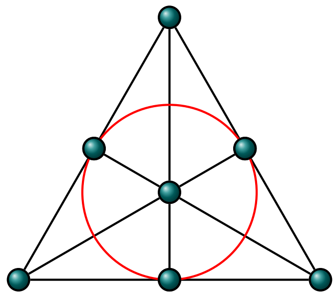


Abbildung 1. Die Fano-Ebene.

Die Punkte der Fano-Ebene sind die 7 Punkte des Bildes. Die Geraden sind symbolisiert durch die Dreiecksseiten, die Seitenhalbierenden und durch den (roten) Inkreis. Jede der Geraden enthält 3 Punkte.

Mit einem Blick auf die Skizze erkennen wir die Fano-Ebene als ein Beispiel für eine ebene Inzidenz-Geometrie, die wie folgt definiert ist.

Definition 1.2 (Ebene Inzidenz-Geometrie)**.** Eine **ebene Inzidenz-Geometrie** besteht aus einer Menge $\mathscr{P}$, einer Menge $\mathscr{G}$ und einer Inzidenzrelation genannte Teilmenge

$$I \subseteq \mathscr{P} \times \mathscr{G}.$$

Ein $P \in \mathscr{P}$ heißt **Punkt**, ein $g \in \mathscr{G}$ heißt **Gerade**. Wir schreiben

$$P \in g :\iff (P, g) \in I$$

und sagen **„P liegt auf g“** oder **„g geht durch P“**. Ansonsten schreiben wir $P \notin g$, wenn $(P, g) \notin I$. In einer ebenen Inzidenz-Geometrie gelten die folgenden Axiome:

(I1) Durch zwei Punkte geht genau eine Gerade: zu $P, Q \in \mathscr{P}$, $P \neq Q$ gibt es genau ein $g \in \mathscr{G}$ mit $P \in g$ und $Q \in g$.

(I2) Auf jeder Geraden $g \in \mathscr{G}$ liegen mindestens 2 Punkte.

(I3) Es gibt $P \in \mathscr{P}$ und $g \in \mathscr{G}$ mit $P \notin g$.

Notation 1.3. In einer Inzidenz-Geometrie bezeichne (AB) die nach Axiom (I1) eindeutige Gerade durch die Punkte $A \neq B$.

Eine Gerade wird durch die Menge der auf ihr liegenden Punkte eindeutig festgelegt.

Lemma 1.4. *Die Zuordnung* $\pi : \mathscr{G} \to \{X \ ; \ X \subseteq \mathscr{P}\}$,

$$\pi(g) = \{P \ ; \ P \in g\}$$

ist injektiv. Es gilt sogar für $g, h \in \mathscr{G}$

$$\pi(g) \subseteq \pi(h) \implies g = h.$$

Beweis. Angenommen die Geraden g und h haben $\pi(g) \subseteq \pi(h)$. Nach Axiom (I2) gibt es $P, Q \in g$, $P \neq Q$. Dann gilt aber auch $P, Q \in h$. Axiom (I1) zeigt $g = h$. $\square$

Lemma 1.4 besagt, daß man in der Definition einer ebenen Inzidenz-Geometrie die Geraden $\mathscr{G}$ durch eine Menge von Teilmengen von $\mathscr{P}$ ersetzen könnte, so daß $(P, g) \in I \iff P \in g$. Daher auch unsere Notation.

$\boxed{\text{Ab jetzt betrachten wir eine Gerade } g \text{ als eine Teilmenge } g \subseteq \mathscr{P} \text{ der Menge der Punkte.}}$

Definition 1.5. Wir legen weiter die folgende Terminologie fest:

(1) Wenn für $P \in \mathscr{P}$ und $g, h \in \mathscr{G}$ gilt $P \in g$ und $P \in h$, dann sagen wir, g und h **schneiden sich** im **Schnittpunkt** P.
(2) Eine Menge von Punkten $M \subseteq \mathscr{P}$ mit $|M| \geq 3$ heißt **kollinear**, wenn es eine Gerade g gibt mit $P \in g$ für alle $P \in M$.

Lemma 1.6. *Zwei Geraden schneiden sich in höchstens einem Punkt: zu $g \neq h \in \mathscr{G}$ ist*

$$|\{P \in \mathscr{P} \ ; \ P \in g \textit{ und } P \in h\}| \leq 1.$$

Beweis. Das ist eine unmittelbare Konsequenz von Axiom (I1). $\square$

Das Axiom (I3) verhindert, daß alle Punkte kollinear sind.

Proposition 1.7. *Die Menge aller Punkte $\mathscr{P}$ ist nicht kollinear und $|\mathscr{P}| \geq 3$.*

Beweis. Nach Axiom (I3) gibt es P und g mit $P \notin g$. Da g nach Axiom (I2) mindestens zwei weitere Punkte hat, folgt $|\mathscr{P}| \geq 3$.

Wenn $\mathscr{P}$ kollinear ist, gibt es demnach eine Gerade g, die alle Punkte enthält. Nach Lemma 1.4 gibt es dann überhaupt nur die eine Gerade g. Dies ist ein Widerspruch zu Axiom (I3). $\square$

2. Projektive Geometrie

Historisch gesehen spielt die Frage nach parallelen Geraden eine gewisse Rolle in der Frage der Axiomatisierung von Geometrie.

Definition 2.1. Zwei Geraden g, h sind **parallel**, wenn entweder $g = h$ oder wenn g und h keinen Schnittpunkt haben. Wir notieren g, h parallel als $g \parallel h$.

In der projektiven Ebene gibt es keine parallelen Geraden.

Definition 2.2. Eine **projektive Ebene** ist eine ebene Inzidenz-Geometrie mit Punkten $\mathscr{P}$ und Geraden $\mathscr{G}$, für die neben den Axiomen (I1)–(I3) die folgenden Eigenschaften gelten:

(P1) Je zwei Geraden schneiden sich.

(P2) Auf jeder Geraden liegen mindestens 3 Punkte.

Beispiel 2.3. Mit einem Blick auf Abbildung 1 erkennen wir die Fano-Ebene als ein Beispiel für eine projektive Ebene.

Lemma 2.4. *In einer projektiven Ebene schneiden sich zwei verschiedene Geraden in genau einem Punkt.*

Beweis. Nach Axiom (P1) schneiden sich die Geraden g, h in mindestens einem Punkt. Wenn es mehr als einer wäre, dann ist $g = h$ nach dem ersten Axiom bzw. nach Lemma 1.6. $\square$

2.1. **Der projektive Raum.** Sei K ein Körper. Der projektive Raum $\mathbb{P}^n(K)$ der Dimension n ist die Menge

$$\mathbb{P}^n(K) = \{P \subseteq K^{n+1} \ ; \ P \text{ ist } K\text{-Unterraum der } \dim_K(P) = 1\}$$

der Ursprungsgeraden in K^{n+1} und kann durch homogene Koordinaten

$$\big(K^{n+1} \setminus \{0\}\big)/K^\times \xrightarrow{\sim} \mathbb{P}^n(K), \quad [x_0 : \ldots : x_n] \mapsto \langle x \rangle_K = K \cdot x$$

mit $x = (x_0, \ldots, x_n)^t$ beschrieben werden. Hier ist $[x_0 : \ldots : x_n] = [\lambda x_0 : \ldots : \lambda x_n]$, für alle $\lambda \in K^\times$ die Restklasse des Spaltenvektors mit den Koordinaten $x_0, \ldots, x_n$.

Definition 2.5. Die projektive Ebene mit Koordinaten aus dem Körper K besteht aus der Menge von Punkten $\mathbb{P}^2(K)$ ausgestattet mit den folgenden Geraden. Zu jedem Unterraum $V \subseteq K^3$ der Dimension $\dim_K(V) = 2$ gehört die **Gerade**

$$\mathbb{P}(V) = \{P \in \mathbb{P}^2(K) \ ; \ P \in V\}$$

bestehend aus den Ursprungsgeraden in V.

Proposition 2.6. *Zu jeder Gerade $\mathbb{P}(V) \subseteq \mathbb{P}^2(K)$ gibt es $a, b, c \in K$ nicht alle 0 mit*

$$[x : y : z] \in \mathbb{P}(V) \iff ax + by + cz = 0.$$

Die Koeffizienten der linearen Gleichung sind eindeutig bis auf Skalieren: a, b, c und a', b', c' beschreiben die gleiche Gerade genau dann, wenn es ein $\lambda \in K^\times$ gibt mit

$$(a, b, c) = \lambda(a', b', c') \in \mathrm{M}_{1\times 3}(K).$$

Beweis. Es ist

$$[x : y : z] \in \mathbb{P}(V) \iff \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in V.$$

Jeder 2-dimensionale Unterraum $V \subseteq K^3$ ist Kern einer surjektiven linearen Abbildung

$$\pi_V : K^3 \twoheadrightarrow K^3/V \simeq K,$$

in Matrixschreibweise

$$\pi_V(\begin{pmatrix} x \\ y \\ z \end{pmatrix}) = \begin{pmatrix} a & b & c \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = ax + by + cz$$

für $a, b, c \in K^3$. Die Abbildung π_V ist surjektiv, wenn die Matrix $(a, b, c) \neq 0$ ist.

Zwei homogene nichttriviale lineare Gleichungen $ax + by + cz = 0$ und $a'x + b'y + c'z = 0$ haben denselben Kern, wenn die Matrix beider Gleichungen zusammen

$$\begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix}$$

immer noch den Rang 1 hat, d.h. die Zeilen linear abhängig sind. □

Proposition 2.7. *Mit dieser Struktur von Geraden ist $\mathbb{P}^2(K)$ eine projektive Ebene.*

Beweis. (i) Seien $P, Q \in \mathbb{P}^2(K)$, $P \neq Q$ zwei Punkte. Dann sind die zugehörigen 1-dimensionalen Unterräume $P, Q \subseteq K^3$ linear unabhängig. Eine Gerade zu $V \subseteq K^3$ enthält P und Q genau dann, wenn $P + Q \subseteq V$. Aus Dimensionsgründen erfüllt dies nur die Gerade zum Unterraum $V = P \oplus Q$.

(ii) Seien $V, W \subseteq K^3$ zwei Unterräume der Dimension 2. Die zugehörigen Geraden schneiden sich, wenn es einen 1-dimensionalen Unterraum $P \subseteq V \cap W$ gibt. Die Dimensionsformel zeigt

$$\dim_K(V \cap W) = \dim_K(V) + \dim_K(W) - \dim_K(V + W) \geq 4 - \dim_K(K^3) = 1.$$

Eine solche Ursprungsgerade gibt es also.

(iii) Die Anzahl der Punkte P auf der Geraden $\mathbb{P}(V) \simeq \mathbb{P}^1(K)$ ist $|K|+1$, denn $V \simeq K^2$ und damit $\mathbb{P}(V) \simeq \mathbb{P}(K^2) = \mathbb{P}^1(K)$, was durch die disjunkte Vereinigung

$$\mathbb{P}^1(K) = \{[x:1] \ ; \ x \in K\} \cup \{[1:0]\}$$

beschrieben wird. Weil jeder Körper $|K| \geq 2$ Elemente hat, gibt es mindestens 3 Punkte auf jeder Geraden.

(iv) Die Punkte mit den homogenen Koordinaten

$$[1:0:0], \ [0:1:0], \ [0:0:1]$$

liegen nicht auf einer Geraden von $\mathbb{P}^2(K)$. □

Beispiel 2.8. Die Fano-Ebene aus Abbildung 1 ist isomorph zu $\mathbb{P}^2(\mathbb{F}_2)$. Die Fano-Ebene ist die projektive Ebene der Form $\mathbb{P}^2(K)$ mit der kleinsten Anzahl von Punkten.

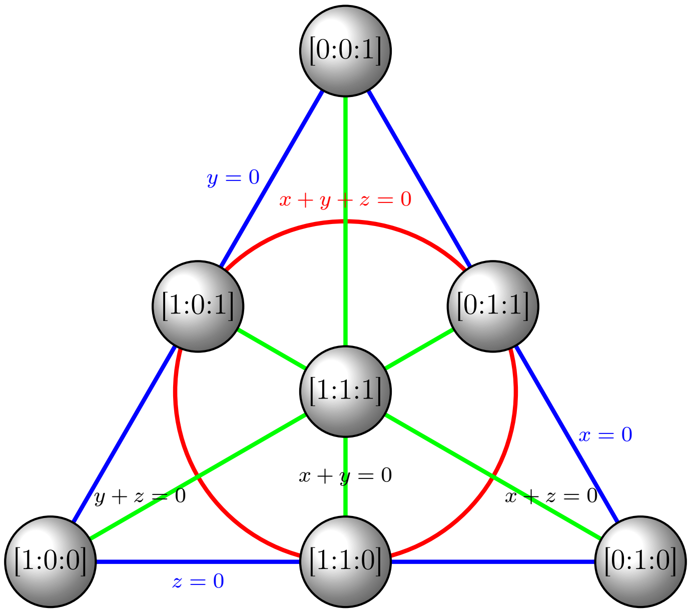


ABBILDUNG 2. Die projektive Ebene $\mathbb{P}^2(\mathbb{F}_2)$.

Die drei Geraden, welche als Dreiecksseiten erscheinen, haben die Gleichungen $x=0$, sowie $y=0$ bzw. $z=0$ für $P=[x:y:z]$. Die drei Geraden, welche als Ecktransversalen[1] auftreten, haben die Gleichungen $x+y=0$, sowie $y+z=0$ bzw. $z+x=0$. Schließlich hat der Kreis die Geradengleichung $x+y+z=0$. Die Gleichungen sind jeweils als Linearformen auf $(\mathbb{F}_2)^3$ gedacht, die einen 2-dimensionalen Unterraum ausschneiden, der seinerseits in $\mathbb{P}^2(\mathbb{F}_2)$ eine Gerade definiert.

Für allgemeines n ist $\mathbb{P}^n(K)$ eine **n-dimensionale Geometrie**. Hier gibt es ausgezeichnete lineare Teilräume von jeder Dimension $0 \leq d \leq n$. Die **d-dimensionalen linearen Unterräume** sind zu einem Unterraum $V \subseteq K^{n+1}$ mit $\dim_K(V) = d+1$ definiert als die Menge der Punkte

$$\mathbb{P}(V) = \{P \in \mathbb{P}^n(K) \ ; \ P \subseteq V\}.$$

[1]Eine Ecktransversale ist in einem Dreieck eine Gerade durch eine Ecken und einen Punkt der gegenüberliegenden Seite.

2.2. **Die projektiv lineare Gruppe.** Man versteht eine Geometrie besser[2], wenn man die geometrieerhaltenden Abbildungen versteht. Auf dem $\mathbb{P}^n(K)$ operiert die projektiv lineare Gruppe

$$\mathrm{PGL}_{n+1}(K) = \mathrm{GL}_{n+1}(K)/K^\times$$

durch $P \mapsto A(P)$ für $A \in \mathrm{GL}_{n+1}(K)$. Der Punkt P ist hier als Unterraum der Dimension 1 aufzufassen und für das Element in $\mathrm{PGL}_{n+1}(K)$ ist ein Vertreter aus $\mathrm{GL}_{n+1}(K)$ zu wählen. Dann ist $A(P)$ das Bild von P unter der linearen Abbildung „Multiplikation mit A“ und auch von Dimension 1, weil Multiplikation mit A ein K-linearer Isomorphismus ist. Das Bild $A(P)$ von P als Gerade in K^{n+1} ist unabhängig von der Wahl des Vertreters A. Offensichtlich werden d-dimensionale lineare Unterräume in ebensolche abgebildet. Und Inzidenzrelationen (ein linearer Raum ist in einem anderen enthalten) bleiben erhalten.

Satz 2.9. *Sei K ein Körper. Die Gruppe $\mathrm{PGL}_2(K)$ operiert exakt 3-fach transitiv auf den Punkten von $\mathbb{P}^1(K)$. Das bedeutet: für drei paarweise verschiedene Punkte $P_0, P_1, P_\infty \in \mathbb{P}^1(K)$ und drei weitere paarweise verschiedene Punkte P_0', P_1', P_∞' gibt es genau ein $A \in \mathrm{PGL}_2(K)$ mit*

$$A(P_i) = P_i'$$

für alle $i = 0, 1, \infty$.

Beweis. Seien $0 = [0:1]$, $1 = [1:1]$ und $\infty = [1:0]$. Es reicht zu zeigen, daß man für drei paarweise verschiedene Punkte $P_0, P_1, P_\infty \in \mathbb{P}^1(K)$ genau ein $A \in \mathrm{PGL}_2(K)$ findet mit $A(i) = P_i$ für alle $i = 0, 1, \infty$.

Sei $P_0 = [b:d]$ und $P_\infty = [a:c]$. Dann sind wegen $P_0 \neq P_\infty$ die Repräsentanten $\binom{a}{c}$ und $\lambda\binom{b}{d}$ für alle $\lambda \in K^\times$ eine Basis von K^2 und

$$A = \begin{pmatrix} a & \lambda b \\ c & \lambda d \end{pmatrix} \in \mathrm{GL}_2(K).$$

Es gilt $A(0) = P_0$ und $A(\infty) = P_\infty$. Sei $P_1 = [x:y]$. Wir müssen λ nun so wählen, daß

$$[a + \lambda b : c + \lambda d] = [x:y].$$

Dies gelingt durch die Lösung der Gleichung

$$(a + \lambda b)y = (c + \lambda d)x$$

mittels

$$\lambda = \frac{ay - cx}{dx - by}.$$

Weder Zähler noch Nenner sind 0, weil

$$[x:y] = P_1 \neq P_\infty = [a:c] \quad \text{und} \quad [b:d] = P_0 \neq P_1 = [x:y].$$

□

Definition 2.10. Punkte $P_0, \ldots, P_d$ in $\mathbb{P}^n(K)$ heißen **in allgemeiner Lage**, wenn der von den entsprechenden Ursprungsgeraden in K^{n+1} aufgespannte Raum die Dimension $d+1$ hat. Das bedeutet, daß jeder lineare Unterraum $L \subseteq \mathbb{P}^n(K)$, der alle P_i, $i = 0, \ldots, d$ enthält, selbst mindestens die Dimension d haben muß. (Die Dimension als linearer Unterraum von $\mathbb{P}^n(K)$ ist definitionsgemäß um 1 kleiner als die des zugehörigen Untervektorraums von K^{n+1}.)

Proposition 2.11. *Die Gruppe $\mathrm{PGL}_{n+1}(K)$ operiert transitiv auf $n+1$ Punkten in allgemeiner Lage in $\mathbb{P}^n(K)$.*

Beweis. Sind $P_0, \ldots, P_n$ in allgemeiner Lage, so gilt $K^{n+1} = P_0 \oplus \ldots \oplus P_n$. Sind $Q_0, \ldots, Q_n$ weitere Punkte in allgemeiner Lage, so ist auch $K^{n+1} = Q_0 \oplus \ldots \oplus Q_n$, und es gibt eine lineare Abbildung $A : K^{n+1} \to K^{n+1}$, die die Summanden P_i isomorph auf die Summanden Q_i abbildet. Man wähle entsprechende Basen $p_i \in P_i$ und $q_i \in Q_i$ und definiere A durch $A(p_i) = q_i$. □

Den folgenden Satz beweisen wir nicht.

[2]Das ist eine Doktrin, die im Wesentlichen auf Felix Klein zurückgeht

Theorem 2.12 (Hauptsatz der projektiven Geometrie). *Sei K ein Körper und $n \in \mathbb{N}$, $n \geq 2$. Jede bijektive Abbildung $\mathbb{P}^n(K) \to \mathbb{P}^n(K)$, die Geraden auf Geraden abbildet (eine* ***Kollineation****), wird von einem Körperautomorphismus $\sigma : K \to K$ und einer linearen Abbildung $A \in \mathrm{PGL}_{n+1}(K)$ induziert als*

$$[x_0 : \ldots : x_n] \mapsto A([\sigma(x_0) : \ldots : \sigma(x_n)]).$$

2.3. **Der Satz von Desargues und der Satz von Pappos.** Wir behandeln nun wichtige Sätze, die in der projektiven Ebene $\mathbb{P}^2(K)$ gelten.

Definition 2.13. Wir sagen, die Punkte A, B und C bilden die Ecken eines (nichtdegenerierten) **Dreiecks**, wenn A, B und C paarweise verschieden und nicht kollinear sind. Wir schreiben für das Dreieck mit den Ecken A, B und C einfach $\Delta(ABC)$.

Proposition 2.14. *Seien $A, B, C \in \mathbb{P}^2(K)$. Dann sind äquivalent:*

(a) *A, B und C bilden ein Dreieck.*
(b) *A, B und C sind in allgemeiner Lage.*
(c) *Als Unterräume von K^3 gilt $A \oplus B \oplus C = K^3$.*

Beweis. Das ist eine Übungsaufgabe, weil es fast sofort aus der Definition folgt. □

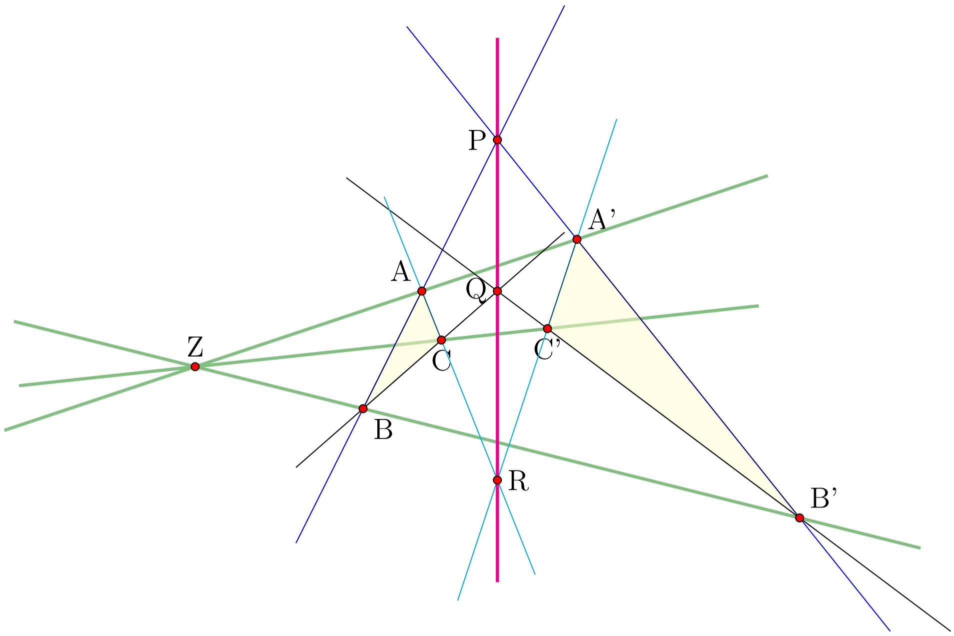


Abbildung 3. Satz von Desargues.

Satz 2.15 (Desargues). *Gegeben seien zwei nichtdegenerierte Dreiecke $\Delta(ABC)$ und $\Delta(A'B'C')$ in $\mathbb{P}^2(K)$. Ferner seien die Ecken A, B, C, A', B', C' paarweise verschieden und die Geraden durch entsprechende Dreiecksseiten seien verschieden:*

$$(AB) \neq (A'B'), \quad (BC) \neq (B'C'), \quad \text{und} \quad (AC) \neq (A'C').$$

Dann sind die folgenden Aussagen äquivalent:

(a) *Die Dreiecke ABC und $A'B'C'$ haben bezüglich eines Punktes $Z \in \mathbb{P}^2(K)$ perspektivische Lage: d.h. es gibt einen Punkt $Z \in \mathbb{P}^2(K)$, der mit A, A', mit B, B' und mit C, C' jeweils kollinear ist.*
(b) *Die Schnittpunkte $P = (AB) \cap (A'B')$, $Q = (BC) \cap (B'C')$ und $R = (AC) \cap (A'C')$ liegen auf einer Geraden.*

Beweis. Wir sprechen die Koordinaten im $\mathbb{P}^2(K)$ als $[x : y : z]$ an. Nach Anwendung einer projektiv linearen Transformation, die an den Aussagen nichts ändert, dürfen wir annehmen, daß

$$A' = [1 : 0 : 0],\ B' = [0 : 1 : 0],\ C' = [0 : 0 : 1].$$

Dann sieht man auf einen Blick die Geradengleichungen für $(A'B')$ etc.:

$$\begin{aligned}(A'B') &= \{[x : y : z]\ ;\ z = 0\},\\ (B'C') &= \{[x : y : z]\ ;\ x = 0\},\\ (A'C') &= \{[x : y : z]\ ;\ y = 0\}.\end{aligned}$$

Weiter setzen wir

$$A = [a_1 : a_2 : a_3],\ B = [b_1 : b_2 : b_3],\ C = [c_1 : c_2 : c_3]$$

und betrachten zugehörige Basen der Ursprungsgeraden

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix},\ \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix},\ \vec{c} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}.$$

Weil das Dreieck ABC nicht ausgeartet ist, ist $(\vec{a}, \vec{b}, \vec{c})$ eine Basis von K^3 und die Matrix $[\vec{a}, \vec{b}, \vec{c}] \in \mathrm{M}_3(K)$ (in Spaltenschreibweise) ist invertierbar:

$$\det([\vec{a}, \vec{b}, \vec{c}]) \neq 0.$$

Der Punkt P liegt in der Geraden $(A'B')$ gegeben durch die Gleichung $z = 0$ und hat als Linearkombination von $\vec{a}$ und $\vec{b}$ einen Vertreter

$$\vec{p} = a_3 \cdot \vec{b} - b_3 \cdot \vec{a} = [\vec{a}, \vec{b}, \vec{c}] \begin{pmatrix} -b_3 \\ a_3 \\ 0 \end{pmatrix},$$

denn dies ist die einzige Linearkombination von $\vec{b}$ und $\vec{a}$ mit $z = 0$. Analog haben Q und R Vertreter

$$\vec{q} = b_1 \cdot \vec{c} - c_1 \cdot \vec{b} = [\vec{a}, \vec{b}, \vec{c}] \begin{pmatrix} 0 \\ -c_1 \\ b_1 \end{pmatrix}, \qquad \vec{r} = c_2 \cdot \vec{a} - a_2\vec{c} = [\vec{a}, \vec{b}, \vec{c}] \begin{pmatrix} c_2 \\ 0 \\ -a_2 \end{pmatrix}.$$

Es sind P, Q und R kollinear genau dann, wenn

$$\det([\vec{q}, \vec{r}, \vec{p}]) = 0.$$

Es gilt $[\vec{q}, \vec{r}, \vec{p}] = [\vec{a}, \vec{b}, \vec{c}] \cdot M$ mit

$$M = \begin{pmatrix} & c_2 & -b_3 \\ -c_1 & & a_3 \\ b_1 & -a_2 & \end{pmatrix}.$$

Die Geradengleichungen für (AA') etc. sind

$$\begin{aligned}(AA') &= \{[x : y : z]\ ;\ a_3y - a_2z = 0\},\\ (BB') &= \{[x : y : z]\ ;\ b_1z - b_3x = 0\},\\ (CC') &= \{[x : y : z]\ ;\ c_2x - c_1y = 0\}.\end{aligned}$$

Dies sind in der Tat Geraden, denn z.B. verschwinden nicht $a_2 = a_3 = 0$, weil sonst $A = A'$ wäre. Diese drei Geraden gehen durch einen gemeinsamen Punkt, wenn das 3×3 Gleichungssystem

$$\left\{\begin{array}{rrrcl} & a_3y & -a_2z & = & 0 \\ -b_3x & & +b_1z & = & 0 \\ c_2x & -c_1y & & = & 0 \end{array}\right.$$

eine nichttriviale Lösung hat. Dies ist genau dann der Fall, wenn $\det(N) = 0$ für

$$N = \begin{pmatrix} & a_3 & -a_2 \\ -b_3 & & b_1 \\ c_2 & -c_1 & \end{pmatrix}$$

Nun setzen wir alles zusammen:

$$\begin{aligned} \text{Aussage (a)} &\iff 0 = \det(N) = b_1c_2a_3 - c_1a_2b_3 \\ &\iff b_1c_2a_3 - c_1a_2b_3 = \det(M) = 0 \\ &\iff \det([\vec{q},\vec{r},\vec{p}]) = \det(M) \cdot \det([\vec{a},\vec{b},\vec{c}]) = 0 \iff \text{Aussage (b).} \end{aligned}$$ □

Bemerkung 2.16. Der Satz von Desargues hat einen wunderbaren dreidimensionalen Beweis. Wir skizzieren diesen für (a) $\Longrightarrow$ (b). Dazu stellen wir uns im Bild unten die Punkte C und C' aus der Ebene heraus über die Zeichnung gezogen.

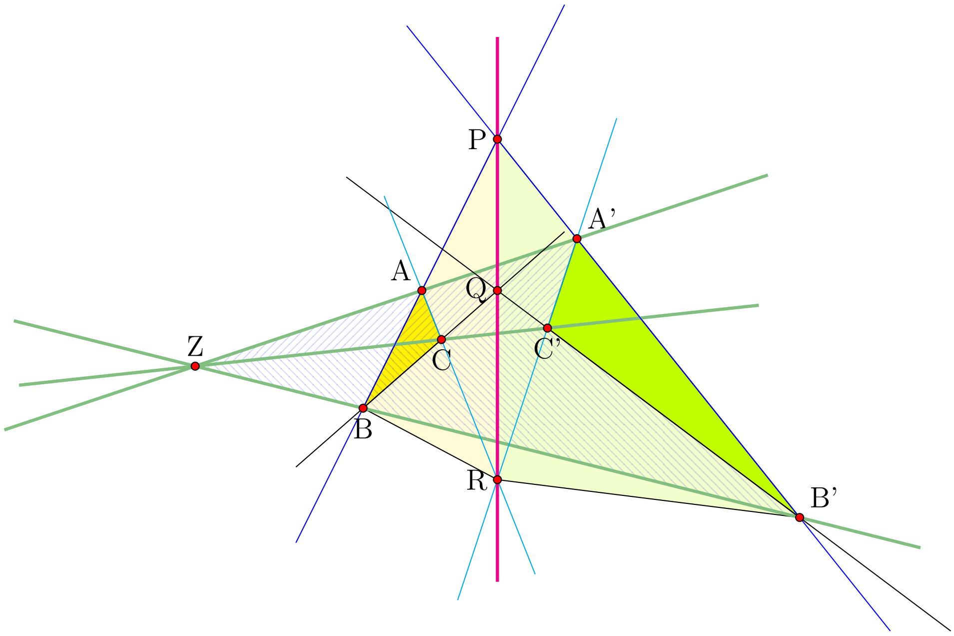


Abbildung 4. 3D - Desargues.

Damit wird Z die Spitze eines Pyramidenstumpfs mit dem grünen Dreieck $\Delta(A'B'C')$ als Basis. Das grüne Dreieck ist auch der Schnitt der hellgrünen Ebene mit dem Pyramidenstumpf. Das gelbe Dreieck $\Delta(ABC)$ ist der Schnitt der hellgelben Ebene mit dem Pyramidenstumpf.

Die Gerade durch P, Q und R liegt auch oberhalb der Zeichenebene als Schnitt der hellgelben Ebene E und der hellgrünen Ebene E'. In der Tat schneiden sich die Geraden (AB) und $(A'B')$, und das ist im dreidimensionalen nicht selbstverständlich, in einem Punkt P, weil nach Voraussetzung A', B' in der vom Dreieck $\Delta(ABZ)$ bestimmten Ebene liegen. Gleiches gilt für $Q = (BC) \cap (B'C')$ und $R = (AC) \cap (A'C')$. Der Schnittpunkt P (bzw. Q und R) liegt auf der Geraden (AB) und damit in der Ebene E und gleichzeitig auf der Geraden $(A'B')$ und damit in der Ebene E'. Somit liegen P, Q, R gemeinsam auf der Schnittgeraden $E \cap E'$. Zurückprojiziert in die Ebene bleibt diese Kollinearität erhalten.

Wenn man hier konsequent im $\mathbb{P}^3(K)$ arbeitet, vermeidet man auch, Sonderfälle betrachten zu müssen (etwa parallele Ebenen).

Bemerkung 2.17. Eine projektive Ebene $\mathscr{P}$, in welcher der Satz von Desargues gilt, erlaubt Koordinaten aus einem Schiefkörper D, d.h. es gibt einen geometrieerhaltenden Isomorphismus

$$\mathscr{P} \simeq \mathbb{P}^2(D).$$

Dazu muß man zwar zunächst die Konstruktion $\mathbb{P}^2(D)$ auf Schiefkörper D (nicht notwendigerweise kommutative Körper) ausdehnen, aber das ist nur eine formale Sache. Der Satz von Desargues gilt auch für $\mathbb{P}^2(D)$ und jeden Schiefkörper D.

Wir wenden uns nun dem Satz von Pappos zu (Es gibt auch die Schreibweise Pappus).

Satz 2.18 (Pappos)**.** *Gegeben seien Geraden $g \neq g'$ im $\mathbb{P}^2(K)$ mit sechs paarweise verschiedenen Punkten $A, B, C \in g$ und $A', B', C' \in g'$, von denen keiner der Schnittpunkt $S = g \cap g'$ ist. Dann sind die Schnittpunkte*

$$P = (AB') \cap (A'B), \quad Q = (BC') \cap (B'C), \quad \textit{und} \quad R = (AC') \cap (A'C)$$

paarweise verschieden und kollinear.

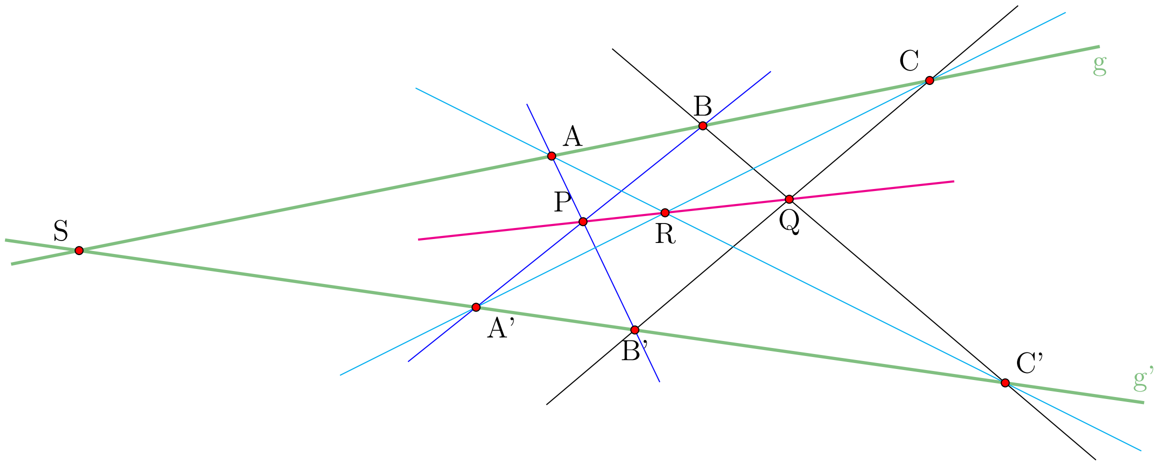


Abbildung 5. Satz von Pappos.

Beweis. Nach Anwendung einer projektiv linearen Transformation, die an den Aussagen nichts ändert, dürfen wir annehmen, daß

$$A = [1:0:0], \; A' = [0:1:0], \; S = [0:0:1].$$

Die Gerade g ist in Koordinaten $[x:y:z] \in \mathbb{P}^2(K)$ beschrieben durch $y = 0$ und die Gerade g' durch $x = 0$.

Also ist $B = [b_1 : 0 : b_3]$ und $B' = [0 : b_2' : b_3']$ mit $b_1, b_3, b_2', b_3' \neq 0$. Nach eventuellem Skalieren mit einer Diagonalmatrix darf man $b_1 = b_3$ und $b_2' = b_3'$ annehmen. Damit ist

$$B = [1:0:1], \quad \text{und } B' = [0:1:1].$$

Weiter gibt es $x, y \in K \setminus \{0,1\}$ mit

$$C = [x:0:1], \quad \text{und } C' = [0:y:1].$$

Eine kurze Rechnung, jeweils zum Schnitt zweier Ebenen in K^3, zeigt

$$\begin{aligned} P &= [1:1:1], \\ Q &= [x(1-y) : y(1-x) : 1-yx] \\ R &= [x:y:1]. \end{aligned}$$

Damit sind augenscheinlich P, Q und R paarweise verschieden. Die Punkte P, Q und R sind kollinear genau dann, wenn die repräsentierenden Vektoren linear abhängig sind. Dazu rechnen wir aus

$$\det \begin{pmatrix} 1 & x & x(1-y) \\ 1 & y & y(1-x) \\ 1 & 1 & 1-yx \end{pmatrix} = \det \begin{pmatrix} 1 & x & -xy \\ 1 & y & -yx \\ 1 & 1 & -yx \end{pmatrix} = 0$$

(wir haben die zweite Spalte von der dritten abgezogen), weil

$$xy = yx$$

und so die dritte Spalte ein Vielfaches der ersten Spalte ist. □

Bemerkung 2.19. Man kann den Satz von Desargues aus dem Satz von Pappos herleiten. Daher gibt es in einer projektive Ebene $\mathscr{P}$, in welcher der Satz von Pappos gilt, wieder Koordinaten. Allerdings kann man zeigen, daß für einen Schiefkörper D der Satz von Pappos in $\mathbb{P}^2(D)$ genau dann gilt, wenn D kommutativ ist. Genauer ist also $\mathscr{P} \simeq \mathbb{P}^2(K)$ für einen Körper K.

Mehr zu den Fragen der Koordinateneinführung bei einer projektiven Ebene finden Sie im Abschnitt §8 des Skripts [Wo08].

3. Affine Geometrie

3.1. **Affine Räume und die affine Ebene.**

Definition 3.1. Ein **affiner Raum** ist eine Menge A mit einer freien und transitiven Operation

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

durch einen Vektorraum V über einem Körper K. Es gilt für alle $a, b \in A$ und $v, w \in V$

$$0 + a = a,$$
$$v + (w + a) = (v + w) + a,$$
$$\exists! u \in V : u + a = b.$$

Der Vektorraum V heißt der **Vektorraum der Translationen** (oder **Translationsraum**) von A. Die **Dimension** von A ist definiert als die Dimension von V.

Um Ausnahmen zu vermeiden, betrachten wir die leere Menge $\emptyset$ auch als affinen Raum.

Beispiel 3.2. (1) Sei $d \in \mathbb{N}$. Die Menge aller normierten Polynome $P(X) \in K[X]$ vom Grad d

$$P(X) = X^d + a_{d-1}X^{d-1} + \ldots + a_1X + a_0$$

ist ein affiner Raum mit Translationen durch den K-Vektorraum der Polynome vom Grad kleiner als d.

(2) Der Lösungsraum einer inhomogenen linearen Gleichung ist ein affiner Raum. Dieser ist entweder leer oder besitzt Translationen durch den Lösungsraum des zugehörigen homogenen Gleichungssystems.

(3) Seien V ein K-Vektorraum und $U \subseteq V$ ein Unterraum. Für $a \in V$ ist die Nebenklasse $a + U$ ein affiner Raum mit Translationen durch U.

Definition 3.3. Sei K ein Körper und $n \in \mathbb{N}$. Dann ist

$$\mathbb{A}^n(K) = K^n$$

der **affine (Koordinaten-)Raum** der Dimension n. Der Vektorraum K^n operiert auf $\mathbb{A}^n(K)$ durch Translation mittels Vektorraumaddition.

Allgemeiner ist für einen K-Vektorraum V die Menge

$$\mathbb{A}(V) = V$$

ein affiner Raum mit Translationen durch V mittels Vektoraddition.

Definition 3.4. Ein Isomorphismus **affiner Räume** A mit Translationsraum V und B mit Translationsraum W besteht aus einer Bijektion

$$f : A \to B,$$

so daß für alle $v \in V$ und $a \in A$ die Formel

$$f(v + a) = \varphi(v) + f(a)$$

einen wohldefinierten K-Vektorraumisomorphismus $\varphi : V \to W$ definiert.

Proposition 3.5. *Jeder affine Raum A mit Vektorraum V der Translationen von $n = \dim_K(V)$ ist als affiner Raum isomorph zu $\mathbb{A}^n(K)$. Ein Isomorphismus hängt ab von der Wahl eines* ***Ursprungs*** *$a_0 \in A$ und einer Basis $\mathscr{B} = (v_1, \ldots, v_n)$ von V und wird durch die folgende Formel gegeben:*

$$f : \mathbb{A}^n(K) \to A, \qquad f(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}) = \Big(\sum_{i=1}^n x_i v_i\Big) + a_0.$$

Beweis. Dies drückt nur aus, daß die Operation transitiv ist, also nur eine Bahn $V + a_0$ hat, und daß die Operation frei ist: aus $v + a_0 = w + a_0$ folgt $v = w$. □

Bemerkung 3.6. Proposition 3.5 beschreibt, wie man auf einem affinen Raum Koordinaten einführen kann.

Definition 3.7. Ein **affiner Unterraum** eines affinen Raums A mit Translationen V ist eine Teilmenge von der Form $U + a$ für $a \in A$ und einem Unterraum $U \subseteq V$, oder aber die leere Menge.

Beispiel 3.8. Sei $f : V \to W$ eine lineare Abbildung. Dann sind die Urbildmengen $f^{-1}(y)$ zu $y \in W$ affine Unterräume von $V = \mathbb{A}(V)$. Der Translationsraum von $f^{-1}(y)$ ist $\ker(f)$, sofern $f^{-1}(y)$ nicht leer ist.

Lemma 3.9. *Der Schnitt von affinen Unterräumen ist ein affiner Unterraum.*

Beweis. Sei A der affine Raum mit Translationsraum V und seien $A_i = U_i + a_i$ die zu schneidenden affinen Unterräume. Wenn der Schnitt leer ist, haben wir nichts zu tun. Ansonsten wählen wir

$$x \in \bigcap_i A_i.$$

Dann ist für alle i der Unterraum $A_i = U_i + a_i = U_i + x$, denn Bahnen sind entweder gleich oder disjunkt. Dann ist

$$\bigcap_i A_i = \bigcap_i U_i + x = \Big(\bigcap_i U_i\Big) + x$$

ein affiner Unterraum mit dem Schnitt $\bigcap_i U_i$ als Translationsraum. □

Definition 3.10. Eine **affine Ebene** ist eine ebene Inzidenz-Geometrie mit Punkten $\mathscr{P}$ und Geraden $\mathscr{G}$, für die neben den Axiomen (I1)–(I3) das starke Parallelenaxiom gilt:

(A1) zu einem Punkt P und einer Gerade g gibt es **genau eine** Gerade h mit
- (i) g und h sind parallel, und
- (ii) $P \in h$.

Definition 3.11. Die **affine Ebene mit Koordinaten aus dem Körper** K besteht aus der Menge von Punkten $\mathbb{A}^2(K)$ ausgestattet mit den affinen Unterräumen der Dimension 1 als Geraden.

Lemma 3.12. *Eine Gerade in $\mathbb{A}^2(K)$ ist dasselbe wie der Lösungsraum einer nichttrivialen inhomogenen linearen Gleichung in zwei Variablen: zu $a_1, a_2, b \in K$ mit $(a_1, a_2) \neq (0, 0)$ die Lösungsmenge*

$$\{(x, y) \in \mathbb{A}^2(K)\ ;\ a_1 x + a_2 y = b\}.$$

Beweis. Trivial nach der Lösungstheorie (in-)homogener linearer Gleichungssysteme. □

Proposition 3.13. *Mit dieser Struktur von Geraden ist $\mathbb{A}^2(K)$ eine affine Ebene.*

Beweis. (i) Durch je zwei Punkte $P \neq Q$ geht genau eine Gerade, nämlich die Gerade

$$P + K \cdot v$$

mit dem eindeutigen $v \in K^2$ mit $Q = v + P$.

(ii) Jede Gerade ist in Bijektion mit einem eindimensionalen K-Vektorraum und hat demnach $|K|$-viele Punkte. Da K ein Körper ist, gibt es mindestens 2 Punkte auf jeder Geraden.

(iii) Schreiben wir einen Punkt in $\mathbb{A}^2(K)$ mit Koordinaten (x, y), $x, y \in K$. Auf der Geraden $x = 0$ liegt nicht der Punkt $(1, 0)$.

(iv) Es bleibt, das starke Parallelenaxiom nachzuweisen. Sei $g = a + V$ eine Gerade in $\mathbb{A}^2(K)$ und $b \in \mathbb{A}^2(K)$ ein weiterer Punkt. Dann ist $b + V$ eine Gerade durch b und

$$a + V \cap b + V = \emptyset \quad \text{oder} \quad a + V = b + V,$$

denn wenn der Schnitt nicht leer ist, dann gibt es $x, y \in V$ mit $a + x = b + y$. Somit folgt $b + z = a + (x - y + z) \in a + V$ für alle $z \in V$, also $b + V \subseteq a + V$ und analog umgekehrt. Kurz: die Bahnen $a + V$ und $b + V$ der Translationsoperation von V auf $\mathbb{A}^2(K)$ sind entweder disjunkt oder gleich.

Jede andere Gerade durch b hat die Form $b + W$ mit einem eindimensionalen Unterraum $W \neq V$. Folglich ist $V \cap W = (0)$ und nach der Dimensionsformel

$$\dim(V + W) = \dim(V) + \dim(W) - \dim(V \cap W) = 2.$$

Also ist $V + W = K^2$ und es gibt $v \in V$ und $w \in W$ mit $a - b = v + w$. Das bedeutet aber

$$x = a + (-v) = b + w \in a + V \cap b + W$$

und so hat diese Gerade einen Schnitt mit $a + V$. Dies zeigt die Eindeutigkeit. □

Für allgemeines n ist $\mathbb{A}^n(K)$ eine **n-dimensionale Geometrie** mit ausgezeichneten linearen Teilräumen von jeder Dimension $0 \leq d \leq n$. Die **d-dimensionalen linearen Unterräume** sind zu einem Unterraum $V \subseteq K^n$ mit $\dim_K(V) = d$ definiert als die Menge der Punkte $V \subseteq \mathbb{A}^n(K)$.

3.2. **Unendlich ferne Punkte.** Sei K ein Körper und $n \in \mathbb{N}$. Die Abbildung

$$\mathbb{A}^n(K) \to \mathbb{P}^n(K), \qquad \begin{pmatrix} x_1 \\ \ldots \\ x_n \end{pmatrix} \mapsto [x_1 : \ldots : x_n : 1]$$

ist injektiv und hat als Bild das Komplement des $n-1$-dimensionalen projektiven Raumes

$$\mathbb{P}^{n-1}(K) \simeq \{[x_0 : x_1 : \ldots : x_{n-1} : 0] \ ; \ x_0, \ldots, x_{n-1} \in K \text{ nicht alle } 0\} \subseteq \mathbb{P}^n(K)$$

Wir sprechen unter dieser Perspektive über $\mathbb{P}^n(K)$ als eine Vervollständigung von $\mathbb{A}^n(K)$ durch die **unendlich fernen Punkte** des $\mathbb{P}^{n-1}(K) \subseteq \mathbb{P}^n(K)$. Iteriert angewandt liefert diese Betrachtung eine disjunkte Vereinigung

$$\mathbb{P}^n(K) = \mathbb{A}^n(K) \cup \mathbb{A}^{n-1}(K) \cup \ldots \cup \mathbb{A}^1(K) \cup \mathbb{A}^0(K)$$

mit

$$\mathbb{A}^d(K) = \{[x_0 : x_1 : \ldots : x_{d-1} : 1 : 0 : \ldots : 0] \in \mathbb{P}^n(K)\}.$$

Speziell für $n = 2$ betten wir die affine Ebene in die projektive Ebene ein

$$\mathbb{A}^2(K) \subseteq \mathbb{P}^2(K)$$

mit den Vorteilen, daß sich zuvor nicht schneidende Geraden aus $\mathbb{A}^2(K)$ nun einen eindeutigen Schnittpunkt in einem unendlich fernen Punkt hinzugewinnen. Die Gerade

$$\mathbb{P}^1(K) \simeq \{[x : y : z] \ ; \ z = 0\} = \mathbb{P}^2(K) \setminus \mathbb{A}^2(K)$$

nennen wir die **unendlich ferne Gerade** der projektiven Ebene $\mathbb{P}^2(K)$.

Proposition 3.14. *Die Abbildung*

$$g = \{[x : y : z] \ ; \ ax + by + cz = 0\} \mapsto g \cap \mathbb{A}^2(K) = \{(x, y) \ ; \ ax + by = -c\}$$

vermittelt eine Bijektion

$$\big\{\text{Geraden } g \text{ in } \mathbb{P}^2(K) \ ; \ g \neq \infty\text{-ferne Gerade}\big\} \leftrightarrow \big\{\text{Geraden in } \mathbb{A}^2(K)\big\}.$$

Beweis. Es liegt (x, y) auf der Gerade mit Gleichung $ax + by = -c$ genau dann, wenn $[x : y : 1]$ auf der Gerade mit $ax + by + cz = 0$ liegt. Daher ist die Beschreibung der Geraden $g \cap \mathbb{A}^2(K)$ in der Proposition korrekt. Es kommt stets eine gültige Geradengleichung heraus, außer im Fall der unendlich fernen Geraden, die durch $z = 0$ gegeben ist. Offensichtlich ist die angegebene Abbildung auf Geradengleichungen eine Bijektion und verträglich mit Skalieren, also eine Bijektion auf den entsprechenden Geraden. □

Proposition 3.15. *Seien $g \neq h$ Geraden in $\mathbb{A}^2(K)$ und $\tilde{g}$ und $\tilde{h}$ die Geraden von $\mathbb{P}^2(K)$ mit $g = \tilde{g} \cap \mathbb{A}^2(K)$ und $h = \tilde{h} \cap \mathbb{A}^2(K)$. Dann sind äquivalent:*

(a) *g und h sind parallel.*

(b) *$\tilde{g}$ und $\tilde{h}$ schneiden sich in einem Punkt der unendlich fernen Geraden $g_\infty \subseteq \mathbb{P}^2(K)$.*

Beweis. Da $g \neq h$ vorausgesetzt ist, gilt

$$g \parallel h \iff g \cap h = \emptyset \iff \tilde{g} \cap \tilde{h} \cap \mathbb{A}^2(K) = \emptyset \iff \tilde{g} \cap \tilde{h} \subseteq g_\infty.$$

Da auch $\tilde{g} \neq \tilde{h}$ gilt, haben $\tilde{g}$ und $\tilde{h}$ in der projektiven Ebene genau einen Schnittpunkt. Folglich ist $\tilde{g} \cap \tilde{h} \subseteq g_\infty$ äquivalent zur Aussage (b). □

Die Sätze von Desargues und Pappos haben Varianten für $\mathbb{A}^2(K)$. Je nachdem, ob ein Punkt oder eine Gerade aus der Aussage im endlichen, d.h. in $\mathbb{A}^2(K)$, oder im unendlichen, d.h. in $\mathbb{P}^1(K) \simeq \mathbb{P}^2(K) \setminus \mathbb{A}^2(K)$ liegt, sehen die affinen Formulierungen ein wenig anders aus.

3.3. **Bewegungen des affinen Raumes.** Die **affin-linearen Bewegungen** des $\mathbb{A}^n(K)$ sind die Abbildungen der Form

$$x \mapsto Ax + b$$

mit $A \in \mathrm{GL}_n(K)$ und $b \in K^n$. Wenn $A = \mathbf{1}$ spricht man von der **Translation** mit b

$$T_b(x) = x + b.$$

Die affin-linearen Bewegungen bilden mit Komposition eine Gruppe, die **affin-lineare Gruppe** der Dimension n

$$\mathrm{Aff}^n(K) = \mathrm{GL}_n(K) \ltimes K^n,$$

wobei $\ltimes$ für das semi-direkte Produkt mit der Verknüpfung

$$(A, b)(C, d) = (AC, b + Ad)$$

steht. In der Tat ist für alle $x \in \mathbb{A}^n(K)$

$$(A, b).\big((C, d).x\big) = (A, b).(Cx + d) = A(Cx + d) + b = ACx + Ad + b = (AC, b + Ad).x$$

Als Komposition von Abbildungen ist die Verknüpfung assoziativ. Das neutrale Element ist $(\mathbf{1}, 0)$

$$(\mathbf{1}, 0).x = \mathbf{1}x + 0 = x.$$

Das Inverse zu (A, b) ist $(A^{-1}, -A^{-1}b)$

$$(A, b)(A^{-1}, -A^{-1}b) = (AA^{-1}, b + A(-A^{-1}b)) = (\mathbf{1}, b - b) = (\mathbf{1}, 0).$$

Proposition 3.16. *Die Translationen bilden eine Untergruppe von* $\mathrm{Aff}^n(K)$ *isomorph zur additiven Gruppe* $(K^n, +)$.

Beweis. Die Abbildung $b \mapsto T_b$ ist ein injektiver Homomorphismus $(K, +) \to \mathrm{Aff}^n(K)$ mit Bild bestehend aus genau der Menge der Translationen. Das folgt sofort aus der Rechnung

$$T_b \circ T_c(x) = (x + c) + b = x + (b + c) = T_{b+c}(x).$$

□

Affin-lineare Bewegungen erhalten Geraden und Schnittpunkte. Aber im Fall $K = \mathbb{R}$, in dem wir wesentlich mehr geometrische Begriffe wie Länge oder Winkel haben, erhält eine allgemeine affin-lineare Bewegung weder Winkel noch Längen. Dazu brauchen wir mehr und eine bilineare Struktur auf dem zugrundeliegenden Vektorraum, und darum kümmern wir uns als nächstes.

Teil 2. Bilinearformen

4. Paarungen von Vektorräumen

4.1. **Das Tensorprodukt.** Wir führen den folgenden multilinearen Begriff ein.

Definition 4.1. Sei K ein Körper und U, V, W seien K-Vektorräume. Eine **(K-)bilineare Abbildung** von U und V nach W ist eine Abbildung von Mengen

$$f : U \times V \to W$$

mit den folgenden Eigenschaften:

(i) Für $u \in U$ und $v_1, v_2 \in V$ gilt

$$f(u, v_1 + v_2) = f(u, v_1) + f(u, v_2).$$

(ii) Für $u_1, u_2 \in U$ und $v \in V$ gilt

$$f(u_1 + u_2, v) = f(u_1, v) + f(u_2, v).$$

(iii) Für $u \in U$, $v \in V$ und $\lambda \in K$ gilt

$$f(\lambda u, v) = f(u, \lambda v) = \lambda f(u, v).$$

Bemerkung 4.2. Man sagt kurz: die (bilineare) Abbildung ist ‚linear in beiden Argumenten', und meint, wenn man ein Argument fixiert, so ist die verbleibende Abbildung im anderen Argument linear. Das ist von der Determinante einer Matrix bekannt, die eine multilineare Abbildung als Funktion der Spalten ist.

Für eine bilineare Abbildung $f : U \times V \to W$ gilt für alle $u \in U$ und $v \in V$

$$f(u, 0) = 0 = f(0, v),$$

denn lineare Abbildungen bilden 0 auf 0 ab.

Beispiel 4.3. Sei K ein Körper und $n, m, r \in \mathbb{N}$. Dann ist Matrixmultiplikation

$$\mathrm{M}_{n\times m}(K) \times \mathrm{M}_{m\times r}(K) \to \mathrm{M}_{n\times r}(K)$$

eine K-bilineare Abbildung.

Es verhält sich nun wie mit linearen Abbildungen, daß eine bilineare Abbildung durch die Vorgabe auf einer Basis eindeutig festgelegt ist und jede solche Festlegung von einer bilinearen Abbildung herrührt.

Proposition 4.4. *Seien U, V, W Vektorräume über dem Körper K. Sei $\mathscr{A} = (a_1, \ldots, a_n)$ eine Basis von U und sei $\mathscr{B} = (b_1, \ldots, b_m)$ eine Basis von V. Dann ist eine bilineare Abbildung*

$$f : U \times V \to W$$

eindeutig durch die Werte

$$(f(a_i, b_j))_{1\le i\le n, 1\le j\le m} \in \prod_{i=1}^{n}\prod_{j=1}^{m} W$$

bestimmt. Jedes $n \times m$-Tupel von Werten wird von einer bilinearen Abbildung realisiert.

Beweis. Sei $(u, v) \in U \times V$ beliebig. Dann gibt es Koordinaten $x_1, \ldots, x_n$ und $y_1, \ldots, y_m$ aus K für $u = \sum_i x_i a_i$ und $v = \sum_j y_j b_j$, und damit gilt

$$f(u, v) = f(\sum_i x_i a_i, v) = \sum_i x_i f(a_i, v) = \sum_i x_i f(a_i, \sum_j y_j b_j) = \sum_{i,j} x_i y_j f(a_i, b_j).$$

Dies zeigt die Eindeutigkeit. Zur Existenz dreht man die Argumentation herum und definiert für Vektoren $w_{ij} \in W$, $1 \leq i \leq n$, $1 \leq j \leq m$

$$f(u,v) := \sum_{i,j} x_i y_j w_{ij}.$$

Dann ist $f(a_i, b_j) = w_{ij}$ und eine leichte Rechnung zeigt, daß ein solches f bilinear ist. $\square$

Definition 4.5. Das **Tensorprodukt** zweier K-Vektorräume U, V ist ein K-Vektorraum T zusammen mit einer bilinearen Abbildung $t : U \times V \to T$, so daß für alle bilinearen Abbildungen $f : U \times V \to W$ eine **eindeutige lineare Abbildung** $F : T \to W$ existiert, für die gilt:

$$f(u,v) = F(t(u,v)) \quad \text{für alle } u \in U, v \in V.$$

Es gilt mit dem üblichen formalen Beweis aufgrund der universellen Eigenschaft eine Eindeutigkeitsaussage (noch bevor wir von der Existenz wissen!).

Satz 4.6. *Das Tensorprodukt ist eindeutig bis auf eindeutigen Isomorphismus.*

Beweis. Sei $s : U \times V \to S$ ein weiteres Tensorprodukt. Dann faktorisiert die bilineare Abbildung s, weil T ein Tensorprodukt ist, über eine lineare Abbildung ψ, und analog faktorisiert t über eine lineare Abbildung φ wie im Diagramm:

$$\begin{tikzcd} & U \times V \arrow[dl, "t"'] \arrow[d, "s"] \arrow[dr, "t"] & \\ T \arrow[r, dotted, "\psi"] & S \arrow[r, dotted, "\varphi"] & T. \end{tikzcd}$$

Damit löst $\varphi \circ \psi$ das von der universellen Eigenschaft gestellte Faktorisierungsproblem für die bilineare Abbildung t in Bezug auf das Tensorprodukt $t : U \times V \to T$ genauso wie $\mathrm{id}_T : T \to T$. Die geforderte Eindeutigkeit erzwingt $\varphi \circ \psi = \mathrm{id}_T$. Aus Symmetrie folgt $\psi \circ \varphi = \mathrm{id}_S$. Dies zeigt, daß φ und ψ sogar zueinander inverse Isomorphismen sind und weiter die Eindeutigkeit des Tensorprodukts.

Außerdem hat man mit den Isomorphismen φ und ψ keine Wahl, wenn der Isomorphismus mit den universellen bilinearen Abbildungen s und t verträglich sein soll. $\square$

Satz 4.7. *Tensorprodukte existieren.*

Beweis. Wir konstruieren ein Tensorprodukt der Vektorräume U und V und beschränken uns auf den Fall endlicher Dimension (der unendlich-dimensionale Fall geht genauso).

Sei $\mathscr{A} = (a_1, \ldots, a_n)$ eine Basis von U und $\mathscr{B} = (b_1, \ldots, b_m)$ sei eine Basis von V. Als Tensorprodukt nehmen wir einen K-Vektorraum mit Basis bestehend aus formalen[3] Ausdrücken

$$a_i \otimes b_j,$$

das ist die direkte Summe

$$U \otimes V := \bigoplus_{ij} K \cdot a_i \otimes b_j$$

mit spezieller Notation für die Basis der eindimensionalen Summanden.

Die universelle bilineare Abbildung ist

$$t : U \times V \to U \otimes V, \qquad (\sum_i x_i a_i, \sum_j y_j b_j) \mapsto \sum_{ij} x_i y_j (a_i \otimes b_j).$$

[3]Der formale Ausdruck $a_i \otimes b_j$ ist nur eine komplizierte Notation für einen Vektor. Im Gegensatz zum Paar (a_i, b_j), das eine a priori Bedeutung hat und zudem zufällig auch im direkten Produkt $U \times V$ auftritt, hat $a_i \otimes b_j$ keine a priori Bedeutung. Den Vektor $a_i \otimes b_j$ gibt es nur als $t(a_i, b_j)$ in einem Tensorprodukt $t : U \times V \to U \otimes V$.

Dies ist die in Proposition 4.4 konstruierte bilineare Abbildung zu den Werten

$$t(a_i, b_j) = a_i \otimes b_j.$$

Sei $f : U \times V \to W$ eine beliebige bilineare Abbildung. Dann hat das einzige F, was in Frage kommt, die Form

$$F : U \otimes V \to W, \qquad F(a_i \otimes b_j) = f(a_i, b_j).$$

Damit ist F durch Angabe der Bilder einer Basis eindeutig festgelegt. Und Proposition 4.4 zeigt $F \circ t = f$ durch den Vergleich der Auswertung auf den verschiedenen (a_i, b_j). □

Notation 4.8. „Das“ Tensorprodukt bezeichnen wir mit $U \otimes V$ (oder auch genauer $U \otimes_K V$) und die universelle bilineare Abbildung schreiben wir als

$$(u, v) \mapsto u \otimes v,$$

nicht nur für die Basisvektoren. Bilinearität bedeutet nun, daß für alle $u, u_1, u_2 \in U$ und $v, v_1, v_2 \in V$ und $\lambda \in K$ gilt

(i) $u \otimes (v_1 + v_2) = u \otimes v_1 + u \otimes v_2$,
(ii) $(u_1 + u_2) \otimes v = u_1 \otimes v + u_2 \otimes v$,
(iii) $(\lambda u) \otimes v = \lambda(u \otimes v) = u \otimes (\lambda v)$.

Wegen Bilinearität gilt dann für alle $u \in U$ und $v \in V$

$$u \otimes 0 = 0 = 0 \otimes v.$$

Ist $u = \sum_{i=1}^n x_i a_i$ für eine Basis $\mathscr{A} = (a_1, \ldots, a_n)$ von U und $v = \sum_{j=1}^m y_j b_j$ für eine Basis $\mathscr{B} = (b_1, \ldots, b_m)$ von V, so gilt aufgrund der Bilinearität

$$u \otimes v = \sum_{i=1}^n \sum_{j=1}^m x_i y_j a_i \otimes b_j.$$

Allgemein ist jeder Vektor $w \in U \otimes V$ eindeutig zu schreiben als

$$w = \sum_{i=1}^n \sum_{j=1}^m w_{ij} a_i \otimes b_j$$

mit „Koordinaten“ $w_{ij} \in K$. Benutzt man andere Basen $\mathscr{A}' = (a'_1, \ldots, a'_n)$ von U und $\mathscr{B}' = (b'_1, \ldots, b'_m)$ von V, so gibt es eine weitere Darstellung desselben w als

$$w = \sum_{i=1}^n \sum_{j=1}^m w'_{ij} a'_i \otimes b'_j.$$

Die w'_{ij} berechnen sich aus den w_{ij} mit Hilfe der Basiswechselmatrizen

$$S = (s_{ij}) = \mathrm{M}^{\mathscr{A}}_{\mathscr{A}'}(\mathrm{id}_U) \quad \text{und } T = (t_{ij}) = \mathrm{M}^{\mathscr{B}}_{\mathscr{B}'}(\mathrm{id}_V)$$

durch

$$w'_{ij} = \sum_{k=1}^n \sum_{l=1}^m s_{ik} w_{kl} t_{jl}.$$

Es gilt nämlich

$$a_k = \sum_{i=1}^n s_{ik} a'_i \quad \text{und} \quad b_l = \sum_{j=1}^m t_{jl} b'_j,$$

und damit folgt

$$\begin{aligned} w &= \sum_{k=1}^{n}\sum_{l=1}^{m} w_{kl} a_k \otimes b_l = \sum_{k=1}^{n}\sum_{l=1}^{m} w_{kl} \sum_{i=1}^{n} s_{ik} a_i' \otimes \sum_{j=1}^{m} t_{jl} b_j' \\ &= \sum_{i=1}^{n}\sum_{j=1}^{m} \Big(\sum_{k=1}^{n}\sum_{l=1}^{m} s_{ik} w_{kl} t_{jl}\Big) a_i' \otimes b_j' = \sum_{i=1}^{n}\sum_{j=1}^{m} w_{ij}' a_i' \otimes b_j'. \end{aligned}$$

Korollar 4.9. *Ist $\mathscr{A} = (a_1, \ldots, a_n)$ eine Basis von U und $\mathscr{B} = (b_1, \ldots, b_m)$ sei eine Basis von V, so sind die Vektoren $a_i \otimes b_j \in U \otimes V$ mit $1 \leq i \leq n$, $1 \leq j \leq m$ eine Basis von $U \otimes V$. Insbesondere gilt*

$$\dim_K(U \otimes V) = \dim_K(U) \cdot \dim_K(V).$$

Beweis. Klar. □

Bemerkung 4.10. Nicht jeder Vektor in $U \otimes V$ ist von der Form $u \otimes v$ für ein $u \in U$ und $v \in V$. Wenn $K = \mathbb{F}$ ein endlicher Körper mit q Elementen ist und $\dim_{\mathbb{F}}(U) = n$ sowie $\dim_{\mathbb{F}}(V) = m$, dann vergleichen wir bei $n, m > 2$ und daher $nm > 2\max\{n, m\} \geq n + m$

$$|\{u \otimes v \ ; \ u \in U, \ v \in V\}| \leq q^n \cdot q^m = q^{n+m} < q^{nm} = |U \otimes V|.$$

Genauer gilt: $\sum_{ij} x_{ij} a_i \otimes b_j$ ist von der Form $u \otimes v$ genau dann, wenn die Matrix $(x_{ij}) \in \mathrm{M}_{n\times m}(K)$ Rang ≤ 1 hat. Und das sind nicht alle bei $n, m \geq 2$.

4.2. **Paarungen.**

Definition 4.11. Sei K ein Körper und V, W seien K-Vektorräume. Eine **Paarung** von V mit W ist eine K-bilineare Abbildung $f : V \times W \to K$.

Notation 4.12. Manchmal ziehen wir der Notation $f : V \times W \to K$ eine Notation mit Klammern vor. Wir schreiben dann

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

Beispiel 4.13. (1) Das **Standardskalarprodukt** auf K^n ist die Bilinearform $K^n \times K^n \to K$ definiert durch

$$\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \rangle := \sum_{i=1}^{n} x_i y_i.$$

Der Nachweis der Bilinearität ist eine einfache Übung im Distributivgesetz von K und gelingt am ökonomischsten mit der Schreibweise

$$\langle x, y\rangle = x^t \cdot y,$$

welche das Standardskalarprodukt durch Matrixmultiplikation mit dem transponierten Vektor beschreibt:

$$x^t = (x_1, \ldots, x_n).$$

(2) Die Spur quadratischer $n \times n$-Matrizen ist eine K-lineare Abbildung

$$\mathrm{Sp} : \mathrm{M}_n(K) \to K,$$

nämlich die Summe der Einträge auf der Diagonalen: für die Matrix $A = (a_{ij})$ ist

$$\mathrm{Sp}(A) = \sum_{i=1}^{n} a_{ii}.$$

Seien $V = \mathrm{M}_{n\times m}(K)$ und $W = \mathrm{M}_{m\times n}(K)$ Vektorräume von Matrizen mit Einträgen im Körper K. Für $A \in V$ und $B \in W$ ist $AB \in \mathrm{M}_n(K)$ eine quadratische Matrix und hat demnach eine Spur. Dann definiert

$$(A, B) \mapsto \mathrm{Sp}(AB)$$

eine Paarung $V \times W \to K$. Wir zeigen exemplarisch für $A_1, A_2 \in V$
$$\mathrm{Sp}((A_1 + A_2)B) = \mathrm{Sp}(A_1B + A_2B) = \mathrm{Sp}(A_1B) + \mathrm{Sp}(A_2B).$$

(3) Das tautologische Beispiel einer Paarung ist die **Auswertung**
$$\begin{aligned} V^* \times V &\to K \\ (f, v) &\mapsto f(v), \end{aligned}$$
Linearität in $v \in V$ ist äquivalent zur Linearität von f. Linearität in $f \in V^*$ ist gerade die Definition der K-Vektorraumstruktur auf dem Dualraum $V^* = \mathrm{Hom}_K(V, K)$.

(4) Sei $f(x) \in C_c(\mathbb{R})$ eine stetige Funktion mit kompaktem Träger, d.h. für ein geeignetes von f abhängendes $R > 0$ ist $f(x) = 0$ für alle $|x| > R$. Sei $g(x) \in C(\mathbb{R})$ eine stetige Funktion auf $\mathbb{R}$. Dann ist das Integral
$$(f, g) := \int_{-\infty}^{\infty} f(x)g(x)dx \quad \in \mathbb{R}$$
wohldefiniert, und als Funktion von $f(x)$ und $g(x)$ eine Paarung $C_c(\mathbb{R}) \times C(\mathbb{R}) \to \mathbb{R}$.

4.3. **Matrixbeschreibung.** Wir betrachten ein weiteres Beispiel.

Beispiel 4.14. Mit $A = (a_{ij})_{1 \le i \le n, 1 \le j \le m} \in \mathrm{M}_{n \times m}(K)$ definieren wir eine Paarung durch
$$\begin{aligned} \langle\ ,\ \rangle_A : K^n \times K^m &\to K \\ (x, y) &\mapsto \langle x, y\rangle_A := \langle x, Ay\rangle = x^t Ay = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i y_j a_{ij}, \end{aligned}$$
wobei $x = (x_1, \ldots, x_n)^t$ und $y = (y_1, \ldots, y_m)^t$ die Koordinaten von x und y sind. Wir rechnen exemplarisch für $y, y' \in K^m$ und $\lambda, \mu \in K$:
$$\langle x, \lambda y + \mu y'\rangle_A = x^t A(\lambda y + \mu y') = \lambda x^t Ay + \mu x^t Ay' = \lambda\langle x, y\rangle_A + \mu\langle x, y'\rangle_A.$$

Wir wollen nun einsehen, daß dies kein Beispiel, sondern die Beschreibung einer jeden Paarung zwischen endlichdimensionalen K-Vektorräumen liefert.

Definition 4.15. Die **Gram'sche Matrix** einer Paarung von endlichdimensionalen K-Vektorräumen $f : V \times W \to K$ bezüglich Basen $\mathscr{B} = (b_1, \ldots, b_n)$ von V und $\mathscr{C} = (c_1, \ldots, c_m)$ von W ist die $n \times m$-Matrix
$$\mathrm{M}^{\mathscr{B},\mathscr{C}}(f) = (f(b_i, c_j))_{1 \le i \le n, 1 \le j \le m} \in \mathrm{M}_{n \times m}(K).$$

Notation 4.16. Wir verwenden die Notation aus Definition 4.15 weiter und erinnern an den Koordinatenisomorphismus
$$\kappa_{\mathscr{B}} : V \xrightarrow{\sim} K^n,$$
der einen Vektor $v \in V$, der in der Basis $\mathscr{B}$ als $v = \sum_{i=1}^n x_i b_i$ geschrieben werden kann, auf den Spaltenvektor seiner Koordinaten bezüglich $\mathscr{B}$ abbildet:
$$\kappa_{\mathscr{B}}(v) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$
Es gilt insbesondere $\kappa_{\mathscr{B}}(b_i) = e_i$, wobei e_i der i-te Standardbasisvektor von K^n ist, der außer einer 1 im i-ten Eintrag sonst nur den Eintrag 0 hat.

Beispiel 4.17. Sei $A = (a_{ij}) \in \mathrm{M}_{n \times m}(K)$ und $\langle\ ,\ \rangle_A : K^n \times K^m \to K$ die zugehörige Paarung aus Beispiel 4.14. Die Gram'sche Matrix zu $\langle\ ,\ \rangle_A$ bezüglich der Standardbasis[4] $\mathscr{E} = \{e_1, \ldots, e_i, \ldots\}$ hat als ij-ten Eintrag
$$(e_i, e_j)_A = e_i^t A e_j = a_{ij},$$

[4] Vorsicht: mißbräuchlich gleiche Notation für die Standardbasis von K^n und für K^m.

also

$$\mathrm{M}^{\mathscr{E},\mathscr{E}}(\langle\ ,\ \rangle_A) = A.$$

Proposition 4.18. *Sei $f : V \times W \to K$ eine Paarung endlichdimensionaler K-Vektorräume V und W mit Basen $\mathscr{B} = (b_1, \ldots, b_n)$ von V und $\mathscr{C} = (c_1, \ldots, c_m)$ von W. Sei $A = \mathrm{M}^{\mathscr{B},\mathscr{C}}(f)$ die Gram'sche Matrix der Paarung. Dann ist das Diagramm*

$$\begin{tikzcd} V \times W \arrow[r, "f"] \arrow[d, "\kappa_{\mathscr{B}} \times \kappa_{\mathscr{C}}"'] & K \arrow[d, equal] \\ K^n \times K^m \arrow[r, "\langle\ ,\ \rangle_A"] & K. \end{tikzcd}$$

kommutativ, d.h. für alle $v \in V$ und $w \in W$ mit Koordinaten $\kappa_{\mathscr{B}}(v) = x$ und $\kappa_{\mathscr{C}}(w) = y$ gilt

$$f(v, w) = x^t A y. \tag{4.1}$$

In Koordinaten für V und W wird die Paarung f durch die von der Gram'schen Matrix zu f definierten Paarung beschrieben.

Beweis. Sei $A = (a_{ij}) \in \mathrm{M}_{n\times m}(K)$ die Gram'sche Matrix von f. Wir müssen für $v = \sum_{i=1}^n x_i b_i$ und $w = \sum_{j=1}^m y_j c_j$ die Gleichung (4.1) nachweisen:

$$\begin{aligned} f(v,w) &= f(\sum_{i=1}^n x_i b_i, \sum_{j=1}^m y_j c_j) = \sum_{i=1}^n x_i f(b_i, \sum_{j=1}^m y_j c_j) = \sum_{i=1}^n \sum_{j=1}^m x_i y_j f(b_i, c_j) \\ &= \sum_{i=1}^n \sum_{j=1}^m x_i a_{ij} y_j = \sum_{i=1}^n x_i \big(A \kappa_{\mathscr{C}}(w)\big)_i = \kappa_{\mathscr{B}}(v)^t A \kappa_{\mathscr{C}}(w). \end{aligned}$$

Alternative: Die Paarung f induziert ein K-lineares $F : V \otimes W \to K$ und $\langle\ ,\ \rangle_A$ ein K-lineares $\alpha : K^n \otimes K^m \to K$. Die Abbildung $V \times W \to K^n \otimes K^m$ mit $(v, w) \mapsto \kappa_{\mathscr{B}}(v) \otimes \kappa_{\mathscr{C}}(w)$ ist bilinear und führt aufgrund der Eigenschaft des Tensorprodukts zu einer linearen Abbildung $\kappa_{\mathscr{B}\otimes\mathscr{C}}$ wie im Diagramm:

$$\begin{tikzcd} V \otimes W \arrow[r, "F"] \arrow[d, "\kappa_{\mathscr{B}\otimes\mathscr{C}}"'] & K \arrow[d, equal] \\ K^n \otimes K^m \arrow[r, "\alpha"] & K. \end{tikzcd}$$

Die Behauptung der Proposition ist äquivalent dazu, daß dieses Diagramm kommutiert:

$$F = \alpha \circ \kappa_{\mathscr{B}\otimes\mathscr{C}}.$$

Lineare Abbildungen sind durch den Wert auf einer Basis eindeutig bestimmt. Als Basis von $V \otimes W$ wählen wir $b_i \otimes c_j$ mit $1 \le i \le n$, $1 \le j \le m$. Dann haben wir zu vergleichen

$$F(b_i \otimes c_j) = f(b_i, c_j) = a_{ij} = e_i^t A e_j = \alpha(e_i \otimes e_j) = \alpha\big(\kappa_{\mathscr{B}}(b_i) \otimes \kappa_{\mathscr{C}}(c_j)\big) = \alpha(\kappa_{\mathscr{B}\otimes\mathscr{C}}(b_i \otimes c_j)). \qquad \square$$

Die Gram'sche Matrix einer Paarung ändert sich, wenn man die Basen wechselt, durch eine Formel, welche die zugehörigen Basiswechselmatrizen benötigt.

Proposition 4.19. *Sei $f : V \times W \to K$ eine Paarung endlichdimensionaler K-Vektorräume V und W mit Basen $\mathscr{B} = (b_1, \ldots, b_n)$ von V und $\mathscr{C} = (c_1, \ldots, c_m)$ von W. Sei $A = \mathrm{M}^{\mathscr{B},\mathscr{C}}(f)$ die Gram'sche Matrix.*

Sei $\mathscr{B}' = (b_1', \ldots, b_n')$ eine weitere Basis von V, sei $\mathscr{C}' = (c_1', \ldots, c_m')$ eine weitere Basis von W, und sei $A' = \mathrm{M}^{\mathscr{B}',\mathscr{C}'}(f)$ die Gram'sche Matrix. Dann gilt

$$A' = S^t A T$$

mit den Basiswechselmatrizen

$$S = \mathrm{M}^{\mathscr{B}'}_{\mathscr{B}}(\mathrm{id}_V) \quad \textit{und} \quad T = \mathrm{M}^{\mathscr{C}'}_{\mathscr{C}}(\mathrm{id}_W).$$

Beweis. Für $v \in V$ und $w \in W$ seien $x = \kappa_{\mathscr{B}}(v)$ und $y = \kappa_{\mathscr{C}}(w)$ sowie $x' = \kappa_{\mathscr{B}'}(v)$ und $y' = \kappa_{\mathscr{C}'}(w)$ die Koordinatenvektoren. Dann gilt

$$x = Sx' \quad \text{und} \quad y = Ty'.$$

Nach Proposition 4.18 folgt

$$x'^t A' y' = f(v,w) = x^t A y = (Sx')^t A (Ty') = x'^t (S^t A T) y'.$$

Dies kann man als Gleichheit zweier Paarungen $K^n \times K^m \to K$ betrachten. Daraus folgt die Behauptung $A' = S^t A T$ durch Vergleich der Gram'schen Matrizen bezüglich der Standardbasis von K^n und K^m, siehe Beispiel 4.17. Alternativ wertet man auf $x' = e_i$ und $y' = e_j$ aus, d.h. $v = b'_i$ und $w = c'_j$, und erhält die ij-Einträge der Matrizen. □

Lemma–Definition 4.20. *Seien V, W zwei K-Vektorräume. Die Menge der Paarungen von V mit W bildet einen K-Vektorraum*

$$\mathscr{L}(V, W; K)$$

unter punktweiser Addition und Skalarmultiplikation.

Beweis. Punktweise bedeutet, daß für $\lambda \in K$ und $f, f_1, f_2 \in \mathscr{L}(V, W; K)$ Addition und Skalarmultiplikation von Paarungen als die Funktionen auf $v \in V$ und $w \in W$ wie folgt definiert sind:

$$\big(f_1 + f_2\big)(v,w) := f_1(v,w) + f_2(v,w)$$

und

$$\big(\lambda f\big)(v,w) := \lambda \cdot f(v,w).$$

Daß diese Formeln Paarungen, also Elemente in $\mathscr{L}(V, W; K)$, definieren, ist genauso ausschließlich Fleißarbeit wie der Nachweis, daß damit eine K-Vektorraumstruktur definiert wird. □

Satz 4.21. *Seien V, W zwei K-Vektorräume. Sei $\mathscr{B} = (b_1, \ldots, b_n)$ eine Basis von V und $\mathscr{C} = (c_1, \ldots, c_m)$ eine Basis von W. Dann haben wir Isomorphismen von K-Vektorräumen*

$$\mathrm{Hom}_K(V \otimes W, K) \xleftarrow{\sim} \mathscr{L}(V, W; K) \xrightarrow{\sim} \mathrm{M}_{n \times m}(K),$$

die einer Paarung $f : V \times W \to K$ das Folgende zuordnen:

- *nach links: die zugehörige K-lineare Abbildung $F : V \otimes W \to K$,*
- *nach rechts: ihre Gram'sche Matrix $\mathrm{M}^{\mathscr{B},\mathscr{C}}(f)$.*

Beweis. Die Abbildung nach links $\mathscr{L}(V, W; K) \to \mathrm{Hom}_K(V \otimes W, K)$ ist bijektiv nach der universellen Eigenschaft des Tensorprodukts. Die Definition der K-Vektorraumstrukturen ist in beiden Fällen ‚punktweise', so daß die Abbildung K-linear ist.

Nach rechts: Eine Abbildung nach $\mathrm{M}_{n \times m}(K)$ ist linear, wenn der ij-te Matrixeintrag linear ist für alle $i = 1, \ldots, n$ und $j = 1, \ldots, m$. Dies sind hier die Abbildungen

$$\mathscr{L}(V, W; K) \to K$$
$$f \mapsto f(b_i, c_j)$$

und diese sind per Definition der K-Vektorraumstruktur auf $\mathscr{L}(V, W; K)$ linear.

Die Gleichung (4.1) zeigt, daß man aus der Gram'schen Matrix die Paarung berechnen kann. Daher ist die Abbildung injektiv.

Es fehlt nunmehr nur noch die Surjektivität. Sei $A = (a_{ij}) \in \mathrm{M}_{n \times m}(K)$ beliebig. Dann definiert

$$v, w \mapsto \kappa_{\mathscr{B}}(v)^t A \kappa_{\mathscr{C}}(w)$$

für $v \in V$ und $w \in W$ eine Paarung $f_A : V \times W \to K$. Die Gram'sche Matrix zu f_A bezüglich $\mathscr{B}$ und $\mathscr{C}$ hat den ij-ten Eintrag

$$f_A(b_i, c_j) = \kappa_{\mathscr{B}}(b_i)^t A \kappa_{\mathscr{C}}(c_j) = e_i^t A e_j = a_{ij}.$$

Also ist A die Gram'sche Matrix von f_A, und das zeigt die Surjektivität (vgl. Beispiel 4.17). $\square$

4.4. **Paarungen und der Dualraum.** Die **duale Basis** zu einer Basis $\mathscr{B} = (b_1, \ldots, b_n)$ eines endlichdimensionalen K-Vektorraums V besteht aus dem Tupel $\mathscr{B}^* = (b_1^*, \ldots, b_n^*)$ von Elementen von V^* mit der Eigenschaft

$$b_i^*(b_j) = \begin{cases} 1 & \text{wenn } i = j, \\ 0 & \text{wenn } i \neq j \end{cases} = \delta_{ij}$$

mit dem Kronecker-δ. Die b_i^* sind dadurch eindeutig bestimmt und $\mathscr{B}^*$ ist eine Basis von V^*. Die Koordinatendarstellung von $f \in V^*$ bezüglich $\mathscr{B}^*$ ist $\kappa_{\mathscr{B}^*}(f) = (f(b_1), \ldots, f(b_n))^t$, weil

$$f = \sum_{i=1}^n f(b_i) b_i^*, \tag{4.2}$$

denn beide Seiten nehmen auf der Basis $\mathscr{B}$ dieselben Werte an:

$$\Big(\sum_{i=1}^n f(b_i) b_i^*\Big)(b_j) = \sum_{i=1}^n f(b_i) b_i^*(b_j) = f(b_j).$$

Mittels des Dualraums läßt sich die Bilinearität einer Paarung alternativ wie folgt beschreiben.

Definition 4.22. Zu einer Paarung von K-Vektorräumen

$$f : V \times W \to K$$

gehören die **linkspartielle Auswertung**

$$\ell = \ell_f : V \to W^*, \quad \ell(v) = (w \mapsto f(v, w))$$

und die **rechtspartielle Auswertung**

$$r = r_f : W \to V^*, \quad r(w) = (v \mapsto f(v, w)).$$

Bemerkung 4.23. (1) Die Abbildung $\ell(v)$ ist linear in w, da f linear im zweiten Argument ist. Die Zuordnung $v \mapsto \ell(v)$ ist linear per Definition der K-Vektorraumstruktur des Dualraums W^* vermöge punktweiser Addition und Skalarmultiplikation, weil f linear im ersten Argument ist. Der Fall $W \to V^*$ ist analog.

(2) Wir nennen hier $\ell = \ell_f$ die **linkspartielle Auswertung** und $r = r_f$ die **rechtspartielle Auswertung** von f. Offizielle Namen haben ℓ und r nicht.

Proposition 4.24. *Sei $f : V \times W \to K$ eine Paarung endlichdimensionaler K-Vektorräume, und sei $\mathscr{B} = (b_1, \ldots, b_n)$ eine Basis von V und $\mathscr{C} = (c_1, \ldots, c_m)$ eine Basis von W. Sei $A = \mathrm{M}^{\mathscr{B},\mathscr{C}}(f) \in \mathrm{M}_{n \times m}(K)$ die Gram'sche Matrix.*

(1) Die rechtspartielle Auswertung $r : W \to V^*$ wird beschrieben durch die Matrix

$$\mathrm{M}^{\mathscr{C}}_{\mathscr{B}^*}(r) = A \in \mathrm{M}_{n \times m}(K).$$

(2) Die linkspartielle Auswertung $\ell : V \to W^*$ wird beschrieben durch die Matrix

$$\mathrm{M}^{\mathscr{B}}_{\mathscr{C}^*}(\ell) = A^t \in \mathrm{M}_{m \times n}(K).$$

Beweis. (1) Für die j-te Spalte von $\mathrm{M}^{\mathscr{C}}_{\mathscr{B}^*}(r)$ müssen wir $r(c_j)$ in der Basis $\mathscr{B}^*$ ausdrücken. Aus (4.2) folgt

$$r(c_j) = \sum_{i=1}^n r(c_j)(b_i) b_i^* = \sum_{i=1}^n f(b_i, c_j) b_i^*$$

und damit wie behauptet
$$\mathrm{M}^{\mathscr{C}}_{\mathscr{B}^*}(r) = A \in \mathrm{M}_{n\times m}(K).$$

(2) Für die j-te Spalte von $\mathrm{M}^{\mathscr{B}}_{\mathscr{C}^*}(\ell)$ müssen wir $\ell(b_j)$ in der Basis $\mathscr{C}^*$ ausdrücken. Aus (4.2) folgt
$$\ell(b_j) = \sum_{i=1}^{m} \ell(b_j)(c_i)c_i^* = \sum_{i=1}^{m} f(b_j, c_i)c_i^*$$
und damit wie behauptet
$$\mathrm{M}^{\mathscr{B}}_{\mathscr{C}^*}(\ell) = A^t \in \mathrm{M}_{m\times n}(K).$$ □

Übungsaufgaben zu §4

Übungsaufgabe 4.1. Seien $V = \mathrm{M}_{n\times m}(K)$ und $W = \mathrm{M}_{m\times n}(K)$ Vektorräume von Matrizen mit Einträgen im Körper K. Für $A \in V$ und $B \in W$ definieren
$$\langle A, B\rangle_1 := \mathrm{Sp}(AB)$$
und
$$\langle A, B\rangle_2 := \mathrm{Sp}(BA)$$
Paarungen $V \times W \to K$. Zeigen Sie, daß $\langle\ ,\ \rangle_1 = \langle\ ,\ \rangle_2$ gilt.

Übungsaufgabe 4.2. Sei V ein K-Vektorraum mit Basis $\mathscr{B} = (b_1, \ldots, b_n)$ und dazu dualer Basis $\mathscr{B}^* = (b_1^*, \ldots, b_n^*)$. Zeigen Sie, daß sich der Koordinatenisomorphismus $\kappa_{\mathscr{B}} : V \xrightarrow{\sim} K^n$ als
$$\kappa_{\mathscr{B}}(v) = \begin{pmatrix} b_1^*(v) \\ \vdots \\ b_n^*(v) \end{pmatrix}$$
schreiben läßt. Es gilt also insbesondere
$$v = \sum_{i=1}^{n} b_i^*(v) b_i. \tag{4.3}$$

5. Perfekte Paarungen

5.1. **Nichtausgeartete und perfekte Paarungen.** Nicht alle Paarungen sind gleich ‚gut', das Extrembeispiel ist sicher die Nullpaarung, deren Wert konstant 0 ist. Auf der anderen Seite, und viel nützlicher, befinden sich die nichtausgearteten bzw. perfekten Paarungen.

Definition 5.1. Sei $f : V \times W \to K$ eine Paarung von K-Vektorräumen.

(1) Die Paarung f ist **links-(bzw. rechts-)nichtausgeartet**, wenn es für alle $v \in V$, $v \neq 0$ ein $w \in W$ gibt (bzw. für alle $w \in W$, $w \neq 0$ ein $v \in V$ gibt) mit
$$f(v, w) \neq 0.$$
(2) Die Paarung f heißt **nichtausgeartet**, wenn sie links- und rechts-nichtausgeartet ist. Andernfalls heißt f **ausgeartet**[5]
(3) Die Paarung f ist eine **perfekte Paarung**, wenn die partiellen Auswertungen
$$\ell : V \to W^* \quad \text{und} \quad r : W \to V^*$$
Isomorphismen von K-Vektorräumen sind.

Die Eigenschaft nichtausgeartet wird häufig auf die folgende Art und Weise benutzt.

[5]„Ausgeartet" ist also dasselbe wie „nicht nichtausgeartet".

Lemma 5.2. *Sei $f : V \times W \to K$ eine links-nichtausgeartete Paarung und seien $v_1, v_2 \in V$. Wenn für alle $w \in W$ gilt*

$$f(v_1, w) = f(v_2, w),$$

dann gilt $v_1 = v_2$.

Analog gilt dies bei rechts-nichtausgeartet mit vertauschten Rollen von V und W.

Beweis. Laut Voraussetzung ist für alle $w \in W$

$$f(v_1 - v_2, w) = f(v_1, w) - f(v_2, w) = 0$$

und damit per Definition von links-nichtausgeartet bereits $v_1 - v_2 = 0$. □

Definition 5.3. Sei $f : V \times W \to K$ eine Paarung von K-Vektorräumen.

(1) Der **Linkskern** von f ist der Untervektorraum von V

$$W^\perp = \{v \in V \ ; \ f(v, w) = 0 \text{ für alle } w \in W\} = \ker(\ell_f : V \to W^*).$$

(2) Der **Rechtskern** von f ist der Untervektorraum von W

$$V^\perp = \{w \in W \ ; \ f(v, w) = 0 \text{ für alle } v \in V\} = \ker(r_f : W \to V^*).$$

Proposition 5.4. *Sei $f : V \times W \to K$ eine Paarung von K-Vektorräumen. Dann sind jeweils äquivalent:*

(1) *(a)* *Die linkspartielle Auswertung $\ell : V \to W^*$ ist injektiv.*
(b) $W^\perp = 0$.
(c) *f ist links-nichtausgeartet.*

(2) *(a)* *Die rechtspartielle Auswertung $r : W \to V^*$ ist injektiv.*
(b) $V^\perp = 0$.
(c) *f ist rechts-nichtausgeartet.*

Beweis. Die Äquivalenz der Aussagen sieht man wie folgt. (a) $\iff$ (b) ist eine Eigenschaft des Kerns einer linearen Abbildung. Und (b) $\iff$ (c) folgt unmittelbar aus der Definition von links-nichtausgeartet und $W^\perp$ (bzw. rechts-nichtausgeartet und $V^\perp$). □

Satz 5.5. *Sei $f : V \times W \to K$ eine Paarung von endlichdimensionalen K-Vektorräumen. Es sind äquivalent:*

(a) *f ist perfekt.*
(b) *f ist nichtausgeartet.*
(c) *f ist links-nichtausgeartet und* $\dim_K(V) = \dim_K(W)$.
(d) *f ist rechts-nichtausgeartet und* $\dim_K(V) = \dim_K(W)$.
(e) *Zu jeder Wahl von Basen $\mathscr{B}$ von V und $\mathscr{C}$ von W ist die Gram'sche Matrix von f quadratisch und hat Determinante $\neq 0$.*
(f) *Zu einer Wahl von Basen $\mathscr{B}$ von V und $\mathscr{C}$ von W ist die Gram'sche Matrix von f quadratisch und hat Determinante $\neq 0$.*

Beweis. Wir zeigen (a) $\Longrightarrow$ (b) $\Longrightarrow$ ((c) und (d)) $\Longrightarrow$ ((c) oder (d)) $\Longrightarrow$ (e) $\Longrightarrow$ (f) $\Longrightarrow$ (a).

(a) $\Longrightarrow$ (b): Ist f perfekt, so sind die partiellen Auswertungsabbildungen ℓ und r Isomorphismen, also insbesondere injektiv. Nach Proposition 5.4 ist f nichtausgeartet.

(b) $\Longrightarrow$ ((c) und (d)): Sei f nichtausgeartet. Dann sind $V \to W^*$ und $W \to V^*$ injektiv und somit

$$\dim_K(V) \le \dim_K(W^*) = \dim_K(W) \le \dim_K(V^*) = \dim_K(V),$$

so daß (c) und (d) gelten. Der nächste Schritt nach ((c) oder (d)) ist trivial.

(c) oder (d) $\Longrightarrow$ (e): Angenommen, Aussage (c) gilt. Dann ist

$$\ell : V \to W^*$$

eine injektive lineare Abbildung zwischen endlichdimensionalen Vektorräumen der gleichen Dimension, also ein Isomorphismus. Sei $A = \mathrm{M}^{\mathscr{B},\mathscr{C}}(f)$ die Gram'sche Matrix. Die Matrix von ℓ bezüglich $\mathscr{B}$ und $\mathscr{C}^*$ ist die transponierte A^t, siehe Proposition 4.24. Dann:

$$\ell \text{ Isomorphismus } \iff \det(A^t) \neq 0.$$

Da $\det(A) = \det(A^t)$, folgt Aussage (e). Mit Aussage (d) ist die Argumentation analog.

(e) $\Longrightarrow$ (f): Das ist trivial.

(f) $\Longrightarrow$ (a): Wenn die Gram'sche Matrix $A = \mathrm{M}^{\mathscr{B},\mathscr{C}}(f)$ quadratisch ist und $\det(A) \neq 0$ gilt, dann sind lineare Abbildungen, die in Koordinaten durch Multiplikation mit A oder A^t dargestellt werden, Isomorphismen. Dies trifft nach Proposition 4.24 auf die partiellen Auswertungen ℓ und r zur Paarung f zu. Also ist f eine perfekte Paarung. ☐

Korollar 5.6 (Riesz'scher Darstellungssatz)**.** *Sei $f : V \times W \to K$ eine nichtausgeartete Paarung endlichdimensionaler Vektorräume. Dann ist jede Linearform*

$$\pi : W \to K$$

für ein eindeutiges $v \in V$ von der Form

$$\pi = \ell(v) = f(v, -).$$

Beweis. Nach Satz 5.5 ist f sogar eine perfekte Paarung. Demnach ist $V \to W^*$ ein Isomorphismus und dies beinhaltet genau die Aussage des Korollars. ☐

Korollar 5.7. *Ist $f : V \times W \to K$ eine nichtausgeartete Paarung von K-Vektorräumen endlicher Dimension, dann gilt*

$$\dim_K(V) = \dim_K(W).$$

Beweis. Nach Satz 5.5 (c). ☐

Beispiel 5.8. Das Standardskalarprodukt $\langle\ ,\ \rangle : K^n \times K^n \to K$ hat als Gram'sche Matrix bezüglich der Standardbasis die Einheitsmatrix. Somit definiert das Standardskalarprodukt eine perfekte Paarung. Insbesondere ist jede Linearform auf K^n von der Form

$$\langle v, - \rangle$$

für ein eindeutiges $v \in K^n$.

Bemerkung 5.9. Im Endlichdimensionalen ist der Riesz'sche Darstellungssatz in Form des Korollars 5.6 kein schwieriger Satz. Dies ändert sich, wenn man zu unendlichdimensionalen topologischen Vektorräumen und stetigen Linearformen übergeht.

Beispiel 5.10. Für einen K-Vektorraum V endlicher Dimension ist die Auswertungspaarung

$$V^* \times V \to K$$

eine perfekte Paarung. Die zur Auswertungspaarung gehörenden partiellen Auswertungen ℓ und r sind die Identität $V^* \to V^*$ und die Abbildung $\iota : V \to (V^*)^*$, die $v \in V$ auf

$$\iota(v) : V^* \to K, \qquad \iota(v)(\pi) = \pi(v)$$

abbildet. Man kann dies aus Satz 5.5 folgern: weil $\ell : V^* \to V^*$ die Identität ist und $\dim(V^*) = \dim(V)$, so ist (c) erfüllt, damit die Paarung perfekt und die Abbildung ι ein Isomorphismus.

5.2. **Symmetrische Bilinearformen.** Wir spezialisieren nun Paarungen zu Bilinearformen.

Definition 5.11. Eine **Bilinearform** auf einem K-Vektorraum V ist eine Paarung
$$f : V \times V \to K.$$

Die Bilinearform f heißt

(1) **symmetrisch**, wenn für alle $v, w \in V$ gilt:
$$f(v,w) = f(w,v),$$
(2) **antisymmetrisch** (oder **schiefsymmetrisch**), wenn für alle $v, w \in V$ gilt:
$$f(v,w) = -f(w,v),$$
(3) **alternierend**, wenn für alle $v \in V$ gilt
$$f(v,v) = 0.$$

Im Fall einer symmetrischen Bilinearform f nennt man $V^{\perp}$ (Linkskern ist gleich Rechtskern) auch das **Radikal** von f und verwendet die Notation $\mathrm{Rad}(f) = V^{\perp}$.

Definition 5.12. Eine quadratische Matrix $A \in \mathrm{M}_n(K)$ heißt

(1) **symmetrisch**, wenn $A^t = A$,
(2) **antisymmetrisch** (oder **schiefsymmetrisch**), wenn $A^t = -A$,
(3) **alternierend**, wenn $A^t = -A$ und die Diagonaleinträge von A gleich 0 sind.

Beispiel 5.13. (1) Das Standardskalarprodukt $\langle x, y\rangle = x^t y$ auf K^n ist symmetrisch.

(2) Sei $C^1(\mathbb{R})$ der $\mathbb{R}$-Vektorraum der einmal stetig differenzierbaren reellwertigen Funktionen auf $\mathbb{R}$, und sei $V \subseteq C(\mathbb{R})$ der Unterraum derjenigen Funktionen $f : \mathbb{R} \to \mathbb{R}$ mit $f(0) = f(1)$. Auf V definiert für $f, g \in V$
$$(f,g) := \int_0^1 f(x)g'(x)dx$$
eine antisymmetrische und alternierende Bilinearform. Wegen
$$(f,g) + (g,f) = \int_0^1 f(x)g'(x) + g(x)f'(x)dx = \int_0^1 (fg)'(x)dx = fg(1) - fg(0) = 0$$
liegt Antisymmetrie vor. Alternierend zu sein folgt aus der Kettenregel
$$(f,f) = \int_0^1 f(x)f'(x)dx = \frac{1}{2}f^2(1) - \frac{1}{2}f^2(0) = 0.$$

(3) Auf $V = K^2$ definiert die Determinante eine alternierende Bilinearform
$$f(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}) := \det \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}.$$
Dies folgt aus den Eigenschaften der Determinantenfunktion.

Bemerkung 5.14. Sei A antisymmetrisch und sei a_{ii} der i-te Diagonaleintrag von A. Dann ist $2a_{ii} = 0$ als der i-te Diagonaleintrag von $A^t + A = 0$. Wenn $2 \in K^{\times}$, so folgt $a_{ii} = 0$ und A ist sogar alternierend.

Je nach Anwendungsgebiet ist die Voraussetzung $2 \in K^{\times}$ oft, oder, wenn Sie zum Beispiel ein Computer sind, selten erfüllt.

Proposition 5.15. *Sei V ein K-Vektorraum mit Basis $\mathscr{B}$ und Bilinearform $f : V \times V \to K$. Sei $\mathscr{B}$ eine Basis von V und $A = \mathrm{M}^{\mathscr{B},\mathscr{B}}(f)$ die Gram'sche Matrix. Dann gilt:*

(1) f ist symmetrisch $\iff$ $A^t = A$ ist symmetrisch.
(2) f ist antisymmetrisch $\iff$ $A^t = -A$ ist antisymmetrisch.

Beweis. Wir benutzen das Vorzeichen $\varepsilon = 1$ für den Fall „symmetrisch", und das Vorzeichen $\varepsilon = -1$ für den Fall „antisymmetrisch".

Ist f (anti-)symmetrisch, so gilt

$$f(b_i, b_j) = \varepsilon \cdot f(b_j, b_i) \qquad \text{für alle } 1 \leq i, j \leq n, \tag{5.1}$$

und das ist äquivalent dazu, daß $\mathrm{M}^{\mathscr{B},\mathscr{B}}(f)$ (anti-)symmetrisch ist.

Für die Umkehrung müssen wir aus (5.1) folgern, daß f (anti-)symmetrisch ist. Seien $v = \sum_i x_i b_i$ und $w = \sum_i y_i b_i$ beliebig. Dann gilt

$$\begin{aligned} f(v,w) = f(\sum_i x_i b_i, \sum_j y_j b_j) &= \sum_{i,j} x_i y_j f(b_i, b_j) \\ &= \varepsilon \cdot \sum_{i,j} x_i y_j f(b_j, b_i) = \varepsilon \cdot f(\sum_j y_j b_j, \sum_i x_i b_i) = \varepsilon \cdot f(w,v). \square \end{aligned}$$

Wir zeigen nun eine Variante der binomischen Formeln.

Satz 5.16. *Sei f eine symmetrische Bilinearform auf dem K-Vektorraum V. Dann gilt für alle $v, w \in V$ die* ***Polarisationsformel****:*

$$2f(v,w) = f(v+w, v+w) - f(v,v) - f(w,w)$$

Beweis. Wegen Bilinearität gilt

$$f(v+w, v+w) = f(v, v+w) + f(w, v+w) = f(v,v) + f(v,w) + f(w,v) + f(w,w),$$

und die Symmetrie zeigt $f(v,w) + f(w,v) = 2f(v,w)$. Hieraus folgt die Polarisationsformel. $\square$

Bemerkung 5.17. Die Polarisationsformel zeigt, daß eine symmetrische Bilinearform f durch die Werte $f(v,v)$ für alle $v \in V$ eindeutig bestimmt ist, sofern $2 \in K^\times$ gilt.

Zum Schluß dieses Abschnitts diskutieren wir die Begriffe antisymmetrisch und alternierend.

Proposition 5.18. *Sei V ein K-Vektorraum mit Basis $\mathscr{B}$ und Bilinearform $f : V \times V \to K$.*

(1) f alternierend $\Longrightarrow$ f antisymmetrisch.
(2) Wenn $2 \in K^\times$, dann gilt auch: f antisymmetrisch $\Longrightarrow$ f alternierend.
(3) f ist alternierend $\Longleftrightarrow$ die Gram'sche Matrix von f bezüglich $\mathscr{B}$ ist alternierend.

Beweis. (1) Die Rechnung in der Polarisationsformel benutzt bis kurz vor Schluß nicht die Symmetrie. Es gilt immer

$$f(v+w, v+w) = f(v, v+w) + f(w, v+w) = f(v,v) + f(v,w) + f(w,v) + f(w,w).$$

Ist f alternierend, so bleibt davon nur

$$0 = f(v,w) + f(w,v),$$

und das zeigt die Antisymmetrie von f.

(2) Sei f antisymmetrisch. Dann ist für alle $v \in V$

$$f(v,v) = -f(v,v),$$

woraus $2f(v,v) = 0$ folgt. Wenn $2 \in K^\times$, so kann man mit $1/2$ multiplizieren und f ist alternierend.

(3) Sei f alternierend. Dann ist f auch antisymmetrisch nach (1) und nach Proposition 5.15 die Gram'sche Matrix $A = \mathrm{M}^{\mathscr{B},\mathscr{B}}(f)$ antisymmetrisch. Die Diagonaleinträge sind $f(b_i, b_i) = 0$, wenn $b_i \in \mathscr{B}$ die Basisvektoren sind. Folglich ist A alternierend.

Sei nun $A = (a_{ij}) = \mathrm{M}^{\mathscr{B},\mathscr{B}}(f)$ alternierend und sei $v = \sum_{i=1}^n x_i b_i \in V$ ein beliebiger Vektor. Dann gilt mit $x = (x_1, \ldots, x_n)^t$

$$f(v,v) = x^t A x = \sum_{i=1}^n \sum_{j=1}^n x_i x_j a_{ij} = \sum_{i<j} x_i x_j (a_{ij} + a_{ji}) + \sum_{i=1}^n x_i^2 a_{ii} = 0. \qquad \square$$

5.3. **Duale Basis — ein zweites Mal.** In Gegenwart einer perfekten symmetrischen Bilinearform gibt es nicht nur im Dualraum eine duale Basis. Der Notationsmißbrauch kann verkraftet werden, weil kontextabhängig klar ist, ob die duale Basis des Dualraums oder die duale Basis im folgenden Sinne gemeint ist.

Lemma–Definition 5.19. *Sei $(V, \langle\ ,\ \rangle)$ ein endlich-dimensionaler K-Vektorraum mit einer perfekten symmetrischen Bilinearform. Die* ***duale Basis*** *zu einer Basis $\mathscr{B} = (b_1, \ldots, b_n)$ bezüglich $\langle\ ,\ \rangle$ ist die Basis $\mathscr{B}^* = (b_1^*, \ldots, b_n^*)$ mit der Eigenschaft*

$$\langle b_i, b_j^* \rangle = \begin{cases} 0 & i \neq j, \\ 1 & i = j. \end{cases}$$

Beweis. Sei $(\pi_1, \ldots, \pi_n)$ die zu $\mathscr{B}$ duale Basis von V^*. Der Riesz'sche Darstellungssatz, Korollar 5.6, garantiert die Existenz und Eindeutigkeit von Vektoren b_j^* mit

$$\pi_j = \langle -, b_j^* \rangle = r(b_j^*).$$

Die Vektoren $(b_1^*, \ldots, b_n^*)$ bilden eine Basis von V, weil die rechtspartielle Auswertung $r : V \to V^*$ ein Isomorphismus ist (die Paarung $\langle\ ,\ \rangle$ ist perfekt) und $r(\mathscr{B}^*)$ eine Basis von V^* ist. □

Bemerkung 5.20. (1) Die Gram'sche Matrix $\mathrm{M}^{\mathscr{B},\mathscr{B}^*}(\langle\ ,\ \rangle)$ ist die Einheitsmatrix. Und umgekehrt, wenn für eine Basis $\mathscr{C}$ die Gram'sche Matrix $\mathrm{M}^{\mathscr{B},\mathscr{C}}(\langle\ ,\ \rangle)$ die Einheitsmatrix ist, dann ist $\mathscr{C} = \mathscr{B}^*$ die duale Basis bzgl. $\langle\ ,\ \rangle$.

(2) Jede Basis $\mathscr{B}$ ist gleich ihrer doppelt dualen $(\mathscr{B}^*)^*$, da $\langle\ ,\ \rangle$ symmetrisch ist.

Beispiel 5.21. Die Standardbasis $\mathscr{E}$ des K^n ist gleich der eigenen dualen $\mathscr{E}^* = \mathscr{E}$ bezüglich des Standardskalarprodukts auf K^n.

Bemerkung 5.22. Die duale Basis $\mathscr{B}^*$ erlaubt es, die Koordinaten eines Vektors bezüglich $\mathscr{B}$ hinzuschreiben. Für alle $v \in V$ gilt

$$v = \sum_{i=1}^{n} \langle v, b_i^* \rangle b_i. \tag{5.2}$$

Man bestimmt die Koeffizienten im Ansatz $v = \sum_i \lambda_i b_i$ durch Anwendung der Linearform $\langle -, b_j^* \rangle$ als

$$\langle v, b_j^* \rangle = \langle \sum_i \lambda_i b_i, b_j^* \rangle = \sum_i \lambda_i \langle b_i, b_j^* \rangle = \sum_i \lambda_i \delta_{ij} = \lambda_j.$$

5.4. **Adjungierte Abbildungen.** Eine lineare Abbildung hat bezüglich perfekter Paarungen auf Quelle und Ziel einen Partner, die adjungierte Abbildung.

Satz–Definition 5.23. *Seien $\langle\ ,\ \rangle_V : V \times V \to K$ und $\langle\ ,\ \rangle_W : W \times W \to K$ perfekte symmetrische Bilinearformen von endlich-dimensionalen K-Vektorräumen, und sei $f : V \to W$ eine K-lineare Abbildung. Dann gibt es eine eindeutige K-lineare Abbildung*

$$f^* : W \to V$$

mit der Eigenschaft, daß für alle $v \in V$ und $w \in W$ gilt:

$$\langle f(v), w \rangle_W = \langle v, f^*(w) \rangle_V.$$

Dieses f^* heißt die zu f ***adjungierte Abbildung****.*

Beweis. Jedes $w \in W$ definiert eine Linearform $L_w : V \to K$

$$L_w(v) = \langle f(v), w \rangle_W.$$

Der Riesz'sche Darstellungssatz, Korollar 5.6, liefert ein eindeutiges Element $f^*(w) \in V$ mit

$$L_w = \langle -, f^*(w) \rangle_V.$$

Die so definierte Abbildung $f^* : W \to V$ erfüllt für alle $v \in V$ und $w \in W$

$$\langle f(v), w \rangle_W = \langle v, f^*(w) \rangle_V.$$

Der Eindeutigkeitsteil des Korollars 5.6 zeigt, daß f^* bereits hierdurch eindeutig festgelegt ist.

Es bleibt zu zeigen, daß f^* linear ist. Dazu betrachten wir $w_1, w_2 \in W$ und $\lambda_1, \lambda_2 \in K$. Da für alle $v \in V$ gilt

$$\begin{aligned}\langle v, \lambda_1 f^*(w_1) + \lambda_2 f^*(w_2)\rangle_V &= \lambda_1\langle v, f^*(w_1)\rangle_V + \lambda_2\langle v, f^*(w_2)\rangle_V \\ &= \lambda_1\langle f(v), w_1\rangle_W + \lambda_2\langle f(v), w_2\rangle_W = \langle f(v), \lambda_1 w_1 + \lambda_2 w_2\rangle_W,\end{aligned}$$

werden die Anforderung an $f^*(\lambda w_1 + \mu w_2)$ durch $\lambda_1 f^*(w_1) + \lambda_2 f^*(w_2)$ erfüllt. Die besagte Eindeutigkeit des Riesz'schen Darstellungssatzes zeigt

$$f^*(\lambda_1 w_1 + \lambda_2 w_2) = \lambda_1 f^*(w_1) + \lambda_2 f^*(w_2),$$

somit ist f^* linear. $\square$

Beispiel 5.24. Seien $V = K^n$ und $\langle\ ,\ \rangle_V = \langle\ ,\ \rangle$ das Standardskalarprodukt auf K^n, und seien $W = K^m$ und $\langle\ ,\ \rangle_W$ das Standardskalarprodukt auf K^m. Eine lineare Abbildung $f : K^n \to K^m$ ist Matrixmultiplikation mit einem $A \in \mathrm{M}_{m\times n}(K)$. Die Rechnung für alle $x \in K^n$ und $y \in K^m$

$$\langle Ax, y\rangle = (Ax)^t y = x^t(A^t y) = \langle x, A^t y\rangle$$

zeigt, daß die adjungierte Abbildung $f^* : K^m \to K^n$ durch Multiplikation mit der transponierten Matrix A^t vermittelt wird.

Wir sind eigentlich an Spezialfällen interessiert, siehe unten, aber es ist ein übliches Phänomen, daß die algebraischen Eigenschaften klarer sind, wenn man unnötige Identifikationen vermeidet.

Satz 5.25 (Funktorialität)**.** *Seien $\langle\ ,\ \rangle_U$, $\langle\ ,\ \rangle_V$ und $\langle\ ,\ \rangle_W$ perfekte symmetrische Bilinearformen auf endlich-dimensionalen K-Vektorräumen U, V und W.*

(1) Die Identität $\mathrm{id}_V : V \to V$ ist selbstadjungiert: es gilt

$$(\mathrm{id}_V)^* = \mathrm{id}_V\,.$$

(2) Seien $g : U \to V$ und $f : V \to W$ zwei K-lineare Abbildungen. Dann gilt

$$(fg)^* = g^* f^* : W \to U.$$

*Man spricht von **contravarianter Funktorialität**.*

Beweis. (1) ist trivial und (2) folgt sofort aus der Rechnung für $u \in U$ und $w \in W$

$$\langle u, g^*(f^*(w))\rangle_U = \langle g(u), f^*(w)\rangle_V = \langle f(g(u)), w\rangle_W. \qquad \square$$

Wir erlauben uns nun die Nachlässigkeit, in der Notation die Bilinearformen auf V und auf W nicht mehr zu unterscheiden.

Satz 5.26. *Seien $(V, \langle\ ,\ \rangle)$ und $(W, \langle\ ,\ \rangle)$ zwei endlich-dimensionale K-Vektorräume mit perfekten symmetrischen Bilinearformen. Sei $\mathscr{B} = (b_1, \ldots, b_n)$ eine Basis von V und $\mathscr{C} = (c_1, \ldots, c_m)$ eine Basis von W. Dann gilt*

$$\mathrm{M}^{\mathscr{C}^*}_{\mathscr{B}^*}(f^*) = \mathrm{M}^{\mathscr{B}}_{\mathscr{C}}(f)^t$$

für alle linearen Abbildungen $f : V \to W$.

Beweis. Mit (5.2) folgt für alle $1 \le j \le n$

$$f(b_j) = \sum_{i=1}^{m} \langle f(b_j), c_i^*\rangle c_i.$$

Wegen Symmetrie ist $\mathscr{B}$ die duale Basis zu $\mathscr{B}^*$. Mit (5.2) folgt für alle $1 \le i \le m$

$$f^*(c_i^*) = \sum_{j=1}^{n} \langle f^*(c_i^*), b_j\rangle b_j^*$$

Damit berechnet sich der ij-Matrixeintrag zu

$$\big(\mathrm{M}^{\mathscr{B}}_{\mathscr{C}}(f)\big)_{ij} = \langle f(b_j), c_i^*\rangle = \langle c_i^*, f(b_j)\rangle = \langle f^*(c_i^*), b_j\rangle = \big(\mathrm{M}^{\mathscr{C}^*}_{\mathscr{B}^*}(f^*)\big)_{ji}. \qquad \square$$

Satz 5.27. *Seien $(V, \langle\ ,\ \rangle)$ und $(W, \langle\ ,\ \rangle)$ zwei endlich-dimensionale K-Vektorräume mit perfekten symmetrischen Bilinearformen. Die Zuordnung $f \mapsto f^*$ definiert einen Isomorphismus von K-Vektorräumen*

$$(-)^* : \mathrm{Hom}_K(V, W) \to \mathrm{Hom}_K(W, V).$$

Das Inverse ist wieder $(-)^*$: für alle linearen Abbildungen $f : V \to W$ gilt $(f^*)^* = f$.

Beweis. Sei $f_i : V \to W$ und $\lambda_i \in K$ für $i = 1, 2$. Dann ist für alle $v \in V$ und $w \in W$:

$$\begin{aligned}\langle v, \lambda_1 f_1^*(w) + \lambda_2 f_2^*(w)\rangle &= \lambda_1\langle v, f_1^*(w)\rangle + \lambda_2\langle v, f_2^*(w)\rangle \\ &= \lambda_1\langle f_1(v), w\rangle + \lambda_2\langle f_2(v), w\rangle = \langle(\lambda_1 f_1 + \lambda_2 f_2)(v), w\rangle.\end{aligned}$$

Die Definition der adjungierten Abbildungen zeigt dann für alle $w \in W$

$$(\lambda_1 f_1 + \lambda_2 f_2)^*(w) = \lambda_1 f_1^*(w) + \lambda_2 f_2^*(w),$$

also die Linearität von $(-)^*$.

Die Aussage über das doppelt Adjungierte und damit sofort die Bijektivität von $(-)^*$ folgt aus der Rechnung für alle $v \in V$ und $w \in W$:

$$\langle v, (f^*)^*(w)\rangle = \langle f^*(v), w\rangle = \langle w, f^*(v)\rangle = \langle f(w), v\rangle = \langle v, f(w)\rangle.$$

Hier geht die Symmetrie von $\langle\ ,\ \rangle$ ein. In der Tat zeigt Eindeutigkeit in Korollar 5.6 (alternativ Lemma 5.2) wieder $f(w) = (f^*)^*(w)$ für alle $w \in W$. □

Beispiel 5.28. Wir betrachten die Drehung $R_\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ des $\mathbb{R}^2$ um den Winkel φ. Um die Begriffe Winkel und Drehung kümmern wir uns genauer in den Abschnitten §8.2 und §9.1. Die Drehung R_φ ist bezüglich Standardkoordinaten gegeben durch die Matrixmultiplikation mit

$$D_\varphi = \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}.$$

Wir statten $\mathbb{R}^2$ mit dem Standardskalarprodukt aus und wählen $\mathscr{B} = \mathscr{C}$ gleich der Standardbasis. Dann sind auch die dualen Basen $\mathscr{B}^* = \mathscr{C}^*$ gleich der Standardbasis, und die Matrix zu R_φ^* berechnet sich nach Satz 5.26 durch:

$$D_\varphi^t = \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ -\sin(\varphi) & \cos(\varphi) \end{pmatrix} = \begin{pmatrix} \cos(-\varphi) & -\sin(-\varphi) \\ \sin(-\varphi) & \cos(-\varphi) \end{pmatrix} = D_{-\varphi} = D_\varphi^{-1}.$$

Letzteres benutzt die aus der Analysis bekannte Gleichung $\cos^2(\varphi) + \sin^2(\varphi) = 1$. Es gilt also

$$R_\varphi^* = R_{-\varphi} = R_\varphi^{-1}.$$

Übungsaufgaben zu §5

Übungsaufgabe 5.1. Sei V ein K-Vektorraum. Wir betrachten die Abbildung

$$\begin{aligned}\iota : V &\to (V^*)^* \\ v &\mapsto (f \mapsto f(v)).\end{aligned}$$

Zeigen Sie, daß ι eine injektive K-lineare Abbildung ist, und daß ι sogar ein Isomorphismus ist, sofern $\dim_K(V) < \infty$.

Tipp: Benutzen Sie eine geeignete Paarung.

Übungsaufgabe 5.2. Finden Sie eine nichtausgeartete Paarung, die nicht perfekt ist.

Übungsaufgabe 5.3. Auf den stetig differenzierbaren periodischen Funktionen

$$C^1(\mathrm{S}^1, \mathbb{R}) := \{f : \mathbb{R} \to \mathbb{R}\ ;\ f(x+1) = f(x) \text{ für alle } x \in \mathbb{R}, \text{ und } f \text{ stetig differenzierbar}\}$$

sei durch

$$(f, g) := \int_0^1 f(\vartheta) g'(\vartheta) d\vartheta$$

eine Paarung

$$\langle\ ,\ \rangle : C^1(\mathrm{S}^1, \mathbb{R}) \times C^1(\mathrm{S}^1, \mathbb{R}) \to \mathbb{R}$$

definiert. Zeigen Sie, daß die konstanten Funktionen sowohl im Linkskern als auch im Rechtskern liegen.

Übungsaufgabe 5.4. Finden Sie eine perfekte Paarung $\langle\ ,\ \rangle : V \times V \to K$ und einen Unterraum $W \subseteq V$, so daß die Einschränkung von $\langle\ ,\ \rangle$ auf $W \times W \to K$ keine perfekte Paarung ist.

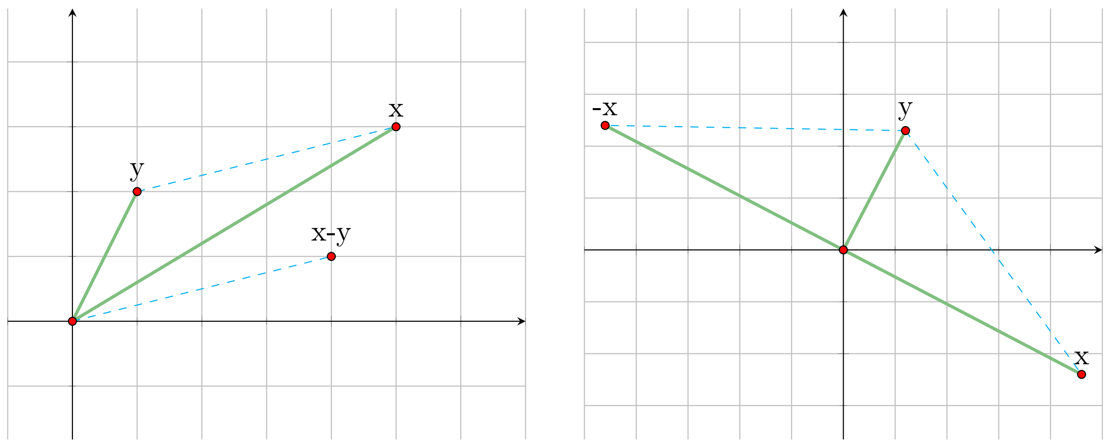
Übungsaufgabe 5.5. Beschreiben Sie eine symmetrische Bilinearform, bezüglich derer die Polarisationsformel aus Satz 5.16 zu einer binomischen Formel wird.

6. Orthogonalität

6.1. **Orthogonalität.** Letztendlich wollen wir mit Bilinearformen Geometrie mit Längen und Winkeln betreiben. Der Begriff der Orthogonalität ist ein Vorgriff hierauf. Zur Motivation betrachten wir das Standardskalarprodukt $\langle\ ,\ \rangle$ auf $\mathbb{R}^n$ und definieren die **Länge eines Vektors** $x = (x_1, \ldots, x_n)^t \in \mathbb{R}^n$ als

$$\|x\| = \sqrt{\langle x, x\rangle} = \Big(\sum_{i=1}^{n} x_i^2\Big)^{1/2}$$

und den Abstand von $x, y \in \mathbb{R}^n$ als $d(x, y) := \|x - y\|$.



(a) Abstand von Punkten im $\mathbb{R}^n$, ein Parallelogramm.

(b) Orthogonale Vektoren.

Abbildung 6. Euklidische Motivation für Länge und Orthogonalität.

Elementargeometrisch motiviert sagen wir, daß x und y orthogonal sind, wenn y denselben Abstand von x wie von $-x$ hat. Dies führt zu

$$\begin{aligned} d(y, -x)^2 - d(y, x)^2 &= \|y + x\|^2 - \|y - x\|^2 = \langle y + x, y + x\rangle - \langle y - x, y - x\rangle \\ &= \langle y, y + x\rangle + \langle x, y + x\rangle - \langle y, y - x\rangle + \langle x, y - x\rangle = \langle y, 2x\rangle + \langle x, 2y\rangle \\ &= 4\langle x, y\rangle, \end{aligned}$$

wobei wir die Symmetrie des Standardskalarproduktes benutzt haben, also

$$x \text{ senkrecht auf } y \iff \langle x, y\rangle = 0.$$

Dies erheben wir in der allgemeinen Situation zur Definition.

Definition 6.1. Sei V ein K-Vektorraum mit einer symmetrischen Bilinearform $\langle\ ,\ \rangle$. Zwei Vektoren $v, w \in V$ heißen **orthogonal** (bezüglich $\langle\ ,\ \rangle$), wenn

$$\langle v, w\rangle = 0$$

gilt. Als Notation vereinbaren wir $v \perp w$, wenn v und w orthogonal sind.

Bemerkung 6.2. (1) Orthogonalität ist kein Begriff des Vektorraums alleine, sondern hängt ab von dem Paar $\big(V, \langle\ ,\ \rangle\big)$.
(2) Da wir $\langle\ ,\ \rangle$ als symmetrisch vorausgesetzt haben, gilt

$$v \perp w \iff w \perp v.$$

Die Relation ‚orthogonal' ist symmetrisch, aber nicht transitiv.

Beispiel 6.3. (1) Bezüglich des Standardskalarprodukts auf K^n sind verschiedene Standardbasisvektoren e_i und e_j für $i \neq j$ orthogonal, denn

$$\langle e_i, e_j\rangle = \delta_{ij}.$$

(2) Ein Vektor kann zu sich selbst orthogonal sein. Wir betrachten die symmetrische Bilinearform $\langle v, w\rangle_A = v^t A w$ auf K^2 bezüglich der Matrix

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

Dies ist eine perfekte symmetrische Bilinearform, weil $A = A^t$ und $\det(A) \neq 0$. Trotzdem ist für $i = 1, 2$

$$\langle e_i, e_i\rangle = 0.$$

Beide Standardbasisvektoren sind orthogonal zu sich selbst. Dies ist auch ein Beispiel einer perfekten symmetrischen Bilinearform ($\det(A) = -1$), die eingeschränkt auf einen linearen Unterraum (hier $Ke_1 \subseteq K^2$) nicht perfekt ist.

Definition 6.4. Sei V ein K-Vektorraum und $\langle\ ,\ \rangle$ eine symmetrische Bilinearform auf V. Der **Orthogonalraum** zu einem K-Untervektorraum $U \subseteq V$ ist

$$U^\perp = \{v \in V\ ;\ \langle v, u\rangle = 0 \text{ für alle } u \in U\}.$$

Dies ist der Kern von $V \to U^*$, $v \mapsto \langle v, -\rangle|_U$, somit ein K-Untervektorraum von V.

Bemerkung 6.5. (1) Weil wir mit einer symmetrischen Bilinearform arbeiten, gilt auch

$$U^\perp = \{v \in V\ ;\ \langle u, v\rangle = 0 \text{ für alle } u \in U\}.$$

Linksorthogonalraum und Rechtsorthogonalraum stimmen überein, weshalb wir nur kurz vom Orthogonalraum sprechen.
(2) Für eine symmetrische Bilinearform $\langle\ ,\ \rangle$ auf V sind Linkskern und Rechtskern identisch und stimmen mit dem Orthogonalraum $V^\perp$ zum ganzen Vektorraum V überein. Die doppeldeutige Notation $V^\perp$ ist also ungefährlich.
(3) Man kann Orthogonalität auch für nicht notwendig symmetrische Bilinearformen oder allgemeiner für Paarungen definieren. Dann muß man aber zwischen linksorthogonal und rechtsorthogonal unterscheiden. Das ist uns zu mühsam.

Proposition 6.6. *Sei V ein K-Vektorraum und $\langle\ ,\ \rangle$ eine symmetrische Bilinearform auf V. Seien U und W Unterräume von V. Dann gilt:*
(1) $U \subseteq W \Longrightarrow W^\perp \subseteq U^\perp$.
(2) $(U + W)^\perp = U^\perp \cap W^\perp$.
(3) $U^\perp + W^\perp \subseteq (U \cap W)^\perp$.
(4) $U \subseteq (U^\perp)^\perp$.

Wenn $\langle\ ,\ \rangle$ zusätzlich eine perfekte Paarung ist, dann gilt

$$\dim(U^\perp) + \dim(U) = \dim(V),$$

in (1) gilt Äquivalenz und in (3) und (4) gilt Gleichheit.

Beweis. (1) folgt sofort aus der Definition.

(2) Wegen $U \subseteq U + W$ folgt $(U+W)^\perp \subseteq U^\perp$. Mit W analog, also $(U+W)^\perp \subseteq U^\perp \cap W^\perp$.

Für die umgekehrte Inklusion müssen wir zeigen, daß jedes $x \in U^\perp \cap W^\perp$ mit einem beliebigen $y \in U + W$ zu 0 paart. Es gibt $u \in U$ und $w \in W$ mit $y = u + w$. Dann folgt

$$\langle x, y\rangle = \langle x, u + w\rangle = \langle x, u\rangle + \langle x, w\rangle = 0.$$

(3) Wegen $U \cap W \subseteq U$ gilt $U^\perp \subseteq (U \cap W)^\perp$ und analog mit W.

(4) folgt aufgrund der Symmetrie sofort aus der Definition.

Sei nun $\langle\ ,\ \rangle$ perfekt. Die Einschränkung auf U nach linkspartieller Auswertung

$$\varphi : V \xrightarrow{\ell} V^* \xrightarrow{(-)|_U} U^*$$

ist surjektiv: jede Linearform auf U läßt sich auf V fortsetzen und ℓ ist ein Isomorphismus nach Voraussetzung. Per Definition ist $\ker(\varphi) = U^\perp$. Aus der Kern/Bild-Dimensionsformel folgt

$$\dim(U^\perp) = \dim(V) - \dim(U).$$

Wenden wir dies auch auf $U^\perp$ an, so folgt

$$\dim\big((U^\perp)^\perp\big) = \dim(V) - \dim(U^\perp) = \dim(V) - \big(\dim(V) - \dim(U)\big) = \dim(U).$$

Weil $U \subseteq (U^\perp)^\perp$, folgt durch Dimensionsvergleich sogar Gleichheit $U = (U^\perp)^\perp$.

Jetzt folgt Gleichheit in (3) aus (2) angewandt auf $U^\perp$ und $W^\perp$:

$$U^\perp + W^\perp = \big((U^\perp + W^\perp)^\perp\big)^\perp = \big((U^\perp)^\perp \cap (W^\perp)^\perp\big)^\perp = (U \cap W)^\perp.$$

Und in (1) folgt die Rückrichtung durch (1) angewandt auf $U^\perp$ und $W^\perp$:

$$W^\perp \subseteq U^\perp \Longrightarrow (U^\perp)^\perp \subseteq (W^\perp)^\perp \iff U \subseteq W.$$ □

6.2. **Orthogonalbasen und Diagonalform.**

Definition 6.7. Eine **Orthogonalbasis** eines K-Vektorraums V mit Bilinearform $\langle\ ,\ \rangle$ ist eine Basis $\mathscr{B}$ von V, so dass für alle verschiedenen Basisvektoren $b \neq b'$ aus $\mathscr{B}$ gilt:

$$\langle b, b'\rangle = 0 = \langle b', b\rangle.$$

Wenn $\langle\ ,\ \rangle$ symmetrisch ist, bedeutet das, $b \perp b'$ sind orthogonal.

Beispiel 6.8. Bezüglich des Standardskalarprodukts auf K^n ist die Standardbasis eine Orthogonalbasis.

Lemma 6.9. *Sei $\langle\ ,\ \rangle$ eine Bilinearform auf dem K-Vektorraum V mit Basis $\mathscr{B}$ und zugehöriger Gram'scher Matrix A. Dann sind äquivalent:*

(a) *$\mathscr{B}$ ist Orthogonalbasis für V.*
(b) *A ist eine Diagonalmatrix.*

Beweis. Dies folgt sofort aus der Definition von Orthogonalbasis und Gram'scher Matrix. □

Lemma 6.10. *Sei $\mathscr{B} = (b_1, \ldots, b_n)$ eine Basis des K-Vektorraums V. Sei i fixiert mit $1 \leq i \leq n$. Sei $x_j \in K$ für alle $j \neq i$. Dann ist auch*

$$\mathscr{B}' = (b_1 + x_1 b_i, \ldots, b_{i-1} + x_{i-1} b_i,\ b_i,\ b_{i+1} + x_{i+1} b_i, \ldots, b_n + x_n b_i)$$

eine Basis von V.

Beweis. Das Tupel $\mathscr{B}'$ hat die richtige Anzahl von Vektoren. Es reicht zu zeigen, daß diese ein Erzeugendensystem sind. Dafür reicht es, die Vektoren der Basis $\mathscr{B}$ als Linearkombination darstellen zu können. Der Vektor b_i ist dabei und für alle $j \neq i$ gilt

$$b_j = (b_j + x_j b_i) - x_j b_i \in \langle \mathscr{B}' \rangle_K. \qquad \square$$

Bemerkung 6.11. Im Folgenden wird für einen Körper oft die Voraussetzung $2 \in K^\times$ gefordert werden. Diese Notation muß erklärt werden. Dazu erinnern wir daran, daß jeder Körper Elemente $0, 1 \in K$ hat, und damit für jedes $n \in \mathbb{N}_0$ ein Element (zur Unterscheidung kurzzeitig n_K benannt)

$$n_K = \underbrace{1 + \ldots + 1}_{n\text{-mal}} \in K.$$

Es ist also $0_K = 0$ und $1_K = 1$ im jeweiligen Körper. Man überzeugt sich leicht, daß für alle $n, m \in \mathbb{N}$ gilt

$$(n + m)_K = n_K + m_K.$$

Allgemeiner, mittels additiven inversen Elementen $(-n)_K = -(n_K)$, erhält man einen Ringhomomorphismus $\mathbb{Z} \to K$

$$n \mapsto n_K.$$

Wenn im Folgenden von $2 \in K$ die Rede ist, so ist stets $2_K = 1+1 \in K$ gemeint. Für die Körper $\mathbb{Q}, \mathbb{R}$ und $\mathbb{C}$ ist $2 \neq 0$, und damit gilt $2 \in K^\times$. Aber für $K = \mathbb{F}_2$ oder auch $\mathbb{F}_q$ mit einer 2er Potenz q ist $2 = 0$, und damit gilt $2 \notin K^\times$.

Theorem 6.12 (Diagonalformensatz)**.** *Sei K ein Körper mit $2 \in K^\times$. Sei $f : V \times V \to K$ eine Bilinearform auf dem K-Vektorraum V mit* $\dim_K(V) < \infty$. *Dann sind äquivalent:*

(a) *Die Bilinearform f ist symmetrisch.*
(b) *V hat eine Orthogonalbasis bezüglich f.*
(c) *Es gibt eine Basis $\mathscr{B}$ von V, so daß die zugehörige Gram'sche Matrix* $\mathrm{M}^{\mathscr{B},\mathscr{B}}(f)$ *eine Diagonalmatrix ist.*

Beweis. Die Äquivalenz (b) $\iff$ (c) folgt aus Lemma 6.9. Da Diagonalmatrizen symmetrisch sind, folgt (c) $\Longrightarrow$ (a) aus Proposition 5.15. Wir müssen nur noch (a) $\Longrightarrow$ (b) zeigen.

Schritt 1: Wenn f die Nullform ist, also $f(v, w) = 0$ für alle $v, w \in V$, dann ist nichts zu tun: jede Basis ist Orthogonalbasis. Andernfalls zeigt die Polarisationsformel (Symmetrie!) aus Satz 5.16

$$f(v, w) = \frac{1}{2}\big(f(v + w, v + w) - f(v, v) - f(w, w)\big)$$

die Existenz eines Vektors $b_1 \in V$ mit $f(b_1, b_1) \neq 0$.

Schritt 2: Wir ergänzen b_1 zu einer Basis $\mathscr{C} = (b_1, c_2, c_3, \ldots, c_n)$ von V. Wir möchten nun die c_i durch Vektoren, die zu b_1 orthogonal sind, ersetzen. Dazu machen wir den Ansatz

$$b_i' = c_i + \lambda_i b_1$$

mit zu bestimmenden $\lambda_i \in K$. Paarung mit b_1 liefert

$$0 = f(c_i, b_1) + \lambda_i f(b_1, b_1),$$

was sich nach Voraussetzung eindeutig nach λ_i auflösen läßt: wir setzen für $i = 2, \ldots, n$

$$b_i' = c_i + \lambda_i b_1 = c_i - \frac{f(c_i, b_1)}{f(b_1, b_1)} b_1.$$

Dann ist $\mathscr{B}' = (b_1, b_2', b_3', \ldots, b_n')$ auch eine Basis von V nach Lemma 6.10, und für $i = 2, \ldots, n$ gilt nach Wahl von λ_i

$$f(b_i', b_1) = f(c_i, b_1) - \frac{f(c_i, b_1)}{f(b_1, b_1)} f(b_1, b_1) = 0.$$

Der Unterraum $W = \langle b_2', \ldots, b_n' \rangle_K \subseteq V$ liegt somit im Orthogonalraum $\langle b_1 \rangle^\perp$.

Schritt 3: Jetzt argumentieren wir per Induktion nach $n = \dim(V)$. Der Induktionsanfang $\dim(V) = 0$ ist trivial.

Die Einschränkung von f auf W ist eine symmetrische Bilinearform auf W, und wegen $\dim(W) = \dim(V) - 1$ gibt es per Induktionsannahme eine Orthogonalbasis $b_2, \ldots, b_n$ von W. Damit ist aber $\mathscr{B} = (b_1, \ldots, b_n)$ eine Orthogonalbasis von V: In der Tat gilt $f(b_1, b_i) = 0$ für $i \geq 2$, da $b_i \in \langle b_1 \rangle^{\perp}$. Und für $i \neq j$ mit $i, j \geq 2$ gilt $f(b_i, b_j) = 0$ per Induktionsannahme. In der linearen Hülle von $\mathscr{B}$ liegt die Basis $\mathscr{B}'$, damit ist $\mathscr{B}$ ein Erzeugendensystem aus $\dim(V)$-vielen Vektoren und somit selbst eine Basis. □

Korollar 6.13. *Sei K ein Körper mit $2 \in K^{\times}$. Sei $f : V \times V \to K$ eine symmetrische Bilinearform auf dem K-Vektorraum V mit $\dim_K(V) = n < \infty$. Dann gibt es eine Basis $\mathscr{B}$ und $\lambda_1, \ldots, \lambda_n \in K$ mit*

$$f(v, w) = \lambda_1 x_1 y_1 + \ldots + \lambda_n x_n y_n$$

für alle $v \in V$ und $w \in W$ dargestellt durch Koordinaten $x = \kappa_{\mathscr{B}}(v) = (x_1, \ldots, x_n)^t$ und $y = \kappa_{\mathscr{B}}(w) = (y_1, \ldots, y_n)^t$.

Die Bilinearform ist perfekt $\iff$ $\lambda_i \neq 0$ für alle $i = 1, \ldots, n$.

Beweis. Dazu eignet sich jede Orthogonalbasis $\mathscr{B}$, deren Existenz durch Theorem 6.12 gesichtert ist. Die Gram'sche Matrix $D = \mathrm{M}^{\mathscr{B},\mathscr{B}}(f)$ ist dann eine Diagonalmatrix $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ und die Formel ergibt sich sofort aus (4.1).

Die Aussage über die Perfektheit der Paarung folgt wegen $\det(D) = \prod_{i=1}^{n} \lambda_i$ sofort aus Satz 5.5. □

Korollar 6.14 (Normalformensatz)**.** *Sei $2 \in K^{\times}$. Zu einer symmetrischen Matrix $A = A^t \in \mathrm{M}_n(K)$ gibt es $S \in \mathrm{GL}_n(K)$ und eine Diagonalmatrix $D \in \mathrm{M}_n(K)$, so daß*

$$A = S^t D S.$$

Beweis. Wir betrachten die Bilinearform $\langle x, Ay \rangle = x^t A y$ auf K^n. Nach Proposition 5.15 ist $\langle\ ,\ \rangle_A$ symmetrisch, weil A symmetrisch ist. Theorem 6.12 zeigt die Existenz einer Orthogonalbasis $\mathscr{B}$ für $\langle\ ,\ \rangle_A$. Dann ist nach Lemma 6.9

$$D = \mathrm{M}^{\mathscr{B},\mathscr{B}}(\langle -, - \rangle_A)$$

eine Diagonalmatrix. Sei $\mathscr{E}$ die Standardbasis und dann $S = \mathrm{M}^{\mathscr{E}}_{\mathscr{B}}(\mathrm{id})$ die Basiswechselmatrix. Die Transformationsformel für Gram'sche Matrizen aus Proposition 4.19 liefert dann

$$A = \mathrm{M}^{\mathscr{E},\mathscr{E}}(\langle -, - \rangle_A) = S^t D S.$$ □

Sei K ein Körper. Ein Vertretersystem von K modulo den Quadraten $(K^{\times})^2$ ist eine Teilmenge $\Lambda \subseteq K$, so daß jedes Element $x \in K$ von der Form

$$x = \lambda a^2$$

mit $a \in K^{\times}$ und eindeutigem $\lambda \in \Lambda$ ist.

Satz 6.15. *Sei K ein Körper mit $2 \in K^{\times}$. Sei $\Lambda \subseteq K$ ein Vertretersystem von K modulo den Quadraten $(K^{\times})^2$. Sei V ein endlichdimensionaler K-Vektorraum.*

Zu jeder symmetrischen Bilinearform $\langle\ ,\ \rangle$ auf V gibt es eine Orthogonalbasis, bezüglich der die Gram'sche Matrix nur Einträge aus Λ hat.

Beweis. Wir wählen zunächst nach Theorem 6.12 eine Orthogonalbasis $\mathscr{C} = (c_1, \ldots, c_n)$. Wir schreiben die Diagonaleinträge $x_i = \langle c_i, c_i \rangle$ in der Form

$$x_i = \lambda_i a_i^2$$

mit $\lambda_i \in \Lambda$ und $a_i \in K^{\times}$. Mit

$$b_i = \frac{1}{a_i} c_i$$

ist $\mathscr{B} = (b_1, \ldots, b_n)$ auch eine Orthogonalbasis, und zwar mit

$$\langle b_i, b_i \rangle = \frac{1}{a_i^2} \langle c_i, c_i \rangle = \frac{x_i}{a_i^2} = \lambda_i.$$

Die Basis $\mathscr{B}$ hat die gesuchten Eigenschaften. □

Beispiel 6.16. (1) Für $K = \mathbb{R}$ kann man $\Lambda = \{-1, 0, 1\}$ nehmen. Jede symmetrische Bilinearform eines endlich-dimensionalen $\mathbb{R}$-Vektorraums hat eine Orthogonalbasis, deren Gram'sche Matrix diagonal ist und nur $1, 0$ oder -1 auf der Diagonalen hat.

(2) Für $K = \mathbb{C}$ kann man $\Lambda = \{0, 1\}$ nehmen.

(3) Für $K = \mathbb{Q}$ kann man als Λ die Menge der quadratfreien ganzen Zahlen und 0 nehmen. Eine ganze Zahl n heißt quadratfrei, wenn es keine ganze Zahl $d > 1$ gibt, so daß $d^2 \mid n$. Äquivalent dazu tritt in der Primfaktorzerlegung von einem quadratfreien $n \in \mathbb{Z}$ jede Primzahl höchstens einmal auf.

6.3. Anisotropie und das Gram-Schmidt'sche Verfahren.

Definition 6.17. Sei V ein K-Vektorraum mit einer symmetrischen Bilinearform $\langle\ ,\ \rangle$.

(1) Ein **anisotroper** Vektor ist ein $v \in V$ mit

$$\langle v, v \rangle \neq 0.$$

Ein **isotroper** Vektor ist ein $v \in V$ mit $\langle v, v \rangle = 0$.

(2) Die Bilinearform $\langle\ ,\ \rangle$ heißt **anisotrop**, wenn alle $v \in V$, $v \neq 0$ anisotrop sind.

Proposition 6.18. *Eine anisotrope symmetrische Bilinearform ist nichtausgeartet.*

Beweis. Zu jedem Vektor $v \neq 0$ ist gerade $w = v$ selbst ein Partner mit $\langle v, w \rangle \neq 0$. □

Beispiel 6.19. Sei $K = \mathbb{R}$.

(1) Das Standardskalarprodukt auf $\mathbb{R}^n$ ist anisotrop, denn für $v = \sum_i x_i e_i \neq 0$ ist stets

$$\langle v, v \rangle = \sum_i x_i^2 > 0.$$

Später spezialisieren wir uns auf allgemeinere Skalarprodukte von $\mathbb{R}$-Vektorräumen, deren Anisotropie auf Positivität beruht, ein Begriff der spezifisch für $\mathbb{R}$ oder allgemeiner für angeordnete Körper ist. Die Voraussetzung ‚anisotrop' ist also eine in der Praxis häufig anzutreffende Eigenschaft.

(2) Das Standardskalarprodukt auf $(\mathbb{F}_2)^2$ ist nicht anisotrop, weil

$$\langle e_1 + e_2, e_1 + e_2 \rangle = 1 + 1 = 0.$$

Allgemeiner ist $v = \binom{x}{y} \in (\mathbb{F}_p)^2$ isotrop, wenn

$$0 = \langle v, v \rangle = x^2 + y^2$$

und solche $v \neq 0$ gibt es genau dann, wenn $p = 2$ oder $p \equiv 1 \pmod 4$. Für die letzte Aussage sei auf die Vorlesung „Elementare Zahlentheorie" verwiesen.

(3) Das Standardskalarprodukt auf $(\mathbb{F}_p)^3$ ist nie anisotrop. Das lernt man auch in der Vorlesung „Elementare Zahlentheorie": die Gleichung

$$x^2 + y^2 + z^2 = 0$$

hat stets Lösungen $0 \neq (x, y, z) \in \mathbb{F}_p^3$.

(4) Die Bilinearform auf $\mathbb{R}^2$ bezüglich der Standardbasis gegeben durch

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

ist nicht anisotrop. Der Vektor $v = \binom{x}{y}$ ist isotrop $\iff$ $x = \pm y$:

$$\langle v, v\rangle_A = x^2 - y^2 = (x-y)(x+y).$$

Insbesondere ist $e_1 + e_2$ isotrop.

Theorem 6.12 hat für anisotrope symmetrische Bilinearformen einen konstruktiven Beweis. Der Beweis beinhaltet einen Algorithmus, der im anisotropen Fall aus einer beliebigen Basis eine Orthogonalbasis macht.

Satz 6.20 (Gram–Schmidt'sches Orthogonalisierungsverfahren)**.** *Sei V ein K-Vektorraum mit einer anisotropen symmetrischen Bilinearform $\langle\ ,\ \rangle$. Sei $\mathscr{C} = (c_1, \ldots, c_n)$ eine Basis von V. Wir konstruieren für $j = 0, \ldots, n$ eine Basis $\mathscr{B}^{(j)} = (b_1, \ldots, b_j, c_{j+1}, \ldots, c_n)$ durch*

$$b_j = c_j - \sum_{i=1}^{j-1} \frac{\langle c_j, b_i\rangle}{\langle b_i, b_i\rangle} b_i.$$

Die Basis $\mathscr{B} = \mathscr{B}^{(n)}$ ist eine Orthogonalbasis von V bezüglich $\langle\ ,\ \rangle$.

Beweis. Per Induktion nach j zeigen wir, daß $\mathscr{B}^{(j)}$ eine Basis ist und die auftretenden Nenner $\neq 0$ sind. Für $j = 0$ gilt dies nach Annahme an $\mathscr{C}$.

Sei $\mathscr{B}^{(j-1)}$ eine Basis. Nach der Definition von $\mathscr{B}^{(j)}$ ist klar, daß die lineare Hülle von $\mathscr{B}^{(j)}$ gleich der linearen Hülle von $\mathscr{B}^{(j-1)}$ ist, denn

$$c_j = b_j + \sum_{i=1}^{j-1} \frac{\langle c_j, b_i\rangle}{\langle b_i, b_i\rangle} b_i \in \langle \mathscr{B}^{(j)}\rangle_K.$$

Daher ist $\mathscr{B}^{(j)}$ ein Erzeugendensystem aus $\dim(V)$-vielen Vektoren, also eine Basis.

Da $\langle\ ,\ \rangle$ anisotrop ist, sind alle auftretenden Nenner von 0 verschieden, denn die Vektoren b_i gehören zu einer Basis und sind daher $b_i \neq 0$. Die Konstruktion ist wohldefiniert.

Wir zeigen nun per Induktion nach j, daß für $k < j$ stets $(b_k, b_j) = 0$ gilt. Für $j = 1$ ist nichts zu zeigen. Wir nehmen an, die Behauptung ist wahr für $j-1$, und rechnen dann

$$\begin{aligned}\langle b_k, b_j\rangle = \langle b_k, c_j - \sum_{i=1}^{j-1} \frac{\langle c_j, b_i\rangle}{\langle b_i, b_i\rangle} b_i\rangle &= \langle b_k, c_j\rangle - \sum_{i=1}^{j-1} \frac{\langle c_j, b_i\rangle}{\langle b_i, b_i\rangle}\langle b_k, b_i\rangle \\ &= \langle b_k, c_j\rangle - \frac{\langle c_j, b_k\rangle}{\langle b_k, b_k\rangle}\langle b_k, b_k\rangle = 0.\end{aligned}$$

Dies zeigt, daß der Algorithmus mit einer Orthogonalbasis endet. □

Bemerkung 6.21. Der Basiswechsel $S = \mathrm{M}^{\mathscr{C}}_{\mathscr{B}}(\mathrm{id})$ von der Basis $\mathscr{C}$ nach $\mathscr{B}$ aus Satz 6.20 ist eine unipotente (nur 1 auf der Diagonalen) obere Dreiecksmatrix. In der Tat gilt

$$c_j = b_j + \sum_{i=1}^{j-1} \frac{\langle c_j, b_i\rangle}{\langle b_i, b_i\rangle} b_i,$$

so daß $S = (s_{ij})_{1 \le i,j \le n}$ mit

$$s_{ij} = \begin{cases} \frac{\langle c_j, b_i\rangle}{\langle b_i, b_i\rangle} & i < j, \\ 1 & i = j, \\ 0 & i > j. \end{cases}$$

Mit S ist auch S^{-1}, die Basiswechselmatrix $\mathrm{M}^{\mathscr{B}}_{\mathscr{C}}(\mathrm{id})$, eine obere Dreiecksmatrix, denn die Menge der oberen Dreiecksmatrizen ist eine Untergruppe von $\mathrm{GL}_n(K)$.

Bemerkung 6.22. Das Gram-Schmidt'sche Orthogonalisierungsverfahren hat in Satz 6.20 als Voraussetzung, daß die Bilinearform $\langle\ ,\ \rangle$ anisotrop sein muß. Dies ist nötig, damit der Algorithmus beweisbar eine Orthogonalbasis liefert und nicht wegen einer undurchführbaren Division durch 0 vorzeitig zum Stehen kommt.

Ohne die Voraussetzung 'anisotrop' kann man den Algorithmus aber **trotzdem auf gut Glück versuchen** in der Hoffnung, daß keiner der produzierten Basisvektoren b_i isotrop ist.

Ist b_i leider isotrop, dann muß man eben wie im Induktionsschritt des Beweises von Theorem 6.12 verfahren und ein neues b_i wählen. Als ersten Versuch permutiere man die noch zu orthogonalisierenden restlichen Basisvektoren (man ändere die Reihenfolge).

6.4. **Orthogonale Summe.** In Linearer Algebra wurde das Konzept der **inneren direkten Summe** behandelt. Ein K-Vektorraum V ist die innere direkte Summe zweier Unterräume W_1 und W_2, wenn $V = W_1 + W_2$ und $W_1 \cap W_2 = (0)$. Wir schreiben

$$V = W_1 \oplus W_2.$$

Dies bedeutet, daß jedes $v \in V$ auf eindeutige Weise als $v = w_1 + w_2$ mit $w_i \in W_i$ geschrieben werden kann: die lineare Abbildung

$$W_1 \times W_2 \to V, \quad (w_1, w_2) \mapsto w_1 + w_2$$

ist ein Isomorphismus.

Dieses wollen wir hier durch die orthogonale innere direkte Summe ergänzen, wenn die Summenzerlegung eine symmetrische Bilinearform berücksichtigt.

Definition 6.23. Ein K-Vektorraum V mit symmetrischer Bilinearform $f : V \times V \to K$ ist die **innere orthogonale Summe**

$$V = W_1 \oplus^{\perp} W_2$$

zweier Unterräume $W_1, W_2 \subseteq V$, wenn

(i) $V = W_1 \oplus W_2$ die innere direkte Summe von Vektorräumen ist,
(ii) und alle $w_1 \in W_1$ und $w_2 \in W_2$ orthogonal sind:

$$f(w_1, w_2) = 0.$$

Bemerkung 6.24. Wenn $V = W_1 \oplus^{\perp} W_2$ gilt unter der Identifikation $W_1 \times W_2 = V$ für alle $v = (v_1, v_2)$, $w = (w_1, w_2) \in W_1 \times W_2 = V$

$$f(v, w) = f(v_1, w_1) + f(v_2, w_2).$$

Notation 6.25. Als Notation für die orthogonale Summe findet man auch $W_1 \perp W_2$.

Man kann eine orthogonale Summenzerlegung der Gram'schen Matrix ablesen, sofern die Basis aus Basen der Summanden durch Vereinigung entsteht.

Proposition 6.26. *Sei V mit symmetrischer Bilinearform $f : V \times V \to K$ die direkte Summe $V = W_1 \oplus W_2$ zweier K-Vektorräume W_i. Seien $\mathscr{B}_i$ Basen von W_i und*

$$A_i = \mathrm{M}^{\mathscr{B}_i, \mathscr{B}_i}(f_i)$$

die Gram'sche Matrix der eingeschränkten Bilinearform $f_i = f|_{W_i \times W_i}$. Dann sind äquivalent:

(a) V ist orthogonale Summe $V = W_1 \oplus^{\perp} W_2$.
(b) Die Gram'sche Matrix von f bezüglich der Basis[6] $\mathscr{B} = \mathscr{B}_1 \cup \mathscr{B}_2$ ist die Blockmatrix

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}.$$

Insbesondere ist f perfekt genau dann, wenn f_1 und f_2 perfekt sind.

[6] Notationsmißbrauch! Eine Basis ist keine Menge, sondern ein Tupel. Die Ordnung ist wichtig, denn sie bestimmt zum Beispiel die Reihenfolge der Koordinaten. Wir verstehen unter $\mathscr{B} = \mathscr{B}_1 \cup \mathscr{B}_2$ dasjenige Tupel, das aus $\mathscr{B}_1$ durch Anhängen des Tupels $\mathscr{B}_2$ entsteht.

Beweis. Die Gram'sche Matrix von f hat Blockform

$$A = \begin{pmatrix} A_1 & B \\ C & A_2 \end{pmatrix}$$

mit $C = B^t$ mit den Einträgen $\langle b, b' \rangle$ für $b \in \mathscr{B}_1$ und $b' \in \mathscr{B}_2$. Die Matrix B beschreibt damit die Einschränkung

$$f_{12} := f|_{W_1 \times W_2} : W_1 \times W_2 \to K.$$

Die direkte Summe ist orthogonal genau dann, wenn $f_{12} = 0$ gilt. Nach Proposition 4.4 ist die Paarung f_{12} die Nullabbildung, wenn die Werte auf Paaren von Basisvektoren 0 sind. Das ist genau die Bedingung $B = 0$.

Behandeln wir nun die Frage, ob f perfekt ist. Als Blockmatrix hat A die Determinante

$$\det(A) = \det(A_1)\det(A_2).$$

Es ist f (bzw. f_i) perfekt genau dann, wenn $\det(A) \in K^\times$ (bzw. $\det(A_i) \in K^\times$). $\square$

Bemerkung 6.27. Sei $2 \in K^\times$ und $(V, \langle\ ,\ \rangle)$ ein endlich-dimensionaler K-Vektorraum mit symmetrischer Bilinearform. Theorem 6.12 besagt nun, daß V eine orthogonale Summe von 1-dimensionalen K-Vektorräumen mit symmetrischer Bilinearform ist. Es gibt nämlich eine Orthogonalbasis $\mathscr{B} = (b_1, \ldots, b_n)$, und dann ist mit $V_i = K b_i$

$$V = V_1 \oplus^\perp \ldots \oplus^\perp V_n$$

eine orthogonale Summe.

Definition 6.28. Ein **orthogonales Komplement** eines Untervektorraums U in einem K-Vektorraum V mit symmetrischer Bilinearform $f : V \times V \to K$ ist ein Untervektorraum $W \subseteq V$, so daß V (mit f) die innere orthogonale Summe ist:

$$V = U \oplus^\perp W.$$

Proposition 6.29. *Sei V ein endlich-dimensionaler K-Vektorraum mit anisotroper symmetrischer Bilinearform $\langle\ ,\ \rangle$, und sei $U \subseteq V$ ein Untervektorraum.*

(1) $U^\perp$ ist ein orthogonales Komplement.
(2) Das orthogonale Komplement ist eindeutig.

Beweis. Für $u \in U \cap U^\perp$ folgt $\langle u, u \rangle = 0$. Da $\langle\ ,\ \rangle$ anisotrop ist, schließen wir $u = 0$. Damit ist $U \cap U^\perp = (0)$.

Anisotrope symmetrische Bilinearformen sind nichtausgeartet, denn für alle $v \in V$, $v \neq 0$ gilt ja $\langle v, v \rangle \neq 0$, also gilt nach Propostion 6.6 und der Dimensionsformel

$$\dim(U + U^\perp) = \dim(U) + \dim(U^\perp) - \dim(U \cap U^\perp) = \dim(V).$$

Daraus folgt $U + U^\perp = V$ und die Summe ist sogar direkt: $U \oplus U^\perp = V$. Per Definition ist die Summe auch orthogonal und $U^\perp$ ein orthogonales Komplement.

Jetzt zeigen wir die Eindeutigkeit des orthogonalen Komplements. Sei W ein orthogonales Komplement von U. Dann ist

$$\dim(W) = \dim(V) - \dim(U) = \dim(U^\perp).$$

Weil $W \subseteq U^\perp$ per Definition gilt, folgt $W = U^\perp$. $\square$

Satz 6.30. *Seien $(V, \langle\ ,\ \rangle_V)$ und $(W, \langle\ ,\ \rangle_W)$ endlich-dimensionale K-Vektorräume mit perfekter symmetrischer Bilinearform, und sei $f : V \to W$ eine K-lineare Abbildung. Dann gilt*

(1) $\ker(f^*) = \operatorname{im}(f)^\perp$,
(2) $\ker(f) = \operatorname{im}(f^*)^\perp$,
(3) f und f^* haben denselben Rang,

Ist $\langle\ ,\ \rangle_W$ auch anisotrop, so gilt darüber hinaus

$$W = \mathrm{im}(f) \oplus^{\perp} \ker(f^*).$$

Beweis. Weil $(f^*)^* = f$ folgt (2) aus (1) angewandt auf $f^* : W \to V$. Aussage (1) folgt aus

$$w \in \ker(f^*) \iff f^*(w) = 0 \iff \langle -, f^*(w)\rangle_V = 0 \iff \langle f(-), w\rangle_W = 0 \iff w \in \mathrm{im}(f)^{\perp}.$$

(3) Es gilt mit Proposition 6.6, Aussage (2) und der Kern/Bild-Dimensionsformel

$$\begin{aligned}\mathrm{rg}(f^*) &= \dim_K(\mathrm{im}(f^*)) = \dim_K(V) - \dim_K(\mathrm{im}(f^*)^{\perp}) \\ &= \dim_K(V) - \dim(\ker(f)) = \dim_K(\mathrm{im}(f)) = \mathrm{rg}(f).\end{aligned}$$

Sei nun $\langle\ ,\ \rangle$ anisotrop. Dann folgt

$$V = \mathrm{im}(f) \oplus^{\perp} \mathrm{im}(f)^{\perp} = \mathrm{im}(f) \oplus^{\perp} \ker(f^*)$$

aus Proposition 6.29 und (1). □

Korollar 6.31. *Sei $A \in \mathrm{M}_{m\times n}(K)$. Dann ist*

(1) $\ker(A^t) = \mathrm{im}(A)^{\perp}$,
(2) $\ker(A) = \mathrm{im}(A^t)^{\perp}$,
(3) *A und A^t haben denselben Rang,*

Ist $\langle\ ,\ \rangle$ anisotrop, so gilt darüber hinaus

$$K^m = \mathrm{im}(A) \oplus^{\perp} \ker(A^t).$$

Beweis. Das ist Satz 6.30 ausformuliert im Beispiel 5.24, denn das Standardskalarprodukt ist eine perfekte symmetrische Bilinearform. □

Weil sich der Beweis aus den Eigenschaften der adjungierten Abbildung sofort ergibt, betonenen wir nochmals ein aus der Linearen Algebra bekanntes Ergebnis:

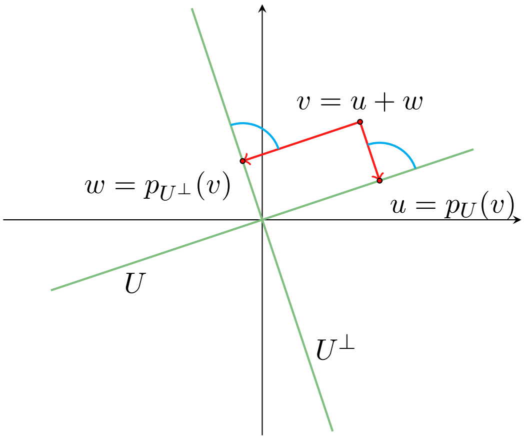
Korollar 6.32. *Sei $A \in \mathrm{M}_{m\times n}(K)$ eine beliebige Matrix. Dann ist der Zeilenrang von A gleich dem Spaltenrang von A.*

Beweis. Der Zeilenrang von A ist gleich dem Spaltenrang von A^t und wird mit $\mathrm{rg}(A^t)$ bezeichnet. Nach Korollar 6.31 (3) ist $\mathrm{rg}(A^t) = \mathrm{rg}(A)$, also gleich dem Spaltenrang von A. □

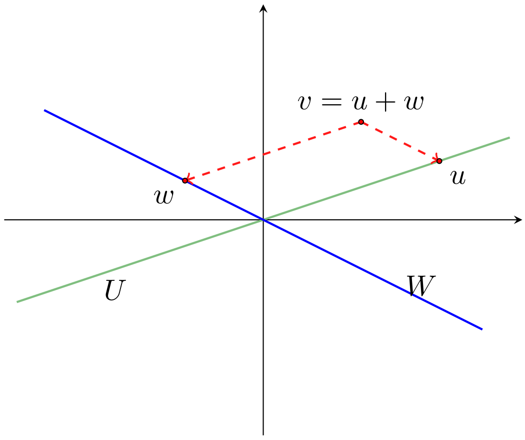
Definition 6.33. Sei $(V, \langle\ ,\ \rangle)$ ein K-Vektorraum mit symmetrischer anisotroper Bilinearform, und sei $U \subseteq V$ ein Unterraum. Zur Zerlegung $V = U \oplus^{\perp} U^{\perp}$ gehört eine **orthogonale Projektion**

$$p_U : V \to V.$$

Für $v = u + w$ mit $u \in U$ und $w \in U^{\perp}$ ist $p_U(v) = u$.



(a) Orthogonale Projektion



(b) Keine orthogonale Projektion

Abbildung 7. Projektion auf $U \subseteq K^2$.

Proposition 6.34. *Sei $(V, \langle\ ,\ \rangle)$ ein K-Vektorraum mit symmetrischer anisotroper Bilinearform, und sei $U \subseteq V$ ein Unterraum. Sei $\mathscr{B} = (b_1, \ldots, b_n)$ eine Orthogonalbasis von V, so daß $(b_1, \ldots, b_r)$ eine Orthogonalbasis von U ist. Die orthogonale Projektion $p_U : V \to V$ auf U ist gegeben durch*

$$p_U(v) = \sum_{i=1}^{r} \langle v, b_i^* \rangle b_i \qquad \text{für alle } v \in V.$$

Beweis. Weil $\mathscr{B}$ eine Orthogonalbasis ist, sind die $b_{r+1}, \ldots, b_n$ in

$$\langle b_1, \ldots, b_r \rangle^\perp = U^\perp.$$

Es reicht nun, p_U mit $\sum_{i=1}^{r} \langle -, b_i^* \rangle b_i$ auf der Basis $\mathscr{B}$ zu vergleichen:

$$\sum_{i=1}^{r} \langle b_j, b_i^* \rangle b_i = \sum_{i=1}^{r} \delta_{ij} b_i = \begin{cases} b_j & j \leq r \\ 0 & j > r \end{cases} = p_U(b_j).$$

□

Proposition 6.35. *Sei $(V, \langle\ ,\ \rangle)$ ein K-Vektorraum mit symmetrischer anisotroper Bilinearform, und sei $U \subseteq V$ ein Unterraum. Dann ist die orthogonale Projektion auf U*

$$p_U : V \to V$$

eindeutig charakterisiert als lineare Abbildung $p : V \to V$ mit den folgenden drei Eigenschaften:

(i) $p^2 = p$,
(ii) $\ker(p) = U^\perp$,
(iii) $p^* = p$.

Beweis. Wir weisen zunächst (i)-(iii) für die orthogonale Projektion nach. Eigenschaft (i) gilt ganz allgemein für eine Projektion auf einen direkten Summanden, und (ii) folgt sofort aus der Definition. Zu (iii) betrachten wir $v_i = u_i + w_i \in V$ mit $u_i \in U$ und $w_i \in U^\perp$ für $i = 1, 2$. Dann ist

$$\langle p_U(v_1), v_2 \rangle = \langle u_1, u_2 + w_2 \rangle = \langle u_1, u_2 \rangle = \langle u_1 + w_1, u_2 \rangle = \langle v_1, p_U(v_2) \rangle,$$

und daraus folgt $p_U^* = p_U$.

Sei nun $p : V \to V$ eine lineare Abbildung mit (i)–(iii). Wegen $p^2 = p$ ist p die Projektion auf $\mathrm{im}(p)$ in der inneren direkten Summe $V = \mathrm{im}(p) \oplus \ker(p)$. Die Behauptung $p = p_U$ folgt nun nach Satz 6.30 und Proposition 6.6 mit

$$\mathrm{im}(p) = \mathrm{im}(p^*) = \ker(p)^\perp = (U^\perp)^\perp = U.$$

□

Bemerkung 6.36. Sei V ein Vektorraum mit anisotroper symmetrischer Bilinearform $\langle\ ,\ \rangle$. Wir geben nun eine Beschreibung des Gram-Schmidt-Orthogonalisierungsverfahrens mittels orthogonaler Projektionen. Sei $\mathscr{C} = (c_1, \ldots, c_n)$ die Ausgangsbasis und $\mathscr{B} = (b_1, \ldots, b_n)$ die Orthogonalbasis, die man rekursiv aus Gram-Schmidt für alle $1 \leq j \leq n$ mittels

$$b_j = c_j - \sum_{k=1}^{j-1} \frac{\langle c_j, b_k\rangle}{\langle b_k, b_k\rangle} b_k$$

bestimmt. Man sieht sofort aus der Orthogonalität von $\mathscr{B}$, daß die zu $\mathscr{B}$ duale Basis durch

$$b_i^* = \frac{1}{\langle b_i, b_i\rangle} b_i$$

gegeben ist. Wir bezeichnen die lineare Hülle der ersten i Vektoren von $\mathscr{B}$ als

$$U_i = \langle b_1, \ldots, b_i\rangle_K.$$

Die orthogonale Projektion $p_i = p_{U_i} : V \to V$ auf U_i wird berechnet durch

$$p_i(v) = \sum_{k=1}^{i} \langle v, b_k^*\rangle b_k = \sum_{k=1}^{i} \frac{\langle v, b_k\rangle}{\langle b_k, b_k\rangle} b_k.$$

Folglich berechnet man im Gram-Schmidt-Orthogonalisierungsverfahren

$$b_j = c_j - p_{j-1}(c_j) = (\mathrm{id}_V - p_{j-1})(c_j) = p_{U_{j-1}^\perp}(c_j).$$

Man bestimmt also b_j dadurch, daß man von c_j die orthogonale Projektion auf U_{j-1} abzieht, so daß die Projektion auf den Orthogonalraum $U_{j-1}^\perp$ übrig bleibt. Damit ist offenbar b_j senkrecht auf allen $b_i \in U_{j-1}$ für $i \leq j-1$.

6.5. **Orthonormalbasen und orthogonale Matrizen.** Vektorräume haben lineare Struktur und man betrachtet Basen, um Koordinaten zu bekommen. Alle Basen sind gleich gut.

In Vektorräumen mit symmetrischer Bilinearform möchten wir Basen betrachten, die an die Bilinearform angepaßt sind. Wenn sie existieren, sind dies die Orthonormalbasen. Die später zu behandelnden euklidischen Vektorräume sind genau die reellen Vektorräume mit Orthonormalbasen. Alle Orthonormalbasen sind gleich gut.

Definition 6.37. Ein **Orthonormalsystem** in einem K-Vektorraum V mit symmetrischer Bilinearform $\langle\ ,\ \rangle$ ist ein Tupel von Vektoren $(v_1, \ldots, v_r)$, für die

$$\langle v_i, v_j\rangle = \delta_{ij} \text{ für alle } 1 \leq i, j \leq r.$$

Lemma 6.38. *Die Vektoren eines Orthonormalsystems $(v_1, \ldots, v_r)$ sind linear unabhängig.*

Beweis. Wenn $\sum_{i=1}^r \lambda_i v_i = 0$ mit $\lambda_i \in K$, dann gilt

$$\lambda_j = \langle v_j, \sum_{i=1}^{r} \lambda_i v_i\rangle = 0.$$

Die Linearkombination ist also trivial. Damit ist das Orthonormalsytem linear unabhängig. □

Definition 6.39. Eine **Orthonormalbasis** eines K-Vektorraums V mit symmetrischer Bilinearform $\langle\ ,\ \rangle$ ist eine Basis $\mathscr{B} = (b_1, \ldots, b_n)$, die ein Orthonormalsystem ist. Wir sagen kurz $\mathscr{B}$ ist eine ONB.

Bemerkung 6.40. Offensichtlich ist jede ONB eine Orthogonalbasis.

Proposition 6.41. *Sei $(V, \langle\ ,\ \rangle)$ ein K-Vektorraum mit symmetrischer Bilinearform und $\mathscr{B} = (b_1, \ldots, b_n)$ eine Basis von V. Wir statten K^n mit dem Standardskalarprodukt aus. Dann sind die folgenden Aussagen äquivalent.*

(a) $\mathscr{B}$ *ist ONB.*

(b) *Die Gram'sche Matrix von $\langle\ ,\ \rangle$ bezüglich $\mathscr{B}$ ist die Einheitsmatrix.*
(c) *$\langle\ ,\ \rangle$ ist perfekt und $\mathscr{B}$ ist seine eigene Dualbasis $\mathscr{B}^* = \mathscr{B}$ bezüglich $\langle\ ,\ \rangle$.*
(d) *Der Koordinatenisomorphismus $\kappa_{\mathscr{B}} : V \to \mathbb{R}^n$ ist eine isometrische Abbildung: für alle $v, w \in V$ gilt*

$$\langle v, w\rangle = \langle \kappa_{\mathscr{B}}(v), \kappa_{\mathscr{B}}(w)\rangle.$$

Beweis. (a) $\iff$ (b) ist klar per Definition.

(b) $\Longrightarrow$ (c): Die Bilinearform ist nach Satz 5.5 perfekt, weil $\det(\mathrm{M}^{\mathscr{B},\mathscr{B}}(\langle\ ,\ \rangle)) = \det(\mathbf{1}) = 1 \neq 0$. Weiter kann man aus der Gram'schen Matrix für alle $1 \leq i, j \leq n$

$$\langle b_i, b_j\rangle = \delta_{ij}$$

ablesen, woraus $b_i^* = b_i$ für alle $1 \leq i \leq n$ folgt.

(c) $\Longrightarrow$ (d): Sei $\mathscr{B}$ selbstdual, also $b_i^* = b_i$ für alle $i = 1, \ldots, n$. Für alle $v \in V$ folgt

$$v = \sum_{i=1}^n \langle v_i, b_i^*\rangle b_i = \sum_{i=1}^n \langle v, b_i\rangle b_i = \sum_{i=1}^n \langle v, b_i\rangle b_i^*, \quad \text{also} \quad \kappa_{\mathscr{B}}(v) = \begin{pmatrix} \langle v, b_1\rangle \\ \vdots \\ \langle v, b_n\rangle \end{pmatrix}.$$

Dann gilt für alle $v, w \in V$:

$$\begin{aligned}\langle v, w\rangle &= \langle \sum_{i=1}^n \langle v, b_i\rangle b_i, \sum_{j=1}^n \langle w, b_j\rangle b_j^*\rangle = \sum_{1\leq i,j\leq n} \langle v, b_i\rangle \langle w, b_j\rangle \langle b_i, b_j^*\rangle \\ &= \sum_{1\leq i,j\leq n} \langle v, b_i\rangle\langle w, b_j\rangle \delta_{ij} = \sum_{i=1}^n \langle v, b_i\rangle\langle w, b_i\rangle = \langle \kappa_{\mathscr{B}}(v), \kappa_{\mathscr{B}}(w)\rangle.\end{aligned}$$

(d) $\Longrightarrow$ (a): Wenn $\kappa_{\mathscr{B}}$ eine isometrische Abbildung ist, dann ist $\mathscr{B}$ eine ONB, weil

$$\langle b_i, b_j\rangle = \langle \kappa_{\mathscr{B}}(b_i), \kappa_{\mathscr{B}}(b_j)\rangle = \langle e_i, e_j\rangle = \begin{cases} 1 & i = j \\ 0 & i \neq j. \end{cases}$$ □

Proposition 6.42. *Sei $f : V \to V$ ein Endomorphismus eines K-Vektorraums V mit symmetrischer Bilinearform $\langle\ ,\ \rangle$ und mit ONB $\mathscr{B}$.*

(1) *Die adjungierte Abbildung hat die transponierte Darstellungsmatrix:*

$$\mathrm{M}^{\mathscr{B}}_{\mathscr{B}}(f^*) = \mathrm{M}^{\mathscr{B}}_{\mathscr{B}}(f)^t.$$

(2) *Es ist f **selbstadjungiert**, d.h. $f = f^*$, genau dann, wenn die Matrix bezüglich einer (äquivalent aller) ONB eine symmetrische Matrix ist.*

Beweis. Das ist Satz 5.26 mit $\mathscr{B}^* = \mathscr{B}$. Die Aussage (2) folgt sofort aus (1). □

Der Basiswechsel zwischen ONB geschieht mit speziellen Matrizen, den orthogonalen Matrizen.

Definition 6.43. Sei K ein Körper.

(1) Eine **orthogonale Matrix** ist eine Matrix in $\mathrm{GL}_n(K)$, so daß $A^tA = \mathbf{1}$ gilt.
(2) Die **orthogonale Gruppe** ist die Untergruppe von $\mathrm{GL}_n(K)$

$$\mathrm{O}_n(K) = \{A \in \mathrm{M}_n(K)\ ;\ A^tA = \mathbf{1}\}.$$

Im speziellen Fall $K = \mathbb{R}$ schreibt man auch

$$\mathrm{O}(n) = \mathrm{O}_n(\mathbb{R}).$$

(3) Die **spezielle orthogonale Gruppe** ist die Untergruppe von $\mathrm{SL}_n(K)$

$$\mathrm{SO}_n(K) = \{A \in \mathrm{SL}_n(K)\ ;\ A^tA = \mathbf{1}\}.$$

Im speziellen Fall $K = \mathbb{R}$ schreibt man auch

$$\mathrm{SO}(n) = \mathrm{SO}_n(\mathbb{R}).$$

Beweis. Wir müssen zeigen, daß $\mathrm{O}_n(K)$ eine Untergruppe von $\mathrm{GL}_n(K)$ ist. Für
$$\mathrm{SO}_n(K) = \mathrm{O}_n(K) \cap \mathrm{SL}_n(K)$$
folgt die Untergruppeneigenschaft dann sofort.

Ein $A \in \mathrm{O}_n(K)$ ist wegen $A^tA = \mathbf{1}$ invertierbar, also in $\mathrm{GL}_n(K)$. Insbesondere folgt automatisch auch $AA^t = \mathbf{1}$. Wenn $A, B \in \mathrm{O}_n(K)$, dann ist
$$(AB)^t(AB) = B^tA^tAB = B^t(A^tA)B = B^tB = \mathbf{1},$$
und auch $AB \in \mathrm{O}_n(K)$. Für das Inverse A^{-1} gilt
$$(A^{-1})^tA^{-1} = (A^t)^{-1}A^{-1} = (AA^t)^{-1} = \mathbf{1},$$
und somit $A^{-1} \in \mathrm{O}_n(K)$. Damit ist $\mathrm{O}_n(K) \subseteq \mathrm{GL}_n(K)$ eine Untergruppe. □

Lemma 6.44. *Sei $A \in \mathrm{M}_n(K)$. Dann*
$$A \text{ orthogonal} \iff A^t \text{ orthogonal.}$$

Beweis. Weil $(A^t)^t = A$ gilt, reicht eine Richtung. Sei A orthogonal. Dann ist $A^t = A^{-1} \in \mathrm{O}_n(K)$ auch orthogonal. □

Satz 6.45. *Sei K ein Körper. Seien $\mathscr{B}$ und $\mathscr{C}$ Basen eines K-Vektorraums V mit perfekter symmetrischer Bilinearform $\langle\ ,\ \rangle$, und sei $S = \mathrm{M}^{\mathscr{B}}_{\mathscr{C}}(\mathrm{id})$ die Basiswechselmatrix. Dann gelten alle drei der folgenden Ausssagen, sobald zwei davon eintreffen (eine „2 aus 3“ Eigenschaft).*

(a) *$\mathscr{B}$ ist ONB.*
(b) *$\mathscr{C}$ ist ONB.*
(c) *$S \in \mathrm{O}_n(K)$ ist orthogonal.*

Beweis. Sei $\mathscr{B} = (b_1, \ldots, b_n)$ und $\mathscr{C} = (c_1, \ldots, c_n)$ und $S = (s_{ij})$. Dann ist $b_j = \sum_{k=1}^n s_{kj}c_k$. Wenn $\mathscr{C}$ ONB ist, dann folgt
$$\langle b_i, b_j\rangle = \langle \sum_{k=1}^n s_{ki}c_k, \sum_{k=1}^n s_{kj}c_k\rangle = \sum_{k=1}^n s_{ki}s_{kj} = (S^tS)_{ij},$$
somit ist S orthogonal genau dann, wenn $\mathscr{B}$ eine ONB ist.

Jetzt müssen wir noch zeigen, daß mit S orthogonal und $\mathscr{B}$ eine ONB dann auch $\mathscr{C}$ eine ONB ist. Aber das folgt aus Symmetrie, wenn wir $\mathscr{B}$ mit $\mathscr{C}$ und dadurch S mit $\mathrm{M}^{\mathscr{C}}_{\mathscr{B}}(\mathrm{id}) = S^{-1} = S^t$ ersetzen, was dann auch eine orthogonale Matrix ist, aus dem bereits bewiesenen. □

Korollar 6.46. *Sei $A \in \mathrm{M}_n(K)$. Dann sind äquivalent:*

(a) *A ist orthogonal.*
(b) *Die Spalten von A sind eine ONB von K^n mit dem Standardskalarprodukt.*
(c) *Die transponierten Zeilen von A sind eine ONB von K^n mit dem Standardskalarprodukt.*

Beweis. Nach Lemma 6.44 gilt $A \in \mathrm{O}_n(K) \iff A^t \in \mathrm{O}_n(K)$. Transponieren vertauscht die Aussagen (b) und (c). Also reicht es (a) $\iff$ (b) zu zeigen.

Sei $A = [a_1, \ldots, a_n]$ die spaltenweise Darstellung. Aus beiden Aussagen folgt, daß $\mathscr{B} = (a_1, \ldots, a_n)$ eine Basis von K^n ist. Weil die Standardbasis des K^n eine ONB ist und A die Basiswechselmatrix von der Basis $(a_1, \ldots, a_n)$ in die Standardbasis ist, folgt die Äquivalenz (a) $\iff$ (b) aus Satz 6.45. □

Proposition 6.47. *Für alle $A \in \mathrm{O}_n(K)$ gilt* $\det(A) = 1$ *oder* $\det(A) = -1$.

Beweis. Das folgt sofort aus
$$\det(A)^2 = \det(A^t) \cdot \det(A) = \det(A^tA) = \det(\mathbf{1}) = 1,$$
denn $\det(A)$ ist Nullstelle des Polynoms $X^2 - 1 = (X-1)(X+1)$. □

Korollar 6.48. *Für alle $n \geq 1$ ist $\mathrm{SO}_n(K)$ eine Untergruppe vom Index 2 in $\mathrm{O}_n(K)$, der Kern des Determinantenhomomorphismus*

$$\det : \mathrm{O}_n(K) \twoheadrightarrow \{\pm 1\}.$$

Beweis. Das folgt sofort aus der Definition von $\mathrm{SO}_n(K)$ und $\det(\mathrm{diag}(-1, 1, \ldots, 1)) = -1$. □

Satz 6.49. *Ist $A \in \mathrm{O}_n(K)$, dann gilt für alle $v, w \in K^n$:*

$$\langle Av, Aw \rangle = \langle v, w \rangle.$$

Beweis. Das folgt wegen $A^t A = \mathbf{1}$ aus $\langle v, w \rangle = \langle \mathbf{1}v, w \rangle = \langle A^t Av, w \rangle = \langle Av, Av \rangle$. □

Übungsaufgaben zu §6

Übungsaufgabe 6.1. Zeigen Sie durch ein Beispiel, daß die Relation $v \perp w$ nicht transitiv ist.

Übungsaufgabe 6.2. Sei K ein Körper mit $2 \in K^\times$. Sei $A = A^t \in \mathrm{M}_n(K)$ eine symmetrische Matrix. Zeigen Sie, daß es ein $S \in \mathrm{GL}_n(K)$ gibt, so daß $S^t AS$ Diagonalmatrix ist.

Übungsaufgabe 6.3. Sei K ein Körper mit $2 = 0$, d.h. ein Körper der Charakteristik 2.

Zeigen Sie, daß die Bilinearform auf K^2 zur Matrix

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

keine Orthogonalbasis besitzt, also bezüglich keiner Basis die Gram'sche Matrix eine Diagonalmatrix ist.

Übungsaufgabe 6.4. Für eine natürliche Zahl $n \in \mathbb{N}$ sei $V_n \subseteq \mathbb{R}[X]$ der Vektorraum der Polynome vom Grad $\leq n$. Für reelle Zahlen $a < b$ statten wir V_n mit der symmetrischen Bilinearform $\langle\ ,\ \rangle_{[a,b]}$ aus, die auf Polynomen $f, g \in \mathbb{R}[X]$ den Wert

$$(f, g)_{[a,b]} := \int_a^b f(x)g(x)dx$$

annimmt.

(1) Zeigen Sie, daß $\langle\ ,\ \rangle_{[a,b]}$ anisotrop ist.
(2) Bestimmen Sie das orthogonale Komplement von V_1 in V_2 für $\langle\ ,\ \rangle_{[0,1]}$.

Übungsaufgabe 6.5. Sei V ein K-Vektorraum mit einer symmetrischen Bilinearform $\langle\ ,\ \rangle$.

Zeigen Sie, daß $v \in V$ genau dann anisotrop ist, wenn $V = \langle v \rangle \oplus v^\perp$ eine direkte Summe ist.

Übungsaufgabe 6.6. Sei V ein K-Vektorraum mit perfekter symmetrischer Bilinearform $\langle\ ,\ \rangle$.

Zeigen Sie, daß jede Orthogonalbasis aus anisotropen Vektoren besteht.

Übungsaufgabe 6.7. Sei $(V, \langle\ ,\ \rangle)$ ein endlich-dimensionaler Vektorraum mit einer perfekten symmetrischen Bilinearform. Sei $f : V \to V$ ein Endomorphismus und $P(X) \in K[X]$ ein Polynom. Zeigen Sie, daß

$$P(f)^* = P(f^*).$$

Teil 3. Euklidische Vektorräume

In diesem Kapitel arbeiten wir mit **Vektorräumen über den reellen Zahlen** $\mathbb{R}$. Der wesentliche Unterschied zwischen $\mathbb{R}$ und einem beliebigen Körper ist die Anordnung: für $x, y \in \mathbb{R}$ kann man von

$$x \leq y$$

sprechen.

Definition 6.50. Eine Anordnung auf einem Körper K ist eine Relation $\leq$ auf K mit den folgenden Eigenschaften: für alle $x, y, z \in K$ gilt

(i) $x \leq x$,
(ii) aus $x \leq y$ und $y \leq x$ folgt $x = y$,
(iii) aus $x \leq y$ und $y \leq z$ folgt $x \leq z$,
(iv) es gilt $x \leq y$ oder $y \leq x$,
(v) aus $x \leq y$ folgt $x + z \leq y + z$,
(vi) aus $x \leq y$ und $0 \leq z$ folgt $xz \leq yz$.

Man schreibt

$$x < y \iff x \leq y \text{ und } x \neq y.$$

Ein **angeordneter Körper** ist ein Körper mit einer Anordnung.

Wir erinnern, daß für $n \in \mathbb{N}$ das Element $n \in K$ bedeutet $1 + \ldots + 1$ (n Summanden).

Proposition 6.51. *Sei K ein Körper mit Anordnung $\leq$.*

(i) *Für alle $n \in \mathbb{N}$ gilt $0 < n$. Insbesondere ist $0 < 1$.*
(ii) *Für alle $x \in K$ gilt $0 < x \iff -x < 0$.*
(iii) *Für alle $x \in K$ gilt $0 \leq x^2$.*

Beweis. Das folgt leicht aus den Axiomen der Anordnung. □

Bemerkung 6.52. Der Körper $\mathbb{C}$ ist kein angeordneter Körper. Denn, sei $\leq$ eine Anordnung auf $\mathbb{C}$, dann folgt stets, egal ob $i > 0$ oder $i < 0$ gilt,

$$0 < i^2 = -1 < 0,$$

ein Widerspruch.

Bemerkung 6.53. Der Körper $\mathbb{F}_p$ ist kein angeordneter Körper. Sei $\leq$ eine Anordnung auf $\mathbb{F}_p$. Dann ist

$$0 < p = 0,$$

ein Widerspruch.

7. Skalarprodukte

7.1. **Definitheit symmetrischer Bilinearformen.** Die zentrale Definition diese Kapitels ist die folgende.

Definition 7.1. Eine symmetrische Bilinearform $\langle \ , \ \rangle$ auf einem $\mathbb{R}$-Vektorraum V heißt

(1) **positiv definit**, wenn für alle $v \in V$, $v \neq 0$ gilt $\langle v, v \rangle > 0$,
(2) **positiv semi-definit**, wenn für alle $v \in V$, $v \neq 0$ gilt $\langle v, v \rangle \geq 0$,
(3) **negativ definit**, wenn für alle $v \in V$, $v \neq 0$ gilt $\langle v, v \rangle < 0$,
(4) **negativ semi-definit**, wenn für alle $v \in V$, $v \neq 0$ gilt $\langle v, v \rangle \leq 0$,
(5) **indefinit**, wenn keiner der Fälle (1)-(4) eintritt.

Eine symmetrische Matrix $A \in \mathrm{M}_n(\mathbb{R})$ heißt **positiv/negativ (semi-)definit** oder **indefinit**, wenn die zugehörige Bilinearform $\langle \ , \ \rangle_A$ auf $\mathbb{R}^n$ entsprechend positiv/negativ (semi-)definit oder indefinit ist.

Wir nennen $\langle \ , \ \rangle$ ein **Skalarprodukt** (oder **inneres Produkt**), wenn $\langle \ , \ \rangle$ eine symmetrische und positiv definite Bilinearform ist.

Beispiel 7.2. Hier sind die typischen Beispiele.

(1) Das Standardskalarprodukt auf $\mathbb{R}^n$ ist positiv definit:
$$\langle v, v\rangle = \sum_{i=1}^n x_i^2 > 0 \quad \text{für alle } v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \neq 0.$$

(2) Die Matrix $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ ist indefinit: $\langle e_1, e_1\rangle_A = 1$ während $\langle e_2, e_2\rangle_A = -1$.

Proposition 7.3. *Eine positiv (bzw. negativ) definite symmetrische Bilinearform ist anisotrop und insbesondere perfekt.*

Beweis. Trivial. □

Proposition 7.4. *Sei λ ein Eigenwert einer positiv (bzw. negativ) definiten Matrix $A \in \mathrm{M}_n(\mathbb{R})$. Dann ist $\lambda > 0$ (bzw. $\lambda < 0$).*

Beweis. Sei $v \in \mathbb{R}^n$ ein zugehöriger Eigenvektor. Dann ist
$$\langle v, v\rangle_A = v^t A v = \langle v, Av\rangle = \langle v, \lambda v\rangle = \lambda \cdot \langle v, v\rangle.$$
Daraus ergibt sich
$$\lambda = \frac{\langle v, Av\rangle}{\langle v, v\rangle}$$
und λ hat das gleiche Vorzeichen wie $\langle v, v\rangle_A$. □

Satz 7.5. *Sei V ein $\mathbb{R}$-Vektorraum mit symmetrischer Bilinearform $\langle\ ,\ \rangle$. Dann gilt*
$$V \text{ hat ONB} \iff \langle\ ,\ \rangle \text{ ist positiv definit.}$$

Beweis. Sei $\mathscr{B}$ eine ONB. Weil das Standardskalarprodukt auf $\mathbb{R}^n$ positiv definit ist, gilt für alle $v \in V$ nach Proposition 6.41 (d)
$$\langle v, v\rangle = \langle \kappa_{\mathscr{B}}(v), \kappa_{\mathscr{B}}(v)\rangle \geq 0$$
mit Gleichheit genau für $\kappa_{\mathscr{B}}(v) = 0$, also $v = 0$. Damit ist $\langle\ ,\ \rangle$ auf V positiv definit.

Sei umgekehrt $\langle\ ,\ \rangle$ positiv definit. Weil $2 \in \mathbb{R}^\times$ gibt es nach dem Diagonalformensatz, Theorem 6.12 eine Orthogonalbasis $\mathscr{B}' = (b_1', \ldots, b_n')$. Weil $\langle\ ,\ \rangle$ positiv definit ist und man in $\mathbb{R}$ Wurzeln aus positiven reellen Zahlen ziehen kann, gibt für es $i = 1, \ldots, n$ Zahlen $\lambda_i \in \mathbb{R}$, $\lambda_i \neq 0$ mit
$$\langle b_i', b_i'\rangle = \lambda_i^2.$$
Dann sind $b_i := \lambda_i^{-1} b_i'$ die Vektoren einer ONB. □

Lemma 7.6. *Sei $(V, \langle\ ,\ \rangle)$ ein endlich-dimensionaler $\mathbb{R}$-Vektorraum mit symmetrischer Bilinearform, und sei A die Gram'sche Matrix bezüglich einer Basis von V. Dann ist das Vorzeichen von*
$$\det(A)$$
unabhängig von der Wahl der Basis (oder der Wert ist immer 0).

Beweis. Nach der Formel aus Proposition 4.19 für den Basiswechsel der Gram'schen Matrix gibt es für die Gram'sche Matrix B zu einer anderen Basis ein $S \in \mathrm{GL}_n(\mathbb{R})$ mit
$$B = S^t A S.$$
Dann ist aber
$$\det(B) = \det(S^t A S) = \det(S^t)\det(A)\det(S) = \det(S)^2 \det(A),$$
und weil $\det(S)^2 > 0$ haben $\det(B)$ und $\det(A)$ das gleiche Vorzeichen (oder beide sind 0 im nicht perfekten Fall). □

Lemma 7.7. *Sei $(V, \langle\ ,\ \rangle)$ ein endlich-dimensionaler $\mathbb{R}$-Vektorraum mit symmetrischer Bilinearform, und sei A die Gram'sche Matrix bezüglich einer Basis von V.*

(1) *Wenn $\langle\ ,\ \rangle$ positiv definit ist, dann gilt*

$$\det(A) > 0.$$

(2) *Wenn $\langle\ ,\ \rangle$ negativ definit ist, dann gilt*

$$(-1)^{\dim(V)} \det(A) > 0.$$

Beweis. Da das Vorzeichen nach Lemma 7.6 nicht von der Wahl der Basis abhängt, dürfen wir nach Theorem 6.12 eine Orthogonalbasis wählen. Die Diagonaleinträge sind von der Form $\langle b, b\rangle$ für Basisvektoren b und sind positiv (bzw. negativ), wenn $\langle\ ,\ \rangle$ positiv (bzw. negativ) definit ist. Das Vorzeichen der Determinante ergibt sich sofort, weil nun $\det(A)$ nur das Produkt der Diagonaleinträge ist. □

Beispiel 7.8. Die Matrix

$$A = \begin{pmatrix} 3 & -1 \\ -1 & 2016 \end{pmatrix}$$

ist positiv definit. Das liegt im Wesentlichen daran, daß die Diagonalterme dominieren und positiv sind. Genauer: für $v = (x, y)^t$ gilt

$$\langle v, v\rangle_A = v^t A v = 3x^2 - 2xy + 2016y^2 = 2x^2 + (x-y)^2 + 2015y^2$$

und dies ist als Summe von Quadraten positiv, außer wenn $x = y = 0$.

Satz 7.9 (Hauptminorenkriterium für positive/negative Definitheit)**.** *Sei $(V, \langle\ ,\ \rangle)$ ein endlich-dimensionaler $\mathbb{R}$-Vektorraum mit symmetrischer Bilinearform, und sei A die Gram'sche Matrix bezüglich einer Basis von V. Sei A_r die Matrix aus den ersten r Zeilen und Spalten von A. Dann gilt:*

(1) *$\langle\ ,\ \rangle$ ist positiv definit $\iff$ für alle $1 \leq r \leq \dim(V)$ gilt $\det(A_r) > 0$.*
(2) *$\langle\ ,\ \rangle$ ist negativ definit $\iff$ für alle $1 \leq r \leq \dim(V)$ gilt $(-1)^r \det(A_r) > 0$.*

Beweis. Sei A die Gram'sche Matrix zur Basis $\mathscr{B} = (b_1, \ldots, b_n)$. Dann ist A_r die Gram'sche Matrix für die Einschränkung von $\langle\ ,\ \rangle$ auf die lineare Hülle $U_r = \langle b_1, \ldots, b_r\rangle_K$. Die Einschrängkung ist wieder positiv (bzw. negativ) definit. Das Vorzeichen von $\det(A_r)$ folgt aus Lemma 7.7.

Für die Umkehrung überlegen wir zuerst, daß die symmetrische Bilinearform $-\langle\ ,\ \rangle$ positiv definit ist genau dann, wenn $\langle\ ,\ \rangle$ negativ definit ist. Sei $B_r = -A_r$ die Gram'sche Matrix zu $-\langle\ ,\ \rangle$ eingeschränkt auf U_r. Dann ist

$$\det(B_r) = \det(-A_r) = (-1)^r \det(A_r).$$

Also folgt (2) aus (1) angewandt auf $-\langle\ ,\ \rangle$. Es reicht also zu beweisen, daß $\langle\ ,\ \rangle$ positiv definit ist, wenn nur alle $\det(A_r) > 0$ sind.

Sei also $\det(A_r) > 0$ für alle $1 \leq r \leq \dim(V)$. Wir zeigen per Induktion nach der Dimension $r = \dim U_r$, daß die Einschränkung von $\langle\ ,\ \rangle$ auf U_r positiv definit ist. Für den Induktionsanfang $r = 1$ haben wir $\det(A_1) > 0$, und die ensprechende symmetrische Bilinearform ist positiv definit wie man sofort sieht.

Sei also nach Induktionsannahme $\langle\ ,\ \rangle$ auf U_{r-1} positiv definit. Nach Theorem 6.12 gibt es eine Orthogonalbasis $(c_1, \ldots, c_{r-1})$ von U_{r-1}. Weil $\langle\ ,\ \rangle$ auf U_{r-1} positiv definit ist, folgt für alle $1 \leq i \leq r-1$

$$\langle c_i, c_i\rangle > 0.$$

Wir erweitern mit dem Basisergänzungssatz zunächst zu einer Basis $(c_1, \ldots, c_{r-1}, v)$ von U_r. Das Gram–Schmidt-Verfahren[7] liefert eine Orthogonalbasis $(c_1, \ldots, c_r)$ von U_r durch

$$c_r = v - \sum_{i=1}^{r-1} \frac{\langle v, c_i\rangle}{\langle c_i, c_i\rangle} c_i \in U_{r-1}^{\perp}.$$

Sei C_{r-1} (bzw. C_r) die Gram'sche Matrix von $\langle\ ,\ \rangle$ auf U_{r-1} bezüglich $(c_1, \ldots, c_{r-1})$ (bzw. auf U_r bezüglich $(c_1, \ldots, c_r)$). Dies sind Diagonalmatrizen mit Diagonaleinträgen $\langle c_i, c_i\rangle$. Dann ist

$$\langle c_r, c_r\rangle = \det(C_r)/\det(C_{r-1}) > 0,$$

denn das Vorzeichen ist nach Lemma 7.6 gleich dem von $\det(A_r)/\det(A_{r-1})$. Also ist für $v = \sum_{i=1}^r v_i c_i$ verschieden von 0

$$\langle v, v\rangle = \sum_{i=1}^r v_i^2 \langle c_i, c_i\rangle > 0. \qquad \square$$

Beispiel 7.10. Die Hauptminoren in Beispiel 7.2 (2) sind $\det(A_1) = 3$ und $\det(A_2) = 6047$. Also ist die Bilinearform $\langle\ ,\ \rangle_A$ positiv definit.

Bemerkung 7.11. Der Beweis von Satz 7.9 übersetzt sich in ein Verfahren, wie man eine symmetrische Bilinearform $\langle\ ,\ \rangle$ auf positive Definitheit prüfen kann, und zwar ohne die Hauptminoren zu berechnen. Dazu wendet man das Gram-Schmidt-Orthogonalisierungsverfahren auf irgendeine Basis $\mathscr{B}$ an. Wenn in einem Schritt Division durch 0 nötig wird, dann ist $\langle\ ,\ \rangle$ nicht positiv definit. Andernfalls erhält man eine Orthgonalbasis $\mathscr{C} = (c_1, \ldots, c_n)$ und $\langle\ ,\ \rangle$ ist positiv definit genau dann, wenn alle Diagonaleinträge $\langle c_i, c_i\rangle$ der Gram'schen Matrix > 0 sind.

7.2. **Signatur.** Sei $(V, \langle\ ,\ \rangle)$ ein endlich-dimensionaler $\mathbb{R}$-Vektorraum mit symmetrischer Bilinearform. Zu einer Orthogonalbasis $\mathscr{B}$ definieren wir

$$\begin{aligned} \mathscr{B}_+ &= \{b \in \mathscr{B}\ ;\ \langle b, b\rangle > 0\},\\ \mathscr{B}_- &= \{b \in \mathscr{B}\ ;\ \langle b, b\rangle < 0\},\\ \mathscr{B}_0 &= \{b \in \mathscr{B}\ ;\ \langle b, b\rangle = 0\}, \end{aligned}$$

und für alle $* \in \{+, 0, -\}$

$$n_*(\mathscr{B}) = |\{b \in \mathscr{B}\ ;\ \langle b, b\rangle > 0\}|.$$

Theorem 7.12 (Trägheitssatz von Sylvester)**.** *Sei $(V, \langle\ ,\ \rangle)$ ein endlich-dimensionaler $\mathbb{R}$-Vektorraum mit symmetrischer Bilinearform und Orthogonalbasis $\mathscr{B}$.*

Die Zahlen $n_+(\mathscr{B})$, $n_-(\mathscr{B})$ und $n_0(\mathscr{B})$ sind von der Wahl der Orthogonalbasis $\mathscr{B}$ unabhängig, und zwar gilt mit

$$\begin{aligned} d_+ &:= \max\{\dim_{\mathbb{R}}(W)\ ;\ W \subseteq V \text{ Unterraum und } \langle\ ,\ \rangle|_{W\times W} \text{ ist positiv definit}\},\\ d_- &:= \max\{\dim_{\mathbb{R}}(W)\ ;\ W \subseteq V \text{ Unterraum und } \langle\ ,\ \rangle|_{W\times W} \text{ ist negativ definit}\},\\ d_0 &:= \dim_{\mathbb{R}} V^{\perp}, \end{aligned}$$

genauer $n_+(\mathscr{B}) = d_+$, $n_-(\mathscr{B}) = d_-$, $n_0(\mathscr{B}) = d_0$ und $\dim_{\mathbb{R}}(V) = d_+ + d_- + d_0$.

Beweis. Sei $\mathscr{B}$ eine Orthogonalbasis von V bezüglich $\langle\ ,\ \rangle$. Wir setzen wie oben $\mathscr{B}_*$ und $n_*(\mathscr{B})$ für $* \in \{+, -, 0\}$ und definieren weiter

$$W_* = \text{ lineare Hülle von } \mathscr{B}_* = \langle \mathscr{B}_*\rangle_{\mathbb{R}}.$$

Dann ist $\langle\ ,\ \rangle$ eingeschränkt auf W_+ (bzw. W_-) positiv (bzw. negativ) definit nach dem Hauptminorenkriterium Satz 7.9 und damit

$$n_\pm(\mathscr{B}) = \dim_{\mathbb{R}}(W_\pm) \le d_\pm. \tag{7.1}$$

[7] Wir wissen zwar nicht, daß $\langle\ ,\ \rangle$ anisotrop ist, haben aber die c_i so gewählt, daß das Gram–Schmidt Verfahren keinen Ärger macht: $\langle c_i, c_i\rangle \neq 0$.

Außerdem ist $W_0 \subseteq V^\perp$ und so

$$n_0(\mathscr{B}) \leq d_0. \tag{7.2}$$

Insgesamt gilt

$$\dim(V) = n_+(\mathscr{B}) + n_-(\mathscr{B}) + n_0(\mathscr{B}) \leq d_+ + d_- + d_0.$$

Sei U_+ (bzw. U_-) ein Unterraum maximaler Dimension d_+ (bzw. d_-) auf dem $\langle\ ,\ \rangle$ positiv (bzw. negativ) definit ist. Dann ist $\langle\ ,\ \rangle$ auf $U_{\leq 0} = U_- + V^\perp$ (bzw. $U_{\geq 0} = U_+ + V^\perp$) positiv (bzw. negativ) semi-definit. Sei beispielsweise $v = u + w$ mit $u \in U_+$ und $w \in V^\perp$, dann gilt

$$\langle v, v\rangle = \langle u + w, u + w\rangle = \langle u, u\rangle \geq 0.$$

Für $v \in U_{\leq 0}$ argumentiert man analog. Somit gilt

$$\begin{aligned} V^\perp \cap U_+ \subseteq U_{\leq 0} \cap U_+ = (0),\\ V^\perp \cap U_- \subseteq U_{\geq 0} \cap U_- = (0), \end{aligned}$$

da $\langle\ ,\ \rangle$ auf dem Schnitt sowohl positiv definit als auch negativ semi-definit (bzw. negativ definit und positiv semi-definit) sein muß. Das geht nur auf dem Null-Vektorraum. Dann können wir nach unten abschätzen

$$\begin{aligned} \dim_{\mathbb{R}}(V) &\geq \dim_{\mathbb{R}}(U_{\leq 0} + U_+)\\ &= \dim_{\mathbb{R}}(U_{\leq 0}) + \dim_{\mathbb{R}}(U_+) - \dim(U_{\leq 0} \cap U_+) = \dim_{\mathbb{R}}(U_- + V^\perp) + \dim_{\mathbb{R}}(U_+)\\ &= \dim_{\mathbb{R}}(U_-) + \dim(V^\perp) - \dim_{\mathbb{R}}(U_- \cap V^\perp) + \dim_{\mathbb{R}}(U_+)\\ &= \dim_{\mathbb{R}}(U_-) + \dim(V^\perp) + \dim_{\mathbb{R}}(U_+) = d_- + d_0 + d_+. \end{aligned}$$

Dies ist nur möglich, wenn in den Ungleichungen (7.1) und (7.2) in Gleichheit gilt.

Die Werte $n_+(\mathscr{B})$, $n_-(\mathscr{B})$ und $n_0(\mathscr{B})$ sind von der Wahl der Orthogonalbasis $\mathscr{B}$ unabhängig, weil die Werte d_+, d_- und d_0 ohne Bezug auf $\mathscr{B}$ allein aus $(V, \langle\ ,\ \rangle)$ heraus definiert sind. □

Symmetrische Bilinearformen auf reellen endlich-dimensionalen Vektorräumen werden über die Signatur klassifiziert.

Definition 7.13. Die **Signatur** einer symmetrischen Bilinearform $\langle\ ,\ \rangle$ auf einem endlichdimensionalen $\mathbb{R}$-Vektorraum V ist das Tupel

$$(d_+, d_-, d_0)$$

aus Theorem 7.12. Im nichtausgearteten Fall, also wenn $d_0 = 0$, wird auch das verkürzte Tupel

$$(d_+, d_-)$$

Signatur genannt. In diesem Fall heißt $\tau(V) = d_+ - d_-$ der **Index** von $\langle\ ,\ \rangle$.

Bemerkung 7.14. Sei $(V, \langle\ ,\ \rangle)$ ein endlich-dimensionaler $\mathbb{R}$-Vektorraum mit symmetrischer Bilinearform $\langle\ ,\ \rangle$ von Signatur (d_+, d_-, d_0). Dann hat die Gram'sche Matrix A (bezüglich irgendeiner Basis) den Rang $\mathrm{rg}(A) = d_+ + d_-$.

Beispiel 7.15. Die Spezielle Relativitätstheorie beschreibt die Raumzeit als einen $\mathbb{R}^4$ mit 3 Raumkoordinaten und einer Zeitkordinate. Abstände werden mit der 'Minkowski-Metrik' bestimmt. Dies ist eine perfekte symmetrische Bilinearform $\eta(u, v) = u^t M v$ der Signatur $(3, 1)$, welche in der Standardbasis aus zuerst drei rechtwinkligen Raumrichtungen und dann einer Zeitrichtung durch die Matrix

$$M = \begin{pmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & -c^2 \end{pmatrix}$$

beschrieben wird. Dabei ist c die Lichtgeschwindigkeit. Sei t der globale Zeitparameter[8] und sei zum Zeitpunkt $t = 0$ ein Photon im Ort (x_0, y_0, z_0, t_0), das sich im Raum in Richtung

$$v = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

bewegt. Wir verlangen

$$c = \sqrt{x^2 + y^2 + z^2},$$

damit sich das Photon mit Lichtgeschwindigkeit ausbreitet, wenn es sich auf der Bahn

$$\varphi(t) = (xt + x_0, yt + y_0, zt + z_0, t + t_0)$$

bewegt. Die Differenz zweier Raumzeitpositionen des Photons ist dann isotrop. Der Einfachheit halber nehmen wir $t_2 = t$ und $t_1 = 0$:

$$\eta(\varphi(t) - \varphi(0), \varphi(t) - \varphi(0)) = \begin{pmatrix} xt \\ yt \\ zt \\ t \end{pmatrix}^t \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -c^2 \end{pmatrix} \begin{pmatrix} xt \\ yt \\ zt \\ t \end{pmatrix} = t^2(x^2 + y^2 + z^2 - c^2) = 0.$$

Die Menge der Punkte, die von einem Punkt im $\mathbb{R}^4$ in einer isotropen Richtung liegen, bilden den Lichtkegel. Auf diesen sind die Bahnen der Photonen beschränkt.

Definition 7.16. Sei K ein Körper. Seien $(V, \langle\ ,\ \rangle)$ und $(W, \langle\ ,\ \rangle)$ K-Vektorräume mit symmetrischen Bilinearformen.

(1) Eine lineare Abbildung $f : V \to W$ heißt **isometrische Abbildung**, wenn für alle $v_1, v_2 \in V$ gilt
$$\langle f(v_1), f(v_2)\rangle = \langle v_1, v_2\rangle.$$

(2) $(V, \langle\ ,\ \rangle)$ und $(W, \langle\ ,\ \rangle)$ heißen **isometrisch**, wenn es wechselseitige inverse isometrische Abbildungen $f : V \to W$ und $f^{-1} : W \to V$ gibt. Das ist äquivalent dazu, daß es eine bijektive isometrische Abbildung $f : V \to W$ gibt.

Beispiel 7.17. Nach Satz 6.49 ist für eine orthogonale Matrix $A \in \mathrm{O}_n(K)$ die Matrixmultiplikation $L_A : K^n \to K^n$ eine bijektive isometrische Abbildung.

Satz 7.18. *Zwei $\mathbb{R}$-Vektorräume mit symmetrischer Bilinearform sind isometrisch genau dann, wenn sie die gleiche Signatur haben.*

Beweis. Seien $(V, \langle\ ,\ \rangle)$ und $(W, \langle\ ,\ \rangle)$ zwei $\mathbb{R}$-Vektorräume mit symmetrischer Bilinearform. Wir nehmen zunächst an, daß V und W isometrisch sind. Sei $f : V \to W$ eine invertierbare isometrische Abbidlung, und sei $\mathscr{B} = (b_1, \ldots, b_n)$ eine Orthogonalbasis von V. Dann ist

$$\mathscr{C} = f(\mathscr{B}) = (f(b_1), \ldots, f(b_n))$$

eine Basis von W, denn f ist Vektorraumisomorphismus. Und $\mathscr{C}$ ist auch Orthogonalbasis, weil f isometrisch ist. Wegen $\langle f(b_i), f(b_i)\rangle = \langle b_i, b_i\rangle$ für $i = 1, \ldots n$ gilt für $* \in \{+, 0, -\}$

$$n_*(\mathscr{B}) = n_*(\mathscr{C})$$

und damit haben V und W die gleiche Signatur.

Haben nun V und W beide die Signatur (r, s, t). Sei $\mathscr{B}' = (b'_1, \ldots, b'_n)$ nach Theorem 7.12 eine Orthogonalbasis von V, so daß

$$(b'_i, b'_i) = \lambda_i \begin{cases} > 0 & \text{für } 1 \le i \le r, \\ < 0 & \text{für } r + 1 \le i \le r + s, \\ = 0 & \text{für } r + s + 1 \le i \le r + s + t = n. \end{cases}$$

[8] Die spezielle Relativitätstheorie kennt eine globale Zeitkoordinate. In der allgemeinen Relativitätstheorie gibt es keine global für alle geltende Zeit, sondern nur noch die vom Beobachter in seinem Koordinatensystem ermittelte (Eigen-)Zeit.

Wir skalieren nun für $1 \leq i \leq r+s$

$$b_i = \frac{1}{\sqrt{|\lambda_i|}} b_i'$$

und $b_i = b_i'$ für $i > r+s$. Bezüglich der neuen Basis $\mathscr{B} = (b_1, \ldots, b_n)$ hat $\langle\ ,\ \rangle$ die Gram'sche Matrix A in Blockform der Größe r, s, t (mit $\mathbf{1}_m$ der Einheitsmatrix der entsprechenden Größe m)

$$A = \begin{pmatrix} \mathbf{1}_r & 0 & 0 \\ 0 & -\mathbf{1}_s & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Sei $\mathscr{C} = (c_1, \ldots, c_n)$ eine entsprechende Basis von W. Dann wird durch

$$f(b_i) := c_i$$

eine $\mathbb{R}$-lineare Abbildung $f : V \to W$ festgelegt. Diese ist invertierbar (Basis geht auf Basis) und isometrisch, denn für Gleichheit der Bilinearform $\langle\ ,\ \rangle$ auf V mit der Bilinearform $\langle f(-), f(-)\rangle$ auf V reicht nach Proposition 4.4 Gleichheit auf Paaren von Basisvektoren, und das ist durch die gleiche Gram'sche Matrix gegeben. $\square$

Definition 7.19. Ein **euklidischer Vektorraum** ist ein endlich-dimensionaler $\mathbb{R}$-Vektorraum V zusammen mit einer positiv definiten, symmetrischen Bilinearform $\langle\ ,\ \rangle$. Wir bezeichnen das Paar $(V, \langle\ ,\ \rangle)$ auch oft der Einfachheit halber nur mit V.

Korollar 7.20. *Alle euklidischen Vektorräume der gleichen Dimension sind isometrisch.*

Beweis. Dies ist der Fall der Signatur $(n, 0)$ bzw. genauer $(n, 0, 0)$ von Satz 7.18. $\square$

Bemerkung 7.21. Wenn zukünftig von einem euklidischen Vektorraum V die Rede sein wird, dann ist es legitim, sich darunter den $\mathbb{R}^n$ mit dem Standardskalarprodukt vorzustellen, wobei $n = \dim_{\mathbb{R}}(V)$, denn V ist isometrisch zu $\mathbb{R}^n$ und Isometrien erhalten alles was durch die lineare Struktur und das Skalarprodukt ausgedrückt werden kann.

Bemerkung 7.22. Die Menge der Skalarprodukte auf einem gegebenen endlich-dimensionalen $\mathbb{R}$-Vektorraum bilden keinen Untervektorraum der (symmetrischen) Bilinearformen. Zum Beispiel ist das Negative eines Skalarprodukts negativ definit und nicht positiv definit. Aber mit positiv definiten symmetrischen Bilinearformen f_1, f_2 und $\lambda, \mu \geq 0$ und nicht beide 0, ist auch

$$f = \lambda f_1 + \mu f_2$$

positiv definit. Eine Teilmenge eines $\mathbb{R}$-Vektorraums, die abgeschlossen ist unter solchen Positivlinearkombinationen, nennt man einen Kegel.

Übungsaufgaben zu §7

Übungsaufgabe 7.1. Zeigen Sie direkt ohne das Hauptminorenkriterium, daß ein Eigenwert einer positiv (bzw. negativ) definiten Matrix A positiv (bzw. negativ) ist.

8. Metrische Eigenschaften in euklidischen Räumen

Euklidische Geometrie ist eine axiomatisch definierte Geometrie, modern fundiert über die Hilbertschen Axiome der euklidischen Geometrie. Dabei modelliert euklidische Geometrie im 2- oder 3-dimensionalen Raum den von uns beobachteten wirklichen Raum, zumindest wie er sich klassisch ohne relativistische Korrekturen darstellt, in den Konzepten

Raum:	Punkt, Gerade, Ebene.
Metrisch:	Abstand, Winkel, Volumen.
Dynamisch:	Orientierung, Bewegungen, relative Lage, Kongruenz.

Hier stoßen wir an philosophische Grenzen, die wir aber getrost für die Belange der Vorlesung ignorieren, wenn wir uns auf den Standpunkt stellen, nur das Modell selbst verstehen zu wollen.

Metrische Begriffe dienen unter anderem dem Messen von Längen und Winkeln. Diese geometrischen Begriffe aus dem uns umgebenden 3D-Anschauungsraum oder der 2D-Anschauungsebene, für die unsere Modelle $\mathbb{A}^3(\mathbb{R})$ und $\mathbb{A}^2(\mathbb{R})$ sind, werden im Folgenden abstrakt eingeführt. Ein Plausibilitätsvergleich mit der Wirklichkeit kann nur in dem Maße erfolgen, wie wir uns über die zu modellierenden Eigenschaften Rechenschaft ablegen.

8.1. **Die euklidische Norm.** Wir beginnen mit dem Satz des Pythagoras im $\mathbb{A}^2(\mathbb{R})$. Sei z der Abstand der Punkte[9] $\binom{x}{y}$ und $\binom{0}{0} \in \mathbb{A}^2(\mathbb{R}) = \mathbb{R}^2$. Dann zeigt eine elementargeometrische Überlegung, daß

$$z^2 = x^2 + y^2.$$

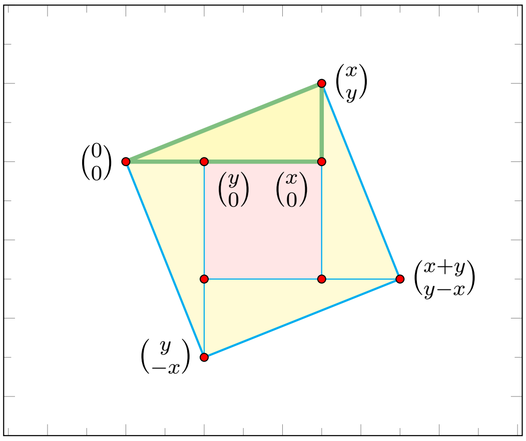


Abbildung 8. Geometrischer Beweis des Satz des Pythagoras.

Die Fläche des großen Quadrats z^2 entspricht der Fläche des kleinen Quadrats $(x-y)^2$ und viermal der Fläche des gelben Dreiecks $xy/2$:

$$z^2 = (x-y)^2 + 4(xy/2) = x^2 + y^2.$$

Hier benutzen wir Addition von Flächeninhalten bei disjunkter Überdeckung, sowie die Flächeninhaltsformel für Quadrate.

Dies übertragen wir nun auf den $\mathbb{A}^3(\mathbb{R}) = \mathbb{R}^3$ und die Punkte

$$O = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \; P = \begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix}, \; Q = \begin{pmatrix} a \\ b \\ 0 \end{pmatrix}, \; R = \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

Das Quadrat des Abstands von O und Q berechnen wir in der Ebene gegeben durch die Gleichung $z = 0$ zu

$$a^2 + b^2.$$

Da die Punkte O, Q und R ebenfalls ein rechtwinkliges Dreieck bilden, jetzt in der Ebene gegeben durch die Gleichung $bx - ay = 0$, folgt für das Quadrat des Abstands von R zu O

$$(a^2 + b^2) + c^2.$$

[9]Ich habe hier absichtlich nicht die Buchstaben a, b und c gewählt, weil es darauf nicht ankommt.

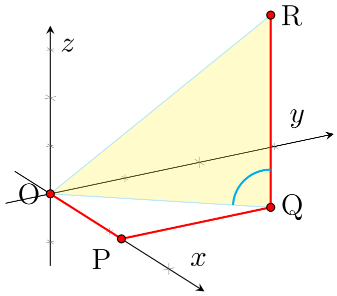


Abbildung 9. Abstand R zu O im $\mathbb{A}^3(\mathbb{R})$.

Diese elementargeometrischen Überlegungen, die an unsere Anschauung appellieren und daher nicht der axiomatischen Methode entsprechen motivieren die folgenden Definitionen.

Definition 8.1. Sei $(V, \langle\ ,\ \rangle)$ ein euklidischer Vektorraum.

(1) Die **Norm** (oder **Länge**) eines Vektors $v \in V$ ist die reelle Zahl
$$\|v\| = \sqrt{\langle v, v\rangle}.$$
Hier bezeichne $\sqrt{x}$ die nicht-negative Wurzel in $\mathbb{R}$ von $x \geq 0$.

(2) Ein **normierter** Vektor in V ist ein Vektor $v \in V$, so daß $\|v\| = 1$.

(3) Der Abstand zweier Vektoren $v, w \in V$ ist
$$d(v, w) = \|v - w\|.$$

(4) Sei A ein affiner Raum mit Translationen durch V. Dann ist der Abstand der Punkte $P, Q \in A$ definiert als
$$d(P, Q) = \|v\|$$
wenn $v \in V$ der eindeutige Vektor ist mit $v + Q = P$.

Bemerkung 8.2. (1) Für den $\mathbb{R}^n$ mit dem Standardskalarprodukt $\langle x, y\rangle = x^t y$ ist mit Koordinaten $x^t = (x_1, \ldots, x_n)$ die Norm durch die folgende Formel gegeben:
$$\|x\| = \sqrt{x_1^2 + \ldots + x_n^2}.$$

(2) Es ist wichtig, daß wir mit einem Skalarprodukt $\langle\ ,\ \rangle$ arbeiten, weil nur dann $\langle v, v\rangle \geq 0$ ist und eine Wurzel in $\mathbb{R}$ existiert.

(3) Daß wir die Norm $\|v\|$ als Länge ansprechen, liegt an der Formel im affinen Raum $\mathbb{A}(V)$:
$$d(v, 0) = \|v\|.$$

(4) Für den affinene Raum $A = \mathbb{A}(V) = V$ stimmen beide Definitionen des Abstands überein, denn
$$(v - w) + w = v.$$

(5) Im $\mathbb{A}^2(\mathbb{R})$ und im $\mathbb{A}^3(\mathbb{R})$ paßt unsere Definition der Länge mit der aus der Anschauung über den Satz des Pythagoras hergeleiteten Länge zusammen.

Proposition 8.3. *Sei V ein euklidischer Vektorraum.*

(1) Für alle $v \in V$ gilt $\|v\| \geq 0$ und
$$\|v\| = 0 \iff v = 0.$$

(2) Für alle $v \in V$ und $\lambda \in \mathbb{R}$ gilt
$$\|\lambda v\| = |\lambda| \cdot \|v\|.$$

Beweis. (1) Es ist $\langle v, v\rangle \geq 0$ und hat daher eine eindeutige nicht-negative Wurzel, weil $\langle\ ,\ \rangle$ ein Skalarprodukt ist. Positive Definitheit zeigt

$$v = 0 \iff \langle v, v\rangle = 0 \iff \|v\| = 0.$$

(2) Wir ziehen die (nicht-negative) Wurzel aus

$$\|\lambda v\|^2 = \langle \lambda v, \lambda v\rangle = \lambda^2\langle v, v\rangle = \lambda^2\|v\|^2.$$ □

Beispiel 8.4. Jedes $v \in V$, $v \neq 0$ hat genau zwei normierte Vielfache. Die Gleichung $\|\lambda v\| = 1$ ist äquivalent zu $|\lambda| \cdot \|v\| = 1$. Gesucht sind die normierten Vektoren

$$\pm\|v\|^{-1} \cdot v.$$

Dieser Trick kam schon im Beweis von Satz 7.18 vor.

Satz 8.5 (Cauchy–Schwarz Ungleichung)**.** *Sei V ein euklidischer Vektorraum. Für alle $v, w \in V$ gilt*

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

Gleichheit gilt genau dann, wenn v und w linear abhängig sind.

Beweis. Wenn v und w linear abhängig sind, also OBdA $w = \lambda v$ mit $\lambda \in K$, dann gilt

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

Seien nun v und w linear unabhängig. Eingeschränkt auf den 2-dimensionalen Unterraum der linearen Hülle von v, w ist $\langle\ ,\ \rangle$ positiv definit, hat also nach Lemma 7.7 bezüglich der Basis v, w eine Gram'sche Matrix mit positiver Determinante:

$$0 < \det\begin{pmatrix} \langle v, v\rangle & \langle v, w\rangle \\ \langle w, v\rangle & \langle w, w\rangle \end{pmatrix} = \|v\|^2\|w\|^2 - \langle v, w\rangle^2.$$

Dies zeigt sofort die Cauchy–Schwarz Ungleichung durch Umstellen und Wurzelziehen. □

Satz 8.6 (Dreiecksungleichung)**.** *Sei V ein euklidischer Vektorraum. Für alle $x, y \in V$ gilt*

$$\|x + y\| \leq \|x\| + \|y\|.$$

Beweis. Da beide Seiten nicht-negativ sind, dürfen wir die Ungleichung quadrieren. Wir haben zu zeigen:

$$\langle x + y, x + y\rangle \leq \langle x, x\rangle + 2\|x\|\|y\| + \langle y, y\rangle.$$

Dies ist nach der Polarisationsformel aus Satz 5.16 äquivalent zu

$$2\langle x, y\rangle \leq 2\|x\|\|y\|,$$

was aus der Cauchy–Schwarz Ungleichung folgt. □

Satz 8.7 (Eigenschaften des Abstands)**.** *Sei A ein affiner Raum mit Vektorraum V der Translationen. Der Abstand hat die folgenden Eigenschaften:*

(1) ***Positiv definit****: für alle $P, Q \in A$ gilt $d(P, Q) \geq 0$, und*

$$d(P, Q) = 0 \iff P = Q.$$

(2) ***Symmetrie****: für alle $P, Q \in A$ ist*

$$d(P, Q) = d(Q, P).$$

(3) ***Dreiecksungleichung****: für alle P, Q und $R \in A$ gilt*

$$d(P, Q) \leq d(P, R) + d(R, Q).$$

Beweis. (1) Sei $v \in V$ der eindeutige Vektor mit $v + Q = P$. Dann ist $d(P,Q) = \|v\| \geq 0$. Und genauer

$$d(P,Q) = 0 \iff \|v\| = 0 \iff v = 0 \iff P = Q.$$

(2) Weiter ist dann $-v + P = -v + (v + Q) = (-v + v) + Q = Q$ und daher

$$d(P,Q) = \|v\| = |-1|\|-v\| = \|-v\| = d(Q,P).$$

(3) Seien nun $x, y \in V$ die eindeutigen Vektoren mit $x + R = P$ und $y + Q = R$. Dann ist $(x + y) + Q = x + (y + Q) = x + R = P$. Damit folgt die Dreiecksungleichung für den Abstand aus der Dreiecksungleichung für die Norm aus Satz 8.6:

$$d(P,Q) = \|x + y\| \leq \|x\| + \|y\| = d(P,R) + d(R,Q).$$ ☐

Bemerkung 8.8. Eine Funktion $d : X \times X \to \mathbb{R}$ auf einer Menge X mit den Eigenschaften von Satz 8.7 nennt man eine Metrik auf X. Das ist ein wichtiger Begriff für die Analysis.

8.2. **Winkel und Orthogonalität.** Wir erinnern, daß zwei Vektoren x, y in einem euklidischen Vektorraum per Definition orthogonal sind, wenn

$$x \perp y :\iff \langle x, y \rangle = 0.$$

Proposition 8.9 (Pythagoras). *Sei V ein euklidischer Vektorraum und seien $x, y \in V$. Dann sind äquivalent:*

(a) $x \perp y$.

(b) Es gilt der Satz des Pythagoras im Dreieck mit den Ecken 0, x und y:

$$d(x,y)^2 = d(x,0)^2 + d(y,0)^2.$$

Beweis. Nach der Polarisationsformel, Satz 5.16, ist

$$2\langle x, y \rangle = \langle x + y, x + y \rangle - \langle x, x \rangle - \langle y, y \rangle.$$

Die linke Seite verschwindet genau dann, wenn $x \perp y$, und die rechte Seite hat den Wert 0 genau dann, wenn der Satz des Pythagoras gilt. ☐

Wir führen nun allgemeiner ein Maß für den Winkel zwischen Vektoren ein.

Definition 8.10. Sei $(V, \langle\ ,\ \rangle)$ ein euklidischer Vektorraum. Der **(ungerichtete) Winkel** zwischen zwei Vektoren $v, w \in V$ mit $v \neq 0$, $w \neq 0$ ist $\measuredangle(v,w) \in [0, \pi]$ mit

$$\cos(\measuredangle(v,w)) := \frac{\langle v, w \rangle}{\|v\| \cdot \|w\|}.$$

Der Winkel ist wohldefiniert, denn $\langle v, w \rangle / (\|v\| \cdot \|w\|)$ liegt nach der Cauchy–Schwarz Ungleichung im Intervall $[-1, 1]$ und die Kosinusfunktion ist bijektiv als Funktion

$$\cos : [0, \pi] \to [-1, 1].$$

Bemerkung 8.11. In einem euklidischen Vektorraum gilt für Vektoren $v, w \in V$, $v \neq 0$, $w \neq 0$:

$$v \perp w \iff \langle v, w \rangle \iff \measuredangle(v,w) = \pi/2,$$

weil für $\varphi \in [0, \pi]$ der Wert $\varphi = \pi/2$ die einzige Nullstelle der Funktion $\cos(\varphi)$ ist.

In der euklidischen Geometrie sollen Vektoren $v \neq 0$ und $w \neq 0$ orthogonal sein, wenn die folgenden Winkel gleich sind:

$$\measuredangle(v,w) = \measuredangle(v,-w).$$

Mit unserer Definition des Winkels ist dies äquivalent zu

$$\frac{\langle v, w \rangle}{\|v\| \cdot \|w\|} = \frac{\langle v, -w \rangle}{\|v\| \cdot \|-w\|},$$

was wiederum zu $(v, w) = 0$ äquivalent ist. Unsere Definition von Winkel und orthogonal sind insoweit mit den geometrischen Modellforderungen konsistent.

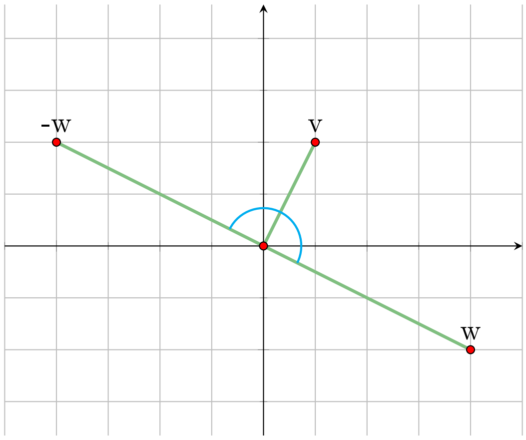
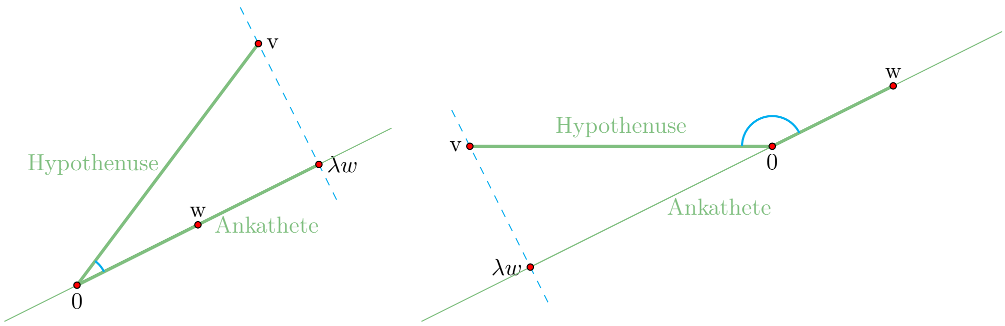


Abbildung 10. Orthogonale Vektoren: $\measuredangle(v,w) = \measuredangle(v,-w)$.

Bemerkung 8.12. (1) Seien $v \neq 0$ und $w \neq 0$ Vektoren eines euklidischen Raumes. Der Kosinus des Winkels $\measuredangle(v,w)$ wird elementargeometrisch als Quotient von Ankathete durch Hypothenuse am Winkel $\measuredangle(v,w)$ in einem rechtwinkligen Dreieck definiert.



(a) 1. Fall

(b) 2.Fall

Abbildung 11. Kosinus des Winkels als Ankathete/Hypothenuse.

Wir suchen auf der Geraden L_w durch 0 und w den Punkt λw, so daß die Gerade durch v und λw orthogonal zu L_w ist. Dies führt zu

$$v - \lambda w \perp w \iff \langle v - \lambda w, w\rangle = 0 \iff \lambda = \frac{\langle v,w\rangle}{\langle w,w\rangle}.$$

Hier tritt nun eine Schwierigkeit auf. Nur für $\lambda \geq 0$ liegt λw auf der gleichen Seite der Geraden L_w wie w (von 0 aus betrachtet). Dann gilt: Das Dreieck $0, v, \lambda w$ hat bei λw einen rechten Winkel, so daß elementargeometrisch

$$\cos(\measuredangle(v,w)) = \cos(\measuredangle(v,\lambda w)) = \frac{\|\lambda w\|}{\|v\|} = \frac{|\lambda|\|w\|}{\|v\|} = \frac{\lambda\|w\|}{\|v\|} = \frac{\langle v,w\rangle}{\|v\|\|w\|}.$$

Auch hier stimmt die abstrakte Definition mit der elementargeometrischen Definition, die modelliert werden soll, überein.

Falls $\lambda < 0$ liegt λw auf der anderen Seite von $0 \in L_w$ wie w. Dann ist

$$\cos(\measuredangle(v,w)) = -\cos(\measuredangle(v,\lambda w)) = -\frac{\|\lambda w\|}{\|v\|} = \frac{-|\lambda|\|w\|}{\|v\|} = \frac{\lambda\|w\|}{\|v\|} = \frac{\langle v,w\rangle}{\|v\|\|w\|}.$$

Man erkennt einen Vorteil der Definition des (Kosinus des) Winkels über das Skalarprodukt: die Fallunterscheidung entfällt!

(2) Sei V ein euklidischer Vektorraum und $v, w \in V$. Wir wählen einen zweidimensionalen Unterraum $U \subseteq V$ mit $v, w \in U$. Eine Wahl haben wir nur, wenn v, w linear abhängig sind, ansonsten ist $U = \langle v, w \rangle$ die lineare Hülle. Der Vektorraum U wird mit der Einschränkung des Skalarprodukts von V selbst ein euklidischer Vektorraum. Nach Korollar 7.20 gibt es eine isometrische lineare Abbildung $f : U \to \mathbb{R}^2$ zum $\mathbb{R}^2$ mit dem Standardskalarprodukt, die ein Isomorphismus ist. Da Abstand, Länge und Winkel nur vom Skalarprodukt abhängen und dies von f definitionsgemäß erhalten bleibt, können wir Abstand, Länge und Winkel von v und w via $f(v)$ und $f(w)$ im $\mathbb{R}^2$ mit dem Standardskalarprodukt berechnen. Insbesondere reicht es aus, die Begriffe Abstand, Länge und Winkel im $\mathbb{R}^2$ als Modelle der 'wirklichen' ebenen euklidischen Geometrie zu erkennen.

Der Satz des Pythagoras wird durch den Kosinussatz verallgemeinert.

Satz 8.13 (Kosinussatz)**.** *Sei V ein euklidischer Vektorraum, und seine $v, w \in V$, $v \neq 0$, $w \neq 0$. Dann gilt:*

$$\|v - w\|^2 = \|v\|^2 + \|w\|^2 - 2\|v\|\|w\| \cos \measuredangle(v, w).$$

Beweis. Wir rechnen mittels der Polarisationsformel, Satz 5.16,

$$\begin{aligned} \|v - w\|^2 &= \langle v - w, v - w \rangle = \langle v, v \rangle + \langle w, w \rangle - 2\langle v, w \rangle \\ &= \|v\|^2 + \|w\|^2 - 2\|v\|\|w\| \cos \measuredangle(v, w). \end{aligned}$$

□

Beispiel 8.14. Wir realisieren den Tetraeder im 3-dimensionalen affinen Raum

$$A = \{x_1 + x_2 + x_3 + x_4 = 1\} \subseteq \mathbb{A}^4(\mathbb{R})$$

als die konvexe Hülle der Standardbasisvektoren:

$$T = \{x = t_1 e_1 + \ldots + t_4 e_4 \in A \ ; \ 0 \leq t_i \leq 1 \text{ für alle } i = 1, \ldots, 4\}.$$

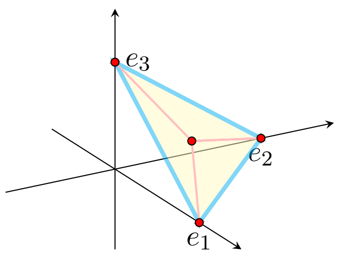


ABBILDUNG 12. 2D-Analogon im $\mathbb{R}^3$ des Tetraeders in $\mathbb{R}^4$.

Diese Beschreibung scheint erst einmal merkwürdig kompliziert, aber Koordinaten der Tetraederecken direkt im $\mathbb{R}^3$ sehen schlimmer aus. Der Translationsraum zu A ist der Unterraum

$$V = \{x_1 + x_2 + x_3 + x_4 = 0\} \subseteq \mathbb{R}^4,$$

der bezüglich der Einschränkung des Standardskalarproduktes des $\mathbb{R}^4$ mit einer positiv definiten, symmetrischen Bilinearform ausgestattet ist: ein euklidischer Vektorraum.

Wir müssen zuerst überlegen, daß die Ecken $e_1, \ldots, e_4$ von T tatsächlich paarweise den gleichen Abstand haben, was T zu einem Tetraeder macht:

$$d(e_i, e_j) = \|e_i - e_j\| = \sqrt{2}.$$

Der entscheidende Vorteil dieser Beschreibung von T ist, daßwir Rechnungen in $\mathbb{R}^4$ mit einfachen Vektoren vornehmen können.

Der Mittelpunkt des Tetraeders ist aus Symmetriegründen gegeben durch den Vektor

$$v_M = \frac{1}{4}(e_1 + e_2 + e_3 + e_4) \in T \subseteq A.$$

In der Tat ist der Abstand

$$d(v_M, e_i) = \|\frac{1}{4}e_1 + \ldots - \frac{3}{4}e_i + \ldots \frac{1}{4}e_4\| = \sqrt{\frac{1}{16} + \ldots + \frac{9}{16} + \ldots + \frac{1}{16}} = \sqrt{\frac{12}{16}} = \frac{1}{2}\sqrt{3}$$

unabhängig von i.

Wir wählen nun v_M als Ursprung für eine Koordinatenbeschreibung von A durch V, also $A = V + v_M$. Der Mittelpunktswinkel im Tetraeder ist nun der Winkel φ, den zwei Ecken vom Mittelpunkt aus einnehmen, also (aus Symmetriegründen ist es egal, welche Ecken wir nehmen)

$$\cos(\varphi) = \frac{\langle e_1 - v_M, e_2 - v_M\rangle}{\|e_1 - v_M\| \cdot \|e_2 - v_M\|} = \frac{\langle \begin{pmatrix} 3/4 \\ -1/4 \\ -1/4 \\ -1/4 \end{pmatrix}, \begin{pmatrix} -1/4 \\ 3/4 \\ -1/4 \\ -1/4 \end{pmatrix} \rangle}{\sqrt{3}/2 \cdot \sqrt{3}/2} = \frac{-4/16}{3/4} = -\frac{1}{3}.$$

Daraus ergibt sich ein Mittelpunktswinkel von ungefähr

$$\varphi \approx 109.471220634491°$$

8.3. **Rechtwinklige Koordinatensysteme.** Ein euklidischer Raum hat lineare Struktur und dazu das Skalarprodukt. Die dem Skalarprodukt angepaßten Basen sind die Orthonormalbasen. Alle Orthonormalbasen sind gleich gut.

Bemerkung 8.15. (1) Wir sprechen aus, was eine ONB, siehe Definition 6.39, in einem in einem euklidischen Vektorraum $(V, \langle\ ,\ \rangle)$ bedeutet. Dies ist eine Basis $\mathscr{B} = (v_1, \ldots, v_n)$, für die gilt:
(i) $v_i \perp v_j$ sind orthogonal für alle $i \neq j$, $1 \leq i, j \leq n$, also $\langle v_i, v_j\rangle = 0$, und
(ii) v_i ist normiert: $\|v_i\| = 1$ für alle $1 \leq i \leq n$.
In einem euklidischen Vektorraum definert eine ONB ein **rechtwinkliges und normiertes Koordinatensystem**.

(2) Es ist konsequent, bei euklidischen Vektorräumen bevorzugt ONBs zu betrachten. Diese sind an die geometrische Struktur des Skalarprodukts angepaßt und werden so der vorhandenen Struktur gerecht. ONB sind für euklidische Vektorräume, was Basen für Vektorräume sind.

(3) Eine Basis $\mathscr{B}$ von V ist eine ONB genau dann, wenn die duale Basis $\mathscr{B}^*$ von V im Sinn von Lemma-Definition 5.19 wieder $\mathscr{B}$ ist, siehe Proposition 6.41. Weiter gilt für alle $v, w \in V$

$$\langle v, w\rangle = \langle \kappa_{\mathscr{B}}(v), \kappa_{\mathscr{B}}(w)\rangle.$$

Man kann das Skalarprodukt mit den Koordinaten in Bezug auf die ONB $\mathscr{B}$ wie mit dem Standardskalarprodukt ausrechnen, siehe Proposition 6.41.

Bemerkung 8.16. Für gewöhnliche Vektorräume benutzt man eine Basis $\mathscr{B}$, um Koordinaten $x = \kappa_{\mathscr{B}}(v)$ einzuführen. Basiswechsel $\mathscr{B} \rightsquigarrow \mathscr{C}$ werden durch Multiplikation $\kappa_{\mathscr{C}}(v) = y = S^{-1}x$ mit einer Basiswechselmatrix $S = \mathrm{M}^{\mathscr{C}}_{\mathscr{B}}(\mathrm{id}) \in \mathrm{GL}_n(K)$ durchgeführt. Für die Darstellungsmatrix $A \in \mathrm{M}_n(K)$ eines Endomorphsimus φ ergibt sich der Basiswechsel

$$A = \mathrm{M}^{\mathscr{B}}_{\mathscr{B}}(\varphi) \rightsquigarrow \mathrm{M}^{\mathscr{C}}_{\mathscr{C}}(\varphi) = S^{-1}AS.$$

Für euklidische Vektorräume nehmen die Orthonormalbasen $\mathscr{B}$ die Rolle ein, die gewöhnliche Basen für Vektorräume spielen. Eine ONB führt rechtwinklige und normierte Koordinaten $x = \kappa_{\mathscr{B}}(v)$ ein, und ein Basiswechsel $\mathscr{B} \rightsquigarrow \mathscr{C}$ in eine weitere ONB wird durch eine orthogonale Matrix $S = \mathrm{M}^{\mathscr{C}}_{\mathscr{B}}(\mathrm{id}) \in \mathrm{O}(n)$ vermittelt, siehe Satz 6.45. Weil S orthogonal ist, wird der Basiswechsel einer Matrix zu einem Endomorphismus φ

$$A = \mathrm{M}^{\mathscr{B}}_{\mathscr{B}}(\varphi) \rightsquigarrow \mathrm{M}^{\mathscr{C}}_{\mathscr{C}}(\varphi) = S^{-1}AS = S^tAS.$$

etwas dadurch vereinfacht, daß sich das Inverse durch Transponieren berechnen läßt. Außerdem sieht das nun aus, wie der Basiswechsel der Gram'schen Matrix einer Bilinearform $f : V \times V \to K$:

$$A = \mathrm{M}^{\mathscr{B},\mathscr{B}}(f) \rightsquigarrow \mathrm{M}^{\mathscr{C},\mathscr{C}}(f) = S^t A S.$$

Die Existenz von Orthonormalbasen in euklidischen Vektorräumen, Satz 7.5, erhalten wir erneut aus dem Gram–Schmidt'schen Ortho**normal**isierungsverfahren.

Satz 8.17 (Gram–Schmidt'sches Orthonormalisierungsverfahren)**.** *Sei $(V, \langle\ ,\ \rangle)$ ein euklidischer Vektorraum, und sei $\mathscr{C} = (c_1, \ldots, c_n)$ eine Basis von V. Wir konstruieren für $i = 0, \ldots, n$ eine Basis $\mathscr{B} = (b_1, \ldots, b_n)$ durch*

$$b_1 = \frac{1}{\|c_1\|} c_1$$

und für $i \geq 2$ durch $b_i = \frac{1}{\|b_i'\|} b_i'$ mit

$$b_i' = c_i - \sum_{j=1}^{i-1} \langle c_i, b_j \rangle b_j.$$

Die Basis $\mathscr{B}$ ist eine Orthonormalbasis von V bezüglich $\langle\ ,\ \rangle$.

Beweis. Dies ist nichts anderes als das Orthogonalisierungsverfahren aus Satz 6.20 mit einem zusätzlichen Normierungsschritt, der b_i' durch ein normiertes Vielfaches. Dadurch verschwindet der Nenner in der Definition der b_i' im Vergleich zu Satz 6.20.

Die Normierung, also die Division durch $\|b_i'\|$, ist möglich, denn aus dem Beweis von Satz 6.20 folgt, daß b_i' Teil einer Basis ist, also $b_i' \neq 0$ und daher $\|b_i'\| \neq 0$.

Wie in Satz 6.20 nachgewiesen, ist $\mathscr{B}$ eine Orthogonalbasis. Da aber alle b_i normiert wurden, handelt es sich um eine ONB. □

Das Gram–Schmidt'sche Orthonormalisierungsverfahren zeigt in Wirklichkeit mehr.

Satz 8.18. *Sei $(V, \langle\ ,\ \rangle)$ ein euklidischer Vektorraum.*

(1) Es gibt Orthonormalbasen für V.
(2) Jedes Orthonormalsystem in V läßt sich zu einer ONB ergänzen.

Beweis. (1) Der Vektorraum V hat eine Basis. Der Algorithmus aus Satz 8.17 transformiert diese in eine ONB, die es damit auch gibt.

(2) Sei $(c_1, \ldots, c_r)$ eine Orthonormalsystem. Dann sind diese Vektoren nach Lemma 6.38 linear unabhängig und können deshalb zu einer Basis $\mathscr{C} = (c_1, \ldots, c_r, c_{r+1}, \ldots, c_n)$ ergänzt werden. Nun wenden wir auf $\mathscr{C}$ den Algorithmus aus Satz 8.17 an. Da $(c_1, \ldots, c_r)$ bereits orthonormal sind, wird dabei (per Induktion zu zeigen) $b_i = c_i$ für $1 \leq i \leq r$, d.h. es passiert zunächst nichts. Der Algorithmus spuckt also eine ONB aus, die das gegebene Orthonormalsystem fortsetzt. □

Proposition 8.19. *Sei $P \in \mathrm{M}_n(\mathbb{R})$ eine quadratische Matrix. Dann sind äquivalent:*

(a) P ist symmetrisch und positiv definit (bzw. semi-definit).
(b) Es gibt eine Matrix $A \in \mathrm{GL}_n(\mathbb{R})$ (bzw. $A \in \mathrm{M}_n(\mathbb{R})$) mit

$$P = A^t A.$$

Beweis. (b) $\Longrightarrow$ (a): Die Matrix $P = A^t A$ ist wegen

$$P^t = (A^t A)^t = A^t (A^t)^t = A^t A = P$$

offensichtlich symmetrisch. Weiter ist für alle $v \in \mathbb{R}^n$

$$\langle v, v \rangle_P = v^t A^t A v = (Av)^t (Av) = \langle Av, Av \rangle \geq 0.$$

Wenn A sogar invertierbar ist, dann folgt für $v \neq 0$ sogar $\langle v, v \rangle_P > 0$.

(a) $\Longrightarrow$ (b): Sei umgekehrt P positiv definit. Die symmetrische Bilinearform $\langle\ ,\ \rangle_P$ besitzt nach Satz 8.18 eine ONB $\mathscr{B}$. Sei $A = \mathrm{M}^{\mathscr{E}}_{\mathscr{B}}(\mathrm{id}) \in \mathrm{GL}_n(\mathbb{R})$ die Basiswechselmatrix von der Standardbasis $\mathscr{E}$ zu $\mathscr{B}$. Dann ist

$$P = \mathrm{M}^{\mathscr{E},\mathscr{E}}(\langle\ ,\ \rangle_P) = A^t\,\mathrm{M}^{\mathscr{B},\mathscr{B}}(\langle\ ,\ \rangle_P)A = A^t\mathbf{1}A = A^tA.$$

Ist P nur positiv semi-definit, dann wählen wir eine Orthogonalbasis $\mathscr{B} = (b_1,\ldots,b_n)$ mit $\langle b_i, b_i\rangle = 1$ für $1 \le i \le r$ und $\langle b_i, b_i\rangle = 0$ für $r+1 \le i \le n$. Es folgt, mit der Matrix in Blockform der Größe $(r, n-r)$

$$E_{r,n-r} := \begin{pmatrix} \mathbf{1}_r & 0 \\ 0 & 0 \end{pmatrix} = E^t_{r,n-r}E_{r,n-r}$$

wie im positiv definiten Fall nur

$$P = A^tE_{r,n-r}A = A^tE^t_{r,n-r}E_{r,n-r}A = B^tB$$

mit $B = E_{r,n-r}A \in \mathrm{M}_n(\mathbb{R})$. $\square$

Bemerkung 8.20. Die Matrix A in Proposition 8.19 kann als obere Dreiecksmatrix gewählt werden. Dazu wählt man wie im Beweis die ONB $\mathscr{B}$ nach dem Gram-Schmidt'schen Orthonormalisierungsverfahren angewandt auf die Standardbasis. Der Basiswechsel hat dann als Basiswechselmatrix $R = A$ eine obere Dreiecksmatrix, vgl. Bemerkung 6.21. Die Zerlegung

$$P = R^tR$$

heißt **Cholesky-Zerlegung** und ist eindeutig, wenn man noch fordert, daß die Diagonaleinträge positiv sind. Die Zerlegung wird auch mit einer unteren Dreiecksmatrix L als $P = LL^t$ geschrieben. Das ist mit $L = R^t$ offensichtlich äquivalent.

8.4. **Die Methode der kleinsten Quadrate.** Die Methode der kleinsten Quadrate ist ein Prinzip zur Bestimmung einer besten Approximation, die auf Carl Friedrich Gauß[10] zurück geht. Genauer ist es die Umsetzung einer Bewertung dessen, was als beste Approximation zu gelten hat.

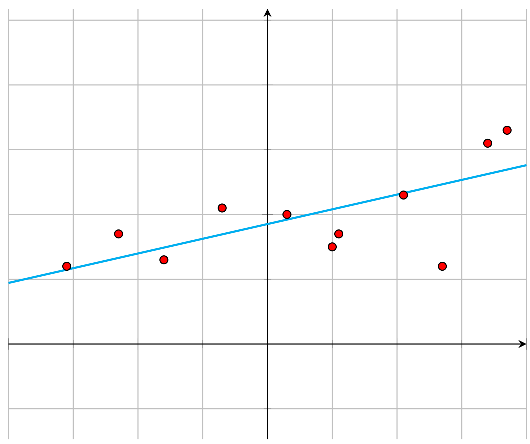


Abbildung 13. Beste Approximation durch eine Gerade.

Beispiel 8.21. Angnommen wir haben einen Datensatz aus Punkten (x_i, y_i) in der Ebene für $i = 1,\ldots,N$ mit N groß. Wir wollen eine Gerade mit Gleichung

$$y = mx + c$$

[10] Carl Friedrich Gauß, 1777–1855, deutscher Mathematiker.

finden, welche die Beziehung zwischen x_i und y_i am besten approximiert. Wir wollen also eigentlich das lineare Gleichungssystem

$$\begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_N & 1 \end{pmatrix} \begin{pmatrix} m \\ c \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix}$$

nach den Variablen m und c lösen. Das ist vermutlich unmöglich, weil die Daten verrauscht sind oder die exakte lineare Abhängigkeit nicht besteht. Es geht um eine approximative Lösung. Allgemeiner haben wir also die Frage, wie man bei einem Gleichungssystem in Matrixform

$$Ax = b,$$

$A \in \mathrm{M}_{n\times m}(\mathbb{R})$, $b \in \mathbb{R}^n$ und x ein Vektor aus Vairablen $x_1, \ldots, x_m$, das keine Lösung hat, eine beste Approximation bestimmt. Die Methode der kleinsten Quadrate schlägt vor, den Fehler $\varepsilon = b - Ax$ zu minimieren. Wir suchen x, so daß

$$\|b - Ax\|^2$$

minimal ist. Im motivierenden Beispiel ist diese Norm

$$(y_1 - (mx_1 + c))^2 + \ldots + (y_N - (mx_N + c))^2,$$

woher sich der Name der **Methode der kleinsten Quadrate** ableitet.

Das Problem hängt zunächst nur vom Bild $Ax \in \mathbb{R}^n$ ab. Wir suchen also einen Vektor im Unterraum $W = \mathrm{im}(A) \subseteq \mathbb{R}^n$, der von b den kleinsten Abstand hat. Dies leistet die orthogonale Projektion.

Proposition 8.22. *Sei $W \subseteq V$ ein Unterraum eines euklidischen Vektorraums V mit orthogonaler Projektion $p_W : V \to V$ auf W. Für alle $b \in V$ ist*

$$\|b - p_W(b)\| = \min_{x \in W} \|b - x\|.$$

Das Minimum wird nur genau in $x = x_0$ angenommen.

Beweis. Der Kern der orthogonalen Projektion ist $\ker(p_W) = W^{\perp}$. Dann ist

$$b - p_W(b) \in \ker(p_W) = W^{\perp}.$$

Wir setzen $x_0 = p_W(b)$ und parametrisieren ein beliebiges $x \in W$ durch $x = x_0 + y$ mit beliebigem $y \in W$. Dann ist $b - x_0 \perp y$ und nach dem Satz des Pythagoras dann

$$\|b - x\|^2 = \|(b - x_0) - y\|^2 = \|b - x_0\|^2 + \|y\|^2 \geq \|b - x_0\|^2$$

mit Gleichheit genau dann, wenn $y = 0$. $\square$

Vor diesem Hintergrund ist es erstrebenswert, die Orthogonale Projektion auf den Spaltenraum $\mathrm{im}(A) \subseteq \mathbb{R}^n$ einer Matrix $A \in \mathrm{M}_{n\times m}$ berechnen zu können.

Methode 1: Der Spaltenraum hat als Erzeugendensystem die Spalten $a_1, \ldots, a_m$ von A. Man wendet auf die Vektoren (e_i ist der i-te Standardbasisvektor von $\mathbb{R}^n$)

$$a_1, \ldots, a_m, e_1, \ldots, e_n$$

das Gram-Schmidt'sche Orthonormalisierungsverfahren an. Wenn in einem Schritt der Nullvektor auftritt, dann war die entsprechende Spalte bereits in der linearen Hülle der Spalten kleineren Index enthalten und wird damit zum Erzeugen des Spaltenraums nicht gebraucht. Im Ergebnis erhalten wir eine ONB von $\mathbb{R}^n$, und zwar genauer sind die Vektoren $b_1, \ldots, b_r$ die aus dem Anfangsstück $a_1, \ldots, a_m$ entstehen eine ONB von

$W = \mathrm{im}(A)$ und der Rest $b_{r+1}, \ldots, b_n$ ist eine ONB von $W^\perp$. Die orthogonale Projektion $p_W : V \to V$ berechnet sich nun durch

$$p_W(v) = \sum_{i=1}^{r} \langle v, b_i \rangle b_i \tag{8.1}$$

gemäß Proposition 6.34, denn eine ONB ist selbstdual, siehe Proposition 6.41.

Weil es gerade paßt, schieben wir die Bessel'sche Ungleichung hier ein, die sofort aus der Formel (8.1) für die orthogonale Projektion folgt.

Proposition 8.23 (Bessel'sche Ungleichung)**.** *Sei $p_W : V \to V$ die orthogonale Projektion auf den Unterraum W eines euklidischen Vektorraums V. Dann gilt für alle $v \in V$:*

$$\|p_W(v)\| \le \|v\|.$$

Beweis. Sei $b_1, \ldots, b_n$ eine ONB von V, so daß $b_1, \ldots, b_r$ eine ONB von W ist. Dann gilt nach Proposition 6.41 und (8.1)

$$\|v\|^2 = \sum_{i=1}^{n} \langle v, b_i \rangle^2 \ge \sum_{i=1}^{r} \langle v, b_i \rangle^2 = \|p_W(v)\|^2. \qquad \square$$

Methode 2: (Spezialfall des Moore–Penrose-Pseudoinversen) Für alle $v \in \mathbb{R}^m$

$$Av = 0 \iff \|Av\|^2 = 0 \iff \langle Av, Av \rangle = 0 \iff \langle A^t Av, v \rangle = 0.$$

Wenn $Av = 0$, dann ist natürlich auch $A^t Av = 0$. Die Kette von Äquivalenzen zeigt auch die Umkehrung:

$$Av = 0 \iff A^t Av = 0.$$

Wir nehmen nun an, daß die Spalten von A linear unabhängig sind, also $\mathrm{rg}(A) = m$ und $\ker(A) = 0$. Dann haben wir gerade gesehen, daß auch $\ker(A^t A) = 0$. Aber $S = A^t A$ ist eine quadratische Matrix, demnach invertierbar $S \in \mathrm{GL}_m(\mathbb{R})$. Wir setzen nun

$$P = AS^{-1}A^t \in \mathrm{M}_n(\mathbb{R})$$

und behaupten, daß P die orthogonale Projektion auf $W = \mathrm{im}(A)$ vermittelt.

Beweis. Dazu rechnen wir

$$P^2 = AS^{-1}A^t AS^{-1}A^t = AS^{-1}SS^{-1}A^t = AS^{-1}A^t = P,$$

also ist P die Matrix einer Projektion. Weil wir bezügich des Standardskalarprodukts und der Standardbasis arbeiten ist P selbstadjungiert:

$$P^* = P^t = (A(A^t A)-1A^t)^t = (A^t)^t((A^t A)^t)^{-1}A^t = A(A^t A)^{-1}A^t = P.$$

Damit ist P die Projektion auf $\mathrm{im}(P)$ bezüglich des Komplements (Satz 6.30)

$$\ker(P) = \mathrm{im}(P^*)^\perp = \mathrm{im}(P)^\perp,$$

also die orthogonale Projektion auf $\mathrm{im}(P)$. Weil A nach Annahmen zu einer injektiven linearen Abbildung gehört und $S \in \mathrm{GL}_m(\mathbb{R})$ folgt $\ker(P) = \ker(A^t)$. Also ist (wieder mit Satz 6.30 und Korollar 6.31)

$$\mathrm{im}(P) = \ker(P^*)^\perp = \ker(P)^\perp = \ker(A^t)^\perp = \mathrm{im}(A). \qquad \square$$

Teile das Arguments stammen aus dem Beweis von Propositon 6.35.

Die nach der Methode der kleinsten Quadrate beste Lösung x_0 zur Gleichung $Ax = b$ löst in Wirklichkeit die Gleichung

$$Ax_0 = Pb.$$

Man kann den Projektor P wieder aus der Gleichung für x_0 eliminieren durch

$$A^t Ax_0 = A^t(Pb) = A^t A(A^t A)^{-1}A^t b = A^t b.$$

Die best Approximation ist somit unter der Rangannahme $\mathrm{rg}(A) = m$ die eindeutige Lösung x_0 der Gleichung
$$(A^t A)x = A^t b.$$
Die Lösung ist eindeutig, weil $S = A^t A \in \mathrm{GL}_m(\mathbb{R})$.

Beispiel 8.24. Beim Approximationsploblem aus Beispiel 8.21 haben wir das 2×2 Gleichungssystem (Summation jeweils $i = 1, \ldots, N$)
$$\begin{pmatrix} \sum x_i^2 & \sum x_i \\ \sum x_i & N \end{pmatrix} \begin{pmatrix} m \\ c \end{pmatrix} = \begin{pmatrix} \sum x_i y_i \\ \sum y_i \end{pmatrix}$$
zu lösen. Mit den Abkürzungen
$$\bar{x} = \frac{1}{N}\sum x_i, \quad q = \frac{1}{N}\sum x_i^2, \quad \bar{y} = \frac{1}{N}\sum y_i, \quad p = \frac{1}{N}\sum x_i y_i$$
ergibt das (z.B. Cramersche Regel)
$$m = \frac{p - \bar{x}\cdot\bar{y}}{q - \bar{x}^2}, \qquad c = \frac{\bar{y}\cdot q - \bar{x}\cdot p}{q - \bar{x}^2}.$$

Im Sinne der Methode der kleinsten Quadrate definieren wir den Abstand eines Punktes von einem affinen Unterraum wie folgt.

Definition 8.25. Sei V ein euklidischer Vektorraum. Der Abstand eines Punktes $v \in V$ von einem nichtleeren affinen Unterraum $B \subseteq V$ ist
$$d(v, B) = \min\{d(v, x) \ ; \ x \in B\}.$$

Proposition 8.26. *Seien V ein euklidischer Vektorraum, $v \in V$ ein Punkt und $B \subseteq V$ ein affiner Unterraum.*

(1) Der Abstand $d(v, B)$ ist wohldefiniert. Das Minimum wird an genau einem Punkt $b \in B$ angenommen.

(2) Sei $b_0 \in B$ und $W \subseteq V$ der Untervektorraum der Translationen von B. Sei $p_W : V \to V$ die orthogonale Projektion auf W. Dann ist
$$b = b_0 + p_W(v - b_0)$$

(3) Es gilt
$$d(v, B) = \|p_{W^\perp}(v - b_0)\|$$

Beweis. Die Abstände sind Translationsinvariant. Nach Translation um b_0 dürfen wir annehmen, daß $B = W$. Dann entsprichen (1) und (2) genau der Behauptung aus Proposition 8.22.

Ausssage (3) folgt aus
$$v - b = v - b_0 - p_W(v - b_0) = p_{W^\perp}(v - b_0).$$
□

Definition 8.27. Eine **affine Hyperebene** im $\mathbb{R}$-Vektorraum V ist ein affiner Unterraum $B = W + b_0$ mit $\dim(W) = \dim(V) - 1$.

Ein **Normalenvektor** an B ist ein Vektor $z \in V$, $z \neq 0$ mit $z \in W^\perp$.

Korollar 8.28. *Wenn speziell $B = W + b_0$ mit $\dim(W) = \dim(V) - 1$, dann ist mit einem normierten Normalenvektor $z \in W^\perp$ der Abstand*
$$d(v, B) = |\langle v - b_0, z\rangle|.$$

Beweis. Der normierte Vektor z ist eine ONB von $W^\perp$, da $\dim(W^\perp) = \dim(V) - \dim(W) = 1$. Die orthogonale Projektion auf $W^\perp$ ist dann gemäß (8.1)
$$p_{W^\perp}(x) = \langle x, z\rangle z$$
und nach Propositon 8.26 folgt
$$d(v, B) = \|p_{W^\perp}(v - b_0)\| = |\langle v - b_0, z\rangle| \cdot \|z\| = |\langle v - b_0, z\rangle|.$$
□

Bemerkung 8.29. Ohne Betragstriche ist

$$\tilde{d}(v,W) = \langle v - b_0, z\rangle = \begin{cases} \langle v - b_0, z\rangle & v \text{ und } z + b_0 \text{ liegen auf der gleichen Seite von } B, \\ -\langle v - b_0, z\rangle & v \text{ und } z + b_0 \text{ liegen auf verschiedenen Seiten von } B. \end{cases}$$

der gerichtete Abstand von v und W.

8.5. **Orientierung.** Im $\mathbb{R}^3$ kann man eine linke Hand von einer rechten Hand unterscheiden. Das dazugehörige Phänomen nennt man Orientierung.

Definition 8.30. Zwei Basen $\mathscr{B} = (b_1, \dots, b_n)$ und $\mathscr{C} = (c_1, \dots, c_n)$ eines $\mathbb{R}$-Vektorraums V heißen **gleich orientiert**, wenn die Basiswechselmatrix $\mathrm{M}^{\mathscr{B}}_{\mathscr{C}}(\mathrm{id})$ eine positive Determinante hat:

$$\det(\mathrm{M}^{\mathscr{B}}_{\mathscr{C}}(\mathrm{id})) > 0.$$

Ansonsten heißen sie **entgegengesetzt orientiert**.

Lemma 8.31. *Gleich orientiert zu sein ist eine Äquivalenzrelation auf Basen eines $\mathbb{R}$-Vektorraums. Es gibt zwei Äquivalenzklassen (außer wenn der Vektorraum der Nullraum ist).*

Beweis. Der Basiswechsel von der Basis $\mathscr{B}$ nach $\mathscr{B}$ ist mittels der Einheitsmatrix $\mathbf{1}$, also ist die Relation reflexiv.

Wenn die Basen $\mathscr{B}$ und $\mathscr{C}$ gleich orientiert sind, und A die Basiswechelmatrix ist, dann ist A^{-1} die Basiswechselmatrix in die umgekehrte Richtung. Wegen

$$\det(A^{-1}) = \det(A)^{-1} > 0$$

sind dann auch $\mathscr{C}$ und $\mathscr{B}$ gleich orientiert. Die Relation ist symmetrisch.

Seien $\mathscr{A}$, $\mathscr{B}$ und $\mathscr{C}$ Basen, und seinen $\mathscr{A}$ und $\mathscr{B}$ sowie $\mathscr{B}$ und $\mathscr{C}$ gleich orientiert. Wenn S die Transfermatrix von $\mathscr{B}$ nach $\mathscr{C}$ und T die Transfermatrix von $\mathscr{A}$ nach $\mathscr{B}$ ist, dann ist ST die Transfermatrix von $\mathscr{A}$ nach $\mathscr{C}$. Wegen

$$\det(ST) = \det(S)\det(T) > 0$$

sind dann auch $\mathscr{A}$ und $\mathscr{C}$ gleich orientiert. Die Relation ist transitiv.

Wir behandeln nun die Anzahl der Äquivalenzklassen. Sei $\mathscr{B}$ eine Basis. Zuerst zeigen wir, daß alle Basen $\mathscr{C}$ und $\mathscr{A}$, die entgegengesetzt zu $\mathscr{B}$ orietntiert sind, zueinander gleich orientiert sind. Mit der Notation aus dem Nachweis der Transitivität gilt nun $\det(S) < 0$ und $\det(T) < 0$, also wieder $\det(ST) > 0$.

Nun brauchen wir noch, daß es überhaupt entgegengesetzt orientierte Basen gibt. Ein Beispiel dazu sind $\mathscr{B} = (b_1, \dots, b_n)$ und $\mathscr{C} = (-b_1, b_2, \dots, b_n)$. □

Bemerkung 8.32. Sei der Einfachheit halber $V = \mathbb{R}^n$. Zwei Basen $\mathscr{B}_0$ und $\mathscr{B}_1$ sollten gleich orientiert sein, wenn man die eine stetig in die andere deformieren kann (Orientierung springt nicht): wenn es eine stetige Funktion

$$\gamma : [0,1] \to \{\text{Basen von } V\}, \quad t \mapsto \mathscr{B}(t)$$

mit $\mathscr{B}(0) = \mathscr{B}_0$ und $\mathscr{B}(1) = \mathscr{B}_1$ gibt. Es stellt sich die Frage, was Stetigkeit hier bedeuten soll. Nun, alle Einträge der Basisvektoren in Bezug auf ein Koordinatensystem sollen stetige Funktionen $[0,1] \to \mathbb{R}$ sein.

Vereinbaren wir, daß für $A \in \mathrm{GL}_n(\mathbb{R})$ und eine Basis $\mathscr{B} = (b_1, \dots, b_n)$ unter $A(\mathscr{B})$ die Basis

$$A(\mathscr{B}) = (Ab_1, \dots, Ab_n)$$

verstehen. Dann ist jede Basis $\mathscr{C}$ mit genau einem A von der Form $A(\mathscr{B})$. Die Abbildung γ kommt daher von einer eindeutigen Abbildung

$$\alpha : [0,1] \to \mathrm{GL}_n(\mathbb{R})$$

mit $\gamma(t) = \mathscr{B}(t) = \alpha(t)(\mathscr{E})$ mit der Standardbasis $\mathscr{E}$. Über die Teilmenge $\mathrm{GL}_n(\mathbb{R}) \subseteq \mathrm{M}_n(\mathbb{R})$ und die Matrixeinträge haben wir nun Koordinaten und einen Begriff von Stetigkeit (aus der Analysis) für die Abbildung α.

Wenn wir Orientierungen diskutieren wollen, interessieren wir uns somit natürlich für die Zusammenhangskomponenten von $\mathrm{GL}_n(\mathbb{R})$: welche Matrizen lassen sich in $\mathrm{GL}_n(\mathbb{R})$ durch einen **Weg** α miteinander verbinden?

Es hat $\mathrm{GL}_n(\mathbb{R})$ zwei Zusammenhangskomponenten, und zwar lassen sich $A, B = SA \in \mathrm{GL}_n(\mathbb{R})$ genau dann verbinden, wenn $\det(S) > 0$ ist. Das Vorzeichen $\det(\alpha(t))$ variiert stetig entlang des Weges α, also gar nicht. Somit ist klar, daß es mindestens die durch das Vorzeichen unterschiedenen Zusammenhangskomponenten gibt. Allerdings läßt sich jede Matrix $A \in \mathrm{GL}_n(\mathbb{R})$ mit $\det(A) > 0$ stetig in die Einheitsmatrix deformieren. Das zeigt mit kurzer Überlegung bereits alles.

Dazu schreiben wir gemäß der Gauß-Elimination zur Zeilenstufenform

$$A = LR$$

mit unterer Dreiecksmatrix L und oberer Dreiecksmatrix R. Weiter zerlegen wir

$$L = S + N, \quad R = T + M$$

mit den Diagonaleinträgen S und T, sowie den Einträgen unterhalb (bzw. oberhalb) der Diagonalen N (bzw. M). Dann ist ein erster Weg

$$\alpha(t) = (S + tN)(T + tM),$$

der die Diagonalmatrix ST mit A verbindet. Sei $ST = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Dann ist ein zweiter zu durchlaufener Weg

$$\beta(t) = \mathrm{diag}(\lambda_1 \cdot |\lambda_1|^{-t}, \ldots, \lambda_n \cdot |\lambda_n|^{-t}),$$

welcher die Diagonaleinträge jeweils zu $\mathrm{sign}(\lambda) \in \{-1, 1\}$ deformiert. Da $\det(A) > 0$ gibt es nun eine gerade Anzahl von Einträgen -1.

Für einen dritten Weg bilden wir Paare von Diagonaleinträgen -1, sagen wir der i und j-te, und bilden die Drehung en um den Winkel $2\pi t$ in $\mathrm{GL}_2(\mathbb{R})$ in der ij-Ebene ab:

$$\gamma_{ij}(t) = \begin{array}{cc} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & \cos(\pi t) & & -\sin(\pi t) & & \\ & & & \ddots & & & \\ & & \sin(\pi t) & & \cos(\pi t) & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} & \begin{matrix} \\ \\ \leftarrow i \\ \\ \leftarrow j \\ \\ \\ \end{matrix} \\ \begin{matrix} \qquad\qquad & \uparrow i & \qquad\qquad & \uparrow j & \qquad\qquad \end{matrix} & \end{array}$$

Der Rest wird mit 1 auf der Diagonale aufgefüllt. Rückwärts durchlaufen und für alle gewählten Paare deformiert dies nach und nach alle -1 Einträge zu 1 und wir landen endlich bei der Einheitsmatrix.

Definition 8.33. Eine **Orientierung** auf einem endlich-dimensionalen $\mathbb{R}$-Vektorraum V ist eine Äquivalenzklasse gleich orientierter Basen. Die Basen in der Orientierung heißen **positiv orientiert**, und die entgegengesetzt orientierten Basen heißen **negativ orientiert**.

Beispiel 8.34. Im $\mathbb{R}^n$ gibt es genau eine Orientierung, die Standardorientierung, bezüglich der die Standardbasis positiv orientiert ist. Für jedes Basis $\mathscr{B} = (b_1, \ldots, b_n)$ gilt dann

$$\mathscr{B} \text{ positiv orientiert} \iff \det([b_1, \ldots, b_n]) > 0.$$

Proposition 8.35. *Sei V ein $\mathbb{R}$-Vektorraum mit einer Orientierung und $\mathscr{B} = (b_1, \ldots, b_n)$ sei eine positiv orientierte Basis. Sei $\sigma \in S_n$ eine Permutation. Dann ist*

$$\sigma(\mathscr{B}) = (b_{\sigma(1)}, \ldots, b_{\sigma(n)}) \text{ positiv orientiert} \iff \mathrm{sign}(\sigma) > 0.$$

Beweis. Die Basiswechselmatrix ist die Permutationsmatrix P_σ zur Permutation σ, und für diese gilt

$$\det(P_\sigma) = \operatorname{sign}(\sigma). \qquad \square$$

Definition 8.36. Eine Matrix $A \in \mathrm{GL}_n(\mathbb{R})$ heißt **orientierungserhaltend**, wenn

$$\det(A) > 0,$$

und **orientierungsumkehrend**, wenn $\det(A) < 0$.

Proposition 8.37. *Die Untergruppe* $\mathrm{SO}(n) \subseteq \mathrm{O}(n)$ *ist die Untergruppe der orientierungserhaltenden orthogonalen Matrizen, und die Menge*

$$\mathrm{GL}_n^+(\mathbb{R}) = \{A \in \mathrm{GL}_n(\mathbb{R}) \ ; \ A \text{ orientierungserhaltend}\}$$

ist eine Untergruppe vom Index 2 *in* $\mathrm{GL}_n(\mathbb{R})$*, der Stabilisator der Standardorientierung auf* $\mathbb{R}^n$*.*

Beweis. Das folgt aus der Definition von $\mathrm{SO}(n)$, und $\mathrm{GL}_n^+(\mathbb{R})$ ist der Kern des surjektiven Gruppenhomomorphismus

$$\mathrm{GL}_n(\mathbb{R}) \xrightarrow{\det} \mathbb{R}^\times \xrightarrow{\operatorname{sign}} \{\pm 1\}. \qquad \square$$

8.6. **Das Volumen eines Parallelotops.** Für Parallelotope wollen wir einen Volumenbegriff bereitstellen.

Definition 8.38. Sei $(V, \langle\ ,\ \rangle)$ ein euklidischer Vektorraum der Dimension $\dim(V) = n$ und seien $v_1, \ldots, v_n \in V$.

(1) Das von $v_1, \ldots, v_n$ aufgespannte **Parallelotop** ist die Teilmenge

$$P(v_1, \ldots, v_n) = \{v \in V \ ; \ v = \sum_{i=1}^n \alpha_i v_i \text{ mit } 0 \le \alpha_i \le 1 \text{ für alle } 1 \le i \le n\}.$$

Das Parallelotop heißt **nicht degeneriert** falls $v_1, \ldots, v_n$ linear unabhängig sind.

(2) Die **Gram-Determinante** des Tupels $(v_1, \ldots, v_n)$ ist die reelle Zahl

$$G(v_1, \ldots, v_n) = \det \begin{pmatrix} \langle v_1, v_1\rangle & \ldots & \langle v_1, v_n\rangle \\ \vdots & \ddots & \vdots \\ \langle v_n, v_1\rangle & \ldots & \langle v_n, v_n\rangle \end{pmatrix} \tag{8.2}$$

Beispiel 8.39. Ein Parallelotop $P = P(v_1, v_2)$ im $\mathbb{R}^2$ beschreibt ein Parallelogramm.

Bemerkung 8.40. Sind $v_1, \ldots, v_n$ linear abhängig

$$x_1 v_1 + \ldots + x_n v_n = 0,$$

so sind die Spalten der Matrix in (8.2) linear abhängig

$$x_1 \begin{pmatrix} \langle v_1, v_1\rangle \\ \vdots \\ \langle v_n, v_1\rangle \end{pmatrix} + \ldots + x_n \begin{pmatrix} \langle v_1, v_n\rangle \\ \vdots \\ \langle v_n, v_n\rangle \end{pmatrix} = \begin{pmatrix} \langle v_1, \sum x_i v_i\rangle \\ \vdots \\ \langle v_n, \sum x_i v_i\rangle \end{pmatrix} = 0$$

und $G(v_1, \ldots, v_n) = 0$. Andernfalls handelt es sich um die Determinante der Gram'schen Matrix der Bilinearform $\langle\ ,\ \rangle$ bezüglich der Basis $(v_1, \ldots, v_n)$. Diese ist nach Lemma 7.7 positiv. Damit ist das Volumen in der folgenden Definition wohldefiniert.

Definition 8.41. Sei $(V, \langle\ ,\ \rangle)$ ein euklidischer Vektorraum der Dimension $n = \dim(V)$. Das **Volumen** eines Parallelotops $P(v_1, \ldots, v_n)$ ist definiert als die nicht-negative reelle Zahl

$$\operatorname{vol}(P(v_1, \ldots, v_n)) = \sqrt{G(v_1, \ldots, v_n)}.$$

Das ist wohldefiniert, weil $\langle\ ,\ \rangle$ positiv definit und damit die Determinante einer Gram'schen Matrix > 0 ist nach Lemma 7.7.

Beispiel 8.42. Ein Parallelotop $P = P(v_1, v_2)$ im $\mathbb{R}^2$ ist ein Rechteck, wenn $v_1 \perp v_2$. Dann gilt

$$\mathrm{vol}(P) = \sqrt{G(v_1, v_2)} = \sqrt{\det \begin{pmatrix} \|v_1\|^2 & 0 \\ 0 & \|v_2\|^2 \end{pmatrix}} = \|v_1\| \cdot \|v_2\|,$$

wie dies zu erwarten war.

Satz 8.43. *Sei V ein euklidischer Vektorraum der Dimension n, seien $v_1, \ldots, v_n \in V$ und sei $\mathscr{B} = (b_1, \ldots, b_n)$ eine ONB. Seien $a_i = \kappa_{\mathscr{B}}(v_i) \in \mathbb{R}^n$ die Koordinatenvektoren der v_i bezüglich $\mathscr{B}$ und $A = [a_1, \ldots, a_n] \in \mathrm{M}_n(\mathbb{R})$ die Matrix mit den Spalten a_i.*

(1) Es gilt

$$\mathrm{vol}(P(v_1, \ldots, v_n)) = |\det A|.$$

(2) Insbesondere gilt

$$\mathrm{vol}(P(b_1, \ldots, b_n)) = 1.$$

Beweis. Aussage (2) ist der Spezialfall von Aussage (1) für $v_i = b_i$. Es ist dann $a_i = e_i$ und $A = \mathbf{1}$, somit $\det(A) = 1$.

(1) Weil $\mathscr{B}$ eine ONB ist, folgt $v_j = \sum_{i=1}^n \langle v_j, b_i \rangle b_i$ für alle $j = 1, \ldots, n$. Demnach ist

$$a_i = \kappa_{\mathscr{B}}(v_i) = \begin{pmatrix} \langle v_i, b_1 \rangle \\ \vdots \\ \langle v_i, b_n \rangle \end{pmatrix}$$

und weiter

$$A = \begin{pmatrix} \langle v_1, b_1 \rangle & \ldots & \langle v_n, b_1 \rangle \\ \vdots & \ddots & \vdots \\ \langle v_1, b_n \rangle & \ldots & \langle v_n, b_n \rangle \end{pmatrix}.$$

Der ij-te Eintrag von $A^t A$ ist

$$(A^t A)_{ij} = \sum_{k=1}^n \langle v_i, b_k \rangle \langle v_j, b_k \rangle = \langle \kappa_{\mathscr{B}}(v_i), \kappa_{\mathscr{B}}(v_j) \rangle = \langle v_i, v_j \rangle.$$

Dies zeigt, daß $A^t A$ diejenige Matrix aus der Definition der Gram-Determinante ist. Folglich gilt

$$|\det(A)|^2 = \det(A)^2 = \det(A^t)\det(A) = \det(A^t A) = G(v_1, \ldots, v_n).$$

Die Formel für das Volumen folgt unmittelbar daraus. □

Korollar 8.44. *Seien $v_1, \ldots, v_n \in \mathbb{R}^n$ als euklidischer Raum mit dem Standardskalarprodukt, und $A = [v_1, \ldots, v_n] \in \mathrm{M}_n(\mathbb{R})$ die entsprechende Matrix. Dann ist*

$$\mathrm{vol}(P(v_1, \ldots, v_n)) = |\det A|.$$

Beweis. Das ist die Formel aus Satz 8.43 für $V = \mathbb{R}^n$ und der Standardbasis als ONB. □

Korollar 8.45. *Sei V ein euklidischer Vektorraum der Dimension n, und seien $v_1, \ldots, v_n \in V$. Dann ist*

$$\mathrm{vol}(P(v_1, \ldots, v_n)) = 0 \iff v_1, \ldots, v_n \textit{ sind linear abhängig.}$$

Beweis. Mit der Notation aus Satz 8.43 folgt

$$\begin{aligned} \mathrm{vol}(P(v_1, \ldots, v_n)) = 0 \iff \det(A) = 0 &\iff a_1, \ldots, a_n \text{ sind linear abhängig} \\ &\iff v_1, \ldots, v_n \text{ sind linear abhängig.} \end{aligned}$$

□

Satz 8.46. *Sei V ein euklidischer Vektorraum und $\dim(V) = n$. Seien $v_1, \ldots, v_n \in V$ und sei $f \in \mathrm{End}_{\mathbb{R}}(V)$. Dann gilt*

$$\mathrm{vol}(P(f(v_1), \ldots, f(v_n))) = |\det(f)| \cdot \mathrm{vol}(P(v_1, \ldots, v_n)).$$

Beweis. Sei $\mathscr{B}$ eine ONB und $A = [a_1, \ldots, a_n]$ mit $a_i = \kappa_{\mathscr{B}}(v_i)$ die Matrix der Koordinatenvektoren der v_i bezüglich $\mathscr{B}$. Sei $F = \mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f)$. Dann ist $FA = [Fa_1, \ldots, Fa_n]$ die Matrix der Koordinatenvektoren der $f(v_i)$ bezüglich $\mathscr{B}$. Nach Satz 8.43 folgt

$$\mathrm{vol}(P(f(v_1), \ldots, f(v_n))) = |\det(FA)| = |\det(F)| \cdot |\det(A)| = |\det(f)| \cdot \mathrm{vol}(P(v_1, \ldots, v_n)). \qquad \square$$

Bemerkung 8.47. Satz 8.46 zeigt, daß diese Festsetzung des Volumens den anschaulichen Volumenbegriff modelliert. Das Parallelotop zu einer ONB ist ein Würfel der Kantenlänge 1, und demnach anschaulichem Volumen 1. Außerdem muß ein Volumenbegriff multilinear in den aufspannenden Vektoren sein, vorausgesetzt, man definiert das Volumen mit einem Vorzeichen, das die Orientierung, welche durch die Reihenfolge der aufspannenden Vektoren definiert wird, berücksichtigt. Dann entfallen in der Formel von Satz 8.43 die Betragstriche.

Im $\mathbb{R}^n$ gilt etwa: Seien $v_1, \ldots, v_n \in \mathbb{R}^n$, und $A = [v_1, \ldots, v_n] \in \mathrm{M}_n(\mathbb{R})$ die entsprechende Matrix. Dann ist das orientierte Volumen

$$\widetilde{\mathrm{vol}}(P(v_1, \ldots, v_n)) = \det A,$$

und wenn $\mathscr{B} = (v_1, \ldots, v_n)$ eine Basis ist, also $A \in \mathrm{GL}_n(\mathbb{R})$, dann ist

$$\widetilde{\mathrm{vol}}(P(v_1, \ldots, v_n)) = \begin{cases} \mathrm{vol}(P(v_1, \ldots, v_n)) & \text{wenn } \mathscr{B} \text{ positiv orientiert ist, und} \\ -\mathrm{vol}(P(v_1, \ldots, v_n)) & \text{wenn } \mathscr{B} \text{ negativ orientiert ist.} \end{cases}$$

Beispiel 8.48. Wir zeigen in diesem Beispiel, wie das Volumen eines Parallelotops in der Ebene als Volumen des zu einem Rechteck gescherten Parallelotops berechnet werden kann. Insbesondere stimmt daher hier das Volumen mit dem Flächeninhalt im herkömmlichen elementargeometrischen Sinne überein.

(1) Die Abbildung $f : \mathbb{R}^2 \to \mathbb{R}^2$ sei die Matrixmultiplikation mit

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

Dies ist die **Scherung** (oder **Transvektion**) entlang der x-Achse, der Fixgeraden der Scherung, mit dem Scherfaktor λ. Wegen $\det(A) = 1$ ändert sich hierbei das Volumen eines Parallelotops nicht.

(2) Zu jedem Parallelotop $P = P(v_1, v_2)$ im $\mathbb{R}^2$ gibt es eine Scherung, die P in ein Rechteck transformiert. Wenn v_1, v_2 linear abhängig sind, hat man nichts zu tun. Ansonsten scheren wir mit Fixgeraden $L = \langle v_1 \rangle$ und einem noch zu bestimmenden Streckfaktor. Sei dazu

$$u_1 = \frac{1}{\|v_1\|} v_1$$

der normierte Vektor in Richtung v_1 auf L, und sei u_2 normiert und orthogonal zu v_1. Mit andern Worten u_2 ergänzt u_1 zu einer ONB $\mathscr{B} = (u_1, u_2)$ von $\mathbb{R}^2$. In dieser Basis ist die gesuchte Scherung f_A von der Form

$$A = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$$

mit einem noch zu bestimmenden $\lambda \in \mathbb{R}$: der Vektor $v = \langle v, u_1 \rangle u_1 + \langle v, u_2 \rangle u_2$ bildet ab auf

$$f_A(v) = \big(\langle v, u_1 \rangle + \lambda \langle v, u_2 \rangle\big) u_1 + \langle v, u_2 \rangle u_2 = v + \lambda \langle v, u_2 \rangle u_1.$$

Hieraus sieht man, daß die Gerade L fix gelassen wird ($f_A(v_1) = v_1$) und ansonsten Punkte v proportional zum Abstand

$$d(v, L) := \min\{d(v, w) \ ; \ w \in L\} = \langle v, u_2 \rangle$$

mit dem Faktor λ in Richtung u_1 bewegt. Diese Scherung macht aus dem Parallelotop $P = P(v_1, v_2)$ ein Rechteck, wenn gilt

$$0 = \langle f_A(v_1), f_A(v_2) \rangle = \langle v_1, v_2 + \lambda \langle v_2, u_2 \rangle u_1 \rangle = \langle v_1, v_2 \rangle + \lambda \langle v_2, u_2 \rangle \langle v_1, u_1 \rangle.$$

Wir setzen $\varphi = \measuredangle(v_1, v_2)$, so daß

$$v_2 = \langle v_2, u_1\rangle u_1 + \langle v_2, u_2\rangle u_2 = \|v_2\| \cdot (\cos(\varphi)u_1 + \sin(\varphi)u_2)$$

und bestimmen den nötigen Scherungsfaktor als

$$\lambda = -\frac{\langle v_1, v_2\rangle}{\langle v_2, u_2\rangle\langle v_1, u_1\rangle} = -\frac{\|v_1\| \cdot \|v_2\| \cdot \cos(\varphi)}{\|v_2\| \sin(\varphi) \cdot \|v_1\|} = -\frac{\cos(\varphi)}{\sin(\varphi)}.$$

Das Bild $f_A(v_2)$ mit diesem λ ist

$$\begin{aligned} f_A(v_2) &= v_2 + \lambda\langle v_2, u_2\rangle u_1 = v_2 - \frac{\langle v_1, v_2\rangle}{\langle v_2, u_2\rangle\langle v_1, u_1\rangle} \cdot \langle v_2, u_2\rangle u_1 \\ &= v_2 - \|v_2\| \cos(\varphi) \cdot u_1 = \|v_2\| \cdot \sin(\varphi)u_2. \end{aligned}$$

Als Volumen (Fläche) von $P = P(v_1, v_2)$ ergibt sich

$$\begin{aligned} \mathrm{vol}(P) &= \mathrm{vol}(f_A(P)) = \|f_A(v_1)\| \cdot \|f_A(v_2)\| \\ &= \|v_1\| \cdot \| \|v_2\| \cdot \sin(\varphi)u_2\| \\ &= \|v_1\| \cdot \|v_2\| \cdot |\sin(\varphi)|. \end{aligned}$$

Dies ist die Formel für den Flächeninhalt eines Parallelogramms als das Produkt aus der Länge der Grundseite $\|v_1\|$ mal der Länge der Höhe über dieser Grundseite $\|v_2\| \cdot |\sin(\varphi)|$.

Übungsaufgaben zu §8

Übungsaufgabe 8.1. Realisieren Sie einen Tetraeder im $\mathbb{R}^3$ mit dem Standardskalarprodukt durch explizite Angabe der Koordinaten der Ecken. Berechnen Sie dann den Mittelpunktswinkel.

Übungsaufgabe 8.2. Beweisen Sie die Cauchy-Schwarz Ungleichung mit Analysis. Zu Vektoren v, w in einem euklidischen Vektorraum V minimieren Sie die Funktion

$$F : \mathbb{R} \to \mathbb{R}, \quad F(t) = \|v + tw\|^2.$$

Werten Sie dann die Ungleichung $F(t_0) \geq 0$ am Minimum t_0 aus.

Übungsaufgabe 8.3. Im Kontext der Methode der kleinsten Quadrate: sei $A \in \mathrm{M}_n(\mathbb{R})$ und $b \in \mathbb{R}^n$. Berechnen sie für alle $v \in \mathbb{R}^n$ die Richtungsableitung der Funktion „Quadrat des Fehlerterms“

$$x \mapsto \|b - Ax\|^2.$$

Zeigen Sie, daß an der Stelle x_0 alle Richtungsableitungen verschwinden genau dann, wenn

$$b - Ax \perp \mathrm{im}(A).$$

Folgern Sie, daß es ein absolutes Minimum gibt und dies genau für die Lösungen $Ax = Pb$ mit der Matrix P der orthogonalen Projektion auf den Spaltenvektorraum $W = \mathrm{im}(A)$ von A angenommen wird.

Übungsaufgabe 8.4 (Hesse'sche Normalform). Schreiben Sie eine Gleichung für eine Gerade in $\mathbb{A}^2(\mathbb{R}) = \mathbb{R}^2$, eine Ebene in $\mathbb{A}^3(\mathbb{R}) = \mathbb{R}^3$, oder allgemeiner einen affinen Unterraum $B \subseteq \mathbb{A}^n(\mathbb{R})$ der Dimension $n - 1$, indem Sie ausdrücken, daß dies genau die Punkte $x \in \mathbb{A}^n(\mathbb{R}) = \mathbb{R}^n$ sind mit Abstand

$$d(x, B) = 0.$$

9. Bewegungen und Isometrien

9.1. Spiegelungen und Drehungen. Wir betrachten zunächst die zwei grundlegenden Beispiele von Isometrien, siehe Definition 9.10.

Beispiel 9.1. Wir betrachten wie in Beispiel 5.28 die **Drehung** $R_\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ des $\mathbb{R}^2$ um den Winkel φ. Diese ist gegeben bezüglich der Standardbasis durch die Matrix

$$D_\varphi = \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}.$$

Wir statten $\mathbb{R}^2$ mit dem Standardskalarprodukt aus. Dann haben wir bereits berechnet, daß $R_\varphi^* = R_{-\varphi} = R_\varphi^{-1}$ wegen $D_\varphi^t = D_{-\varphi} = D_\varphi^{-1}$: die Drehmatrix ist eine orthogonale Matrix. Für alle $v, w \in \mathbb{R}^2$ gilt dann wie in Satz 6.49

$$\langle R_\varphi(v), R_\varphi(w)\rangle = \langle D_\varphi v, D_\varphi w\rangle = \langle v, w\rangle.$$

und damit speziell für $v = w$ nach Wurzelziehen

$$\|R_\varphi(v)\| = \|D_\varphi v\| = \|v\|.$$

Die Drehung $R_\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ ist also eine isometrische Abbildung. Wir schließen, daß in der Tat v und $R_\varphi(v)$ auf dem Kreis mit dem selben Radius $r = \|v\|$ um den Ursprung liegen.

Proposition 9.2. *Der Winkel zwischen v und $R_\varphi(v)$ ist*

$$\measuredangle(v, R_\varphi(v)) = \begin{cases} \varphi & \text{wenn } 0 \le \varphi \le \pi, \\ 2\pi - \varphi & \text{wenn } \pi \le \varphi \le 2\pi. \end{cases}$$

Beweis. Für $v = \binom{x}{y}$ berechnen wir

$$\cos(\measuredangle(v, R_\varphi(v))) = \frac{\langle v, R_\varphi v\rangle}{\|v\| \cdot \|R_\varphi(v)\|} = \frac{(x,y)\binom{x\cdot\cos(\varphi) - y\cdot\sin(\varphi)}{x\cdot\sin(\varphi) + y\cdot\cos(\varphi)}}{\|v\|^2} = \frac{(x^2+y^2)\cdot\cos(\varphi)}{x^2+y^2} = \cos(\varphi).$$

Die Umkehrung des Cosinus zusammen mit der Konvention, daß Winkel in $[0, \pi]$ liegen, führt zur Behauptung. □

Definition 9.3. Die Aussage von Proposition 9.2 wird einheitlicher, wenn man zu orientierten Winkeln übergeht. Dazu betrachten wir $\mathbb{R}^2$ ausgestattet mit der Standardorientierung, bezüglich der die Standardbasis positiv orientiert ist. Sodann definieren wir zu Vektoren $v, w \in \mathbb{R}^2$, $v, w \neq 0$ den **orientierten Winkel** $-\pi < \widetilde{\measuredangle}(v, w) \le \pi$ durch

$$\widetilde{\measuredangle}(v, w) := \begin{cases} \measuredangle(v, w) & \text{wenn } (v, w) \text{ eine positiv orientierte Basis ist,} \\ -\measuredangle(v, w) & \text{wenn } (v, w) \text{ eine negativ orientierte Basis ist,} \\ 0 & \text{wenn } v = \lambda w \text{ mit } \lambda > 0, \\ \pi & \text{wenn } v = \lambda w \text{ mit } \lambda < 0. \end{cases}$$

Proposition 9.4. *Für alle $v, w \in \mathbb{R}^2$, $v, w \neq 0$ und $\varphi \in \mathbb{R}$ gilt*

(1) $|\widetilde{\measuredangle}(v, w)| = \measuredangle(v, w)$,
(2) $\widetilde{\measuredangle}(v, R_\varphi(v)) \equiv \varphi \pmod{2\pi\mathbb{Z}}$.

Beweis. Aussage (1) ist klar, wenn (v, w) eine Basis von $\mathbb{R}^2$ ist. Wenn $v = \lambda w$, dann ist

$$\cos(\measuredangle(v, w)) = \frac{\langle \lambda w, w\rangle}{\|\lambda w\| \cdot \|w\|} = \frac{\lambda}{|\lambda|} = \text{ Vorzeichen von } \lambda.$$

Bei $\lambda > 0$ ist damit $\measuredangle(v, w) = 0$ und bei $\lambda < 0$ ist damit $\measuredangle(v, w) = \pi$.

Für Aussage (2) müssen wir bestimmen, ob $(v, R_\varphi(v))$ positiv oder negativ orientiert ist, sofern es sich um eine Basis handelt. Dazu setzen wir $v = \binom{x}{y}$ berechnen wir die Determinante der Basiswechselmatrix zur Standardbasis

$$\det\left(\begin{pmatrix} x & x\cos(\varphi) - y\sin(\varphi) \\ y & x\sin(\varphi) + y\cos(\varphi) \end{pmatrix}\right) = (x^2 + y^2) \cdot \sin(\varphi).$$

Für $0 < \varphi < \pi$ ist $(v, R_\varphi(v))$ eine positiv orientierte Basis und

$$\widetilde{\measuredangle}(v, R_\varphi(v)) = \measuredangle(v, R_\varphi(v)) = \varphi$$

nach Proposition 9.2. Für $\pi < \varphi < 2\pi$ ist $(v, R_\varphi(v))$ eine positiv orientierte Basis und

$$\widetilde{\measuredangle}(v, R_\varphi(v)) = -\measuredangle(v, R_\varphi(v)) = -(2\pi - \varphi) = \varphi - 2\pi \equiv \varphi \pmod{2\pi\mathbb{Z}}$$

wieder nach Proposition 9.2. Für $\varphi = 0$ und $\varphi = \pi$ folgt Aussage (2) sofort. Da $\varphi \mapsto R_\varphi(v)$ periodisch in π mit Periode 2π ist, folgt damit die Aussage für alle $\varphi \in \mathbb{R}$. $\square$

Beispiel 9.5. Sei $v \in \mathbb{R}^n$, $v \neq 0$. Die **Spiegelung an der Hyperebene orthogonal zu** v also an $H_v = \langle v\rangle^\perp$ ist die lineare Abbildung $S_v : \mathbb{R}^n \to \mathbb{R}^n$

$$S_v(w) = w - 2\frac{\langle v, w\rangle}{\langle v, v\rangle}v.$$

In der Tat sind alle $w \in H_v$ im Eigenraum zum Eigenwert 1 von S_v, oder einfach gesagt: die Vektoren aus der Spiegelhyperebene H_w bewegen sich nicht:

$$S_v(w) = w - 2\frac{\langle v, w\rangle}{\langle v, v\rangle}v = w.$$

In Richtung von v allerdings gilt

$$S_v(v) = v - 2\frac{\langle v, v\rangle}{\langle v, v\rangle}v = v - 2v = -v.$$

Wir bestimmen nun die Matrix von S_v in einer geeigneten ONB. Offensichtlich gilt für $\lambda \in K^\times$

$$S_{\lambda v}(w) = w - 2\frac{\langle \lambda v, w\rangle}{\langle \lambda v, \lambda v\rangle}\lambda v = w - 2\frac{\langle v, w\rangle}{\langle v, v\rangle}v = S_v(w).$$

Es kommt also nur auf die von v aufgespannte Gerade an. Wir nehmen daher OBdA an, daß v normiert ist, und ergänzen zu einer ONB $\mathscr{B} = (b_1 = v, b_2, \ldots, b_n)$. Dann liegt $b_i \in H_v$ für $i = 2, \ldots, n$ und $S_v(b_i) = b_i$. In dieser Basis nimmt S_v die Blockmatrixgestalt

$$\mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(S_v) = \begin{pmatrix} -1 & 0 \\ 0 & \mathbf{1}_{n-1} \end{pmatrix}$$

an, wobei $\mathbf{1}_{n-1}$ die Einheitsmatrix in $\mathrm{M}_{n-1}(\mathbb{R})$ ist. Diese Matrix ist symmetrisch und bezüglich einer ONB aufgestellt, somit gilt nach Proposition 6.42

$$S_v^* = S_v.$$

Wir rechnen nun für $x, y \in \mathbb{R}^n$ beliebig mittels $S_v \circ S_v = \mathrm{id}$

$$\langle S_v(x), S_v(y)\rangle = \langle x, S_v^* S_v(y)\rangle = \langle x, S_v^2(y)\rangle = \langle x, y\rangle.$$

Die Spiegelung S_v ist somit eine isometrische Abbildung.

Wir bestimmen nun O(2) und SO(2). Wir erinnern zunächst daran, daß eine orthogonale Matrix A Determinante $\det(A) = \pm 1$ besitzt, siehe Proposition 6.47. Außerdem erinnern wir an die Additionstheoreme für Sinus und Cosinus.

Proposition 9.6. *Für alle $\varphi, \psi \in \mathbb{R}$ gilt*

(1) $\cos(\varphi + \psi) = \cos(\varphi)\cos(\psi) - \sin(\varphi)\sin(\psi)$,
(2) $\sin(\varphi + \psi) = \sin(\varphi)\cos(\psi) + \cos(\varphi)\sin(\psi)$.

Beweis. Für alle $|ph \in \mathbb{R}$ gilt die Euler'sche Formel

$$e^{i\varphi} = \cos(\varphi) + i\sin(\varphi)$$

mit der komplexen Exponentialfunktion $e^z : \mathbb{C} \to \mathbb{C}$ definiert durch die Potenzreihe

$$e^z = \sum_{n\geq 0} \frac{z^n}{n!}.$$

Die Funktionalgleichung der Exponentialfunktion ($e^{x+y} = e^x \cdot e^y$ für alle $x, y \in \mathbb{C}$) zeigt

$$\cos(\varphi+\psi) + i\sin(\varphi+\psi) = e^{i(\varphi+\psi)} = e^{i\varphi}\cdot e^{i\psi} = \big(\cos(\varphi)+i\sin(\varphi)\big)\cdot\big(\cos(\psi)+i\sin(\psi)\big).$$

Die Behauptung folgt durch Ausmultiplizieren und Vergleich von Real- und Imaginärteil. □

Satz 9.7. *Im zweidimensionalen ist eine orthogonale Matrix eine Drehung oder eine Spiegelung:*

(1) Die Abbildung $\varphi \mapsto D_\varphi$ induziert einen Gruppenisomorphismus

$$\mathbb{R}/2\pi\mathbb{Z} \to \mathrm{SO}(2).$$

Insbesondere ist jede orthogonale Matrix $D \in \mathrm{O}(2)$ mit $\det(D) = 1$ eine Drehung.

(2) Jede orthogonale Matrix $S \in \mathrm{O}(2)$ mit $\det(S) = -1$ ist eine Spiegelung.

Beweis. Gegeben sei eine orthogonale Matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{O}(2).$$

Die definierende Eigenschaft $A^tA = \mathbf{1}$ ist äquivalent zum Gleichungssystem

$$\begin{aligned} a^2 + c^2 &= 1, \\ ab + cd &= 0, \\ b^2 + d^2 &= 1. \end{aligned}$$

Aus $a^2 + c^2 = 1$ und etwas Analysis folgt, daß es einen Winkel $\varphi \in \mathbb{R}$ gibt, mit

$$a = \cos(\varphi) \quad \text{und} \quad c = \sin(\varphi).$$

Da $ab + cd = 0$ folgt die Existenz von $\varepsilon \in \mathbb{R}$ mit

$$b = -\varepsilon\sin(\varphi) \quad \text{und} \quad d = \varepsilon\cos(\varphi).$$

Aus $b^2 + d^2 = 1$ folgt $\varepsilon \in \{\pm 1\}$. Damit ist A die Matrix

$$A = \begin{pmatrix} \cos(\varphi) & -\varepsilon\sin(\varphi) \\ \sin(\varphi) & \varepsilon\cos(\varphi) \end{pmatrix}$$

und $\det(A) = \varepsilon(\cos(\varphi)^2 + \sin(\varphi)^2) = \varepsilon$.

(1) Sei nun $A \in \mathrm{SO}(2)$, also $\varepsilon = \det(A) = 1$. Dann ist $A = D_\varphi$ eine Drehung. Damit haben wir gesehen, daß die Abbildung $D : \mathbb{R} \to \mathrm{SO}(2)$ mit

$$D(\varphi) = D_\varphi$$

surjektiv ist. Die Additionstheoreme für Sinus und Cosinus zeigen, daß D ein Gruppenhomomorphismus ist:

$$\begin{aligned} D_\varphi D_\psi &= \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}\begin{pmatrix} \cos(\psi) & -\sin(\psi) \\ \sin(\psi) & \cos(\psi) \end{pmatrix} \\ &= \begin{pmatrix} \cos(\varphi)\cos(\psi) - \sin(\varphi)\sin(\psi) & -\cos(\varphi)\sin(\psi) - \sin(\varphi)\cos(\psi) \\ \sin(\varphi)\cos(\psi) + \cos(\phi)\sin(\psi) & -\sin(\varphi)\sin(\psi) + \cos(\varphi)\cos(\psi) \end{pmatrix} \\ &= \begin{pmatrix} \cos(\varphi+\psi) & -\sin(\varphi+\psi) \\ \sin(\varphi+\psi) & \cos(\varphi+\psi) \end{pmatrix} = D_{\varphi+\psi}. \end{aligned}$$

Der Rest von (1) folgt aus dem Homomorphiesatz angewandt auf den surjektiven Gruppenhomomorphismus $D : \mathbb{R} \to \mathrm{SO}(2)$, wenn wir den Kern von D bestimmen:

$$\ker(D) = \{\varphi \; ; \; D_\varphi = \mathbf{1}\} = \{\varphi \; ; \; \exists n \in \mathbb{Z} : \varphi = 2\pi n\} = 2\pi\mathbb{Z}.$$

Für Aussage (2) fehlt nur noch, die Matrix A im Fall $\varepsilon = -1$ als Matrix einer Spiegelung zu identifizieren, also

$$A = \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ \sin(\varphi) & -\cos(\varphi) \end{pmatrix}$$

Die ONB aus

$$v = \begin{pmatrix} \sin(\varphi/2) \\ -\cos(\varphi/2) \end{pmatrix} \quad \text{und} \quad w = \begin{pmatrix} \cos(\varphi/2) \\ \sin(\varphi/2) \end{pmatrix}$$

wird durch Multiplikation mit A wie folgt abgebildet (wieder mittels der Additionstheoreme)

$$Av = \begin{pmatrix} \cos(\varphi)\sin(\varphi/2) - \sin(\varphi)\cos(\varphi/2) \\ \sin(\varphi)\sin(\varphi/2) + \cos(\varphi)\cos(\varphi/2) \end{pmatrix} = \begin{pmatrix} \sin(\varphi/2 - \varphi) \\ \cos(\varphi/2 - \varphi) \end{pmatrix} = -v$$

$$Aw = \begin{pmatrix} \cos(\varphi)\cos(\varphi/2) + \sin(\varphi)\sin(\varphi/2) \\ \sin(\varphi)\cos(\varphi/2) - \cos(\varphi)\sin(\varphi/2) \end{pmatrix} = \begin{pmatrix} \cos(\varphi - \varphi/2) \\ \sin(\varphi - \varphi/2) \end{pmatrix} = w$$

Damit beschreibt A eine Spiegelung an der Geraden $\mathbb{R}w$ und nimmt in der ONB (v, w) die Form

$$A \sim \begin{pmatrix} -1 & \\ & 1 \end{pmatrix}$$

an. Also besteht $\mathrm{O}(2) \setminus \mathrm{SO}(2)$ ausschließlich aus Spiegelungen. $\square$

Bemerkung 9.8. Die Gruppe $\mathrm{O}(2)$ ist eine Art kontinuierliche Diedergruppe. Für jedes $n \geq 1$ gibt es einen natürlichen injektiven Gruppenhomomorphismus $D_n \hookrightarrow \mathrm{O}(2)$ der Diedergruppe der Ordnung $2n$.

Korollar 9.9. *Jede Drehung im $\mathbb{R}^2$ ist das Produkt zweier Spiegelungen.*

Beweis. Die Drehung um den Winkel φ schreiben wir als

$$D_\varphi = \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix} = \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ \sin(\varphi) & -\cos(\varphi) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Dabei haben wir im Beweis von Satz 9.7 den ersten Faktor als Spiegelung an der Geraden orthogonal zu $v = \binom{\sin(\varphi/2)}{-\cos(\varphi/2)}$ erkannt, und der zweite Faktor ist der Spezialfall $\varphi = 0$, die Spiegelung an der Geraden orthogonal zu $\binom{0}{1}$. $\square$

9.2. **Bewegungen.** Wir studieren nun die Abbildungen, welche die metrischen Eigenschaften des euklidischen Raumes erhalten.

Definition 9.10. Sei V ein euklidischer Vektorraum.

(1) Eine **Bewegung** von V ist eine abstandserhaltende Abbildung $f : V \to V$ von Mengen: für alle $v, w \in V$ gilt
$$d(f(v), f(w)) = d(v, w).$$
(2) Eine **Isometrie** ist eine Bewegung, die eine lineare Abbildung ist.

Beispiel 9.11. Sei V ein euklidischer Vektorraum und $w \in V$. Die **Translation** mit w ist die Bewegung $T_w : V \to V$,

$$T_w(v) = v + w.$$

In der Tat ist T_w eine Bewegung, denn für alle $x, y \in V$ gilt

$$d(T_w(x), T_w(y)) = \|T_w(x) - T_w(y)\| = \|(x + w) - (y + w)\| = \|x - y\| = d(x, y).$$

Die Translation T_w ist genau dann linear, wenn $w = 0$.

Satz 9.12. *Sei $(V, \langle\ ,\ \rangle)$ ein euklidischer Vektorraum, und $f : V \to V$ eine Abbildung von Mengen. Dann sind äquivalent:*

(a) f *ist Isometrie.*
(b) f *ist Bewegung mit* $f(0) = 0$.
(c) f *erhält das Skalarprodukt: für alle* $v, w \in V$ *gilt* $\langle f(v), f(w)\rangle = \langle v, w\rangle$.
(d) f *ist isometrische Abbildung, also insbesondere linear.*
(e) f *ist linear und normerhaltend, d.h. für alle* $v \in V$ *gilt* $\|f(v)\| = \|v\|$.

Beweis. Wir zeigen die Äquivalenz durch einen Ringschluß:

(a) $\implies$ (b) $\implies$ (c) $\implies$ (d) $\implies$ (e) $\implies$ (a).

(a) $\implies$ (b): Eine Isometrie ist eine Bewegung und erhält als lineare Abbildung die Null.

(b) $\implies$ (c): Sei f eine Bewegung, die den Nullpunkt erhält. Wegen $f(0) = 0$ gilt für alle $v \in V$

$$\|f(v)\| = \|f(v) - f(0)\| = d(f(v), f(0)) = d(v, 0) = \|v - 0\| = \|v\|. \tag{9.1}$$

Die Polarisationsformel, Satz 5.16, angewandt auf v und $-w$ liefert für alle $v, w \in V$

$$2\langle v, w\rangle = \|v\|^2 + \|w\|^2 - \|v - w\|^2 = \|v\|^2 + \|w\|^2 - d(v, w)^2. \tag{9.2}$$

Damit folgt (c) aus (9.1) und (9.2) mit der Rechnung

$$2\langle f(v), f(w)\rangle = \|f(v)\|^2 + \|f(w)\|^2 - d(f(v), f(w))^2 = \|v\|^2 + \|w\|^2 - d(v, w)^2 = 2\langle v, w\rangle.$$

(c) $\implies$ (d): Sei f skalarprodukterhaltend. Dann erhält f wegen $\|v\|^2 = \langle v, v\rangle$ auch die Norm. Aus (9.2) folgt

$$\|v - w\|^2 = \|v\|^2 + \|w\|^2 - 2\langle v, w\rangle, \tag{9.3}$$

und somit für alle $\lambda \in K$ und $v \in V$:

$$\begin{aligned}
\|f(\lambda v) - \lambda f(v)\|^2 &= \|f(\lambda v)\|^2 + \|\lambda f(v)\|^2 - 2\langle f(\lambda v), \lambda f(v)\rangle \\
&= \|f(\lambda v)\|^2 + \lambda^2\|f(v)\|^2 - 2\lambda\langle f(\lambda v), f(v)\rangle \\
&= \|\lambda v\|^2 + \lambda^2\|v\|^2 - 2\lambda\langle \lambda v, v\rangle = \|\lambda v\|^2 + \|\lambda v\|^2 - 2\langle \lambda v, \lambda v\rangle = 0
\end{aligned}$$

Da 0 der einzige Vektor von Norm 0 ist, folgt für alle $\lambda \in K$ und $v \in V$:

$$f(\lambda v) = \lambda f(v).$$

Für alle $v, w \in V$ berechnet sich mit der Polarisationsformel, Satz 5.16, und wieder (9.3)

$$\|f(v + w) - (f(v) + f(w))\|^2$$

$$\begin{aligned}
&= \|f(v + w)\|^2 + \|f(v) + f(w)\|^2 - 2\langle f(v + w), f(v) + f(w)\rangle \\
&= \|v + w\|^2 + \|f(v)\|^2 + \|f(w)\|^2 + 2\langle f(v), f(w)\rangle - 2\langle f(v + w), f(v)\rangle - 2\langle f(v + w), f(w)\rangle \\
&= \|v + w\|^2 + \big(\|v\|^2 + \|w\|^2 + 2\langle v, w\rangle\big) - 2\langle v + w, v\rangle - 2\langle v + w, w\rangle \\
&= \|v + w\|^2 + \|v + w\|^2 - 2\langle v + w, v + w\rangle = 0
\end{aligned}$$

Da 0 der einzige Vektor von Norm 0 ist, folgt für alle $v, w \in V$:

$$f(v + w) = f(v) + f(w).$$

Damit ist f linear, und weil f zudem nach Voraussetzung das Skalarprodukt erhält, handelt es sich bei f um eine isometrische Abbildung im Sinne dieser Vorlesung.

(d) $\implies$ (e): Sei f isometrische Abbildung. Dann gilt für alle $v \in V$

$$\|f(v)\|^2 = \langle f(v), f(v)\rangle = \langle v, v\rangle = \|v\|^2.$$

Somit ist f linear und normerhaltend.

(e) $\Longrightarrow$ (a): Sei f linear und normerhaltend. Dann gilt für alle $v, w \in V$

$$d(f(v), f(w)) = \|f(v) - f(w)\| = \|f(v - w)\| = \|v - w\| = d(v, w).$$

Somit ist f eine lineare Bewegung: eine Isometrie. $\square$

Wir zeigen nun, daß bis auf eine Translation eine Bewegung linear ist.

Proposition 9.13. *Die Komposition von Bewegungen ist eine Bewegung. Die Komposition von Isometrien ist eine Isometrie.*

Beweis. Das ist klar. $\square$

Korollar 9.14. *Sei $(V, \langle\ ,\ \rangle)$ ein euklidischer Vektorraum und $F : V \to V$ eine Bewegung. Dann gibt es $w \in V$ und eine Isometrie $f : V \to V$, so daß für alle $v \in V$*

$$F(v) = (T_w \circ f)(v) = f(v) + w.$$

Der Vektor w und die Isometrie f sind eindeutig.

Beweis. Mit $w = F(0)$ ist $f = T_{-w} \circ F$ eine Bewegung mit $f(0) = 0$, also nach Satz 9.12 eine Isometrie. Es gilt dann

$$F = (T_w \circ T_{-w}) \circ F = T_w \circ (T_{-w} \circ F) = T_w \circ f.$$

Die Eindeutigkeitsaussage bleibt zur Übung. $\square$

Satz 9.15. *Sei $(V, \langle\ ,\ \rangle)$ ein euklidischer Vektorraum.*

(1) Eine Isometrie ist ein Automorphismus des zugrundeliegenden Vektorraums.
(2) Eine Bewegung ist bijektiv.

Beweis. Sei $f : V \to V$ eine Isometrie. Als euklidischer Raum ist V per Definition endlichdimensional. Dann ist f als lineare Abbildung zwischen endlich-dimensionalen Vektorräumen der gleichen Dimension bijektiv genau dann, wenn $\ker(f) = 0$.

Als Isometrie ist f nach Satz 9.12 normerhaltend. Sei $v \in V$ und $f(v) = 0$. Dann gilt

$$\|v\| = \|f(v)\| = 0,$$

also $v = 0$. Dies zeigt Aussage (1).

(2) Wir schreiben eine Bewegung F nach Korollar 9.14 als $F = T_w \circ f$ mit Translation T_w und Isometrie f. Da Translationen bijektiv sind, folgt (2) aus (1). $\square$

Beispiel 9.16. Sei V der $\mathbb{R}$-Vektorraum der ℓ^2-Folgen, also der Folgen $(a_n)_{n\in\mathbb{N}}$ mit

$$\sum_{i=1}^{\infty} a_i^2 < \infty.$$

Auf V ist das ℓ^2-Skalarprodukt definiert als

$$\langle (a_n)_{n\in\mathbb{N}}, (b_n)_{n\in\mathbb{N}} \rangle_{\ell^2} := \sum_{i=1}^{\infty} a_i b_i.$$

Der Shift-Operator $S : V \to V$ definiert durch

$$S((a_1, a_2, a_3, \ldots)) = (0, a_1, a_2, \ldots)$$

ist eine Bewegung, die nicht bijektiv ist. Das ℓ^2-Skalarprodukt ist eine positiv definite, symmetrische Bilinearform. Trotzdem ist V kein euklidischer Raum, denn diese sind per Definition endlich dimensional.

9.3. **Isometrien.** Nun charakterisieren wir Isometrien in den linearen Abbildungen.

Satz 9.17. *Sei $(V, \langle\ ,\ \rangle)$ ein euklidischer Vektorraum mit ONB $\mathscr{B}$ und sei $f : V \to V$ eine lineare Abbildung. Sei $A = \mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f)$ die Darstellungsmatrix. Dann sind äquivalent:*

(a) *f ist Isometrie.*
(b) *f führt die ONB $\mathscr{B} = (b_1, \ldots, b_n)$ in eine ONB $f(\mathscr{B}) = (f(b_1), \ldots, f(b_n))$ über.*
(c) *A ist orthogonal.*
(d) *Die Spalten von A sind eine ONB von $\mathbb{R}^n$ mit dem Standardskalarprodukt.*
(e) *Die transponierten Zeilen von A sind eine ONB von $\mathbb{R}^n$ mit dem Standardskalarprodukt.*
(f) *$f^*f = \mathrm{id}_V$.*
(g) *f ist invertierbar und $f^{-1} = f^*$.*
(h) *$ff^* = \mathrm{id}_V$.*

Beweis. Wir zeigen die Äquivalenz durch die folgenden Implikationen:

$$\begin{tikzcd} (a) \arrow[r, Rightarrow] & (b) \arrow[dr, Leftrightarrow] & & & (g) \arrow[r, Leftrightarrow] & (h) \\ (e) \arrow[r, Leftrightarrow] & (d) \arrow[r, Leftrightarrow] & (c) \arrow[r, Leftrightarrow] & (f) \arrow[r, Rightarrow] \arrow[ur, Leftrightarrow] & (a) \end{tikzcd}$$

(a) $\Longrightarrow$ (b): Sei $\mathscr{B} = (b_1, \ldots, b_n)$ eine ONB und f eine Isometrie. Dann ist f skalarprodukterhaltend nach Satz 9.12 und $f(\mathscr{B})$ eine ONB, weil

$$\langle f(b_i), f(b_j)\rangle = \langle b_i, b_j\rangle.$$

(b) $\Longleftrightarrow$ (c): folgt aus Satz 6.45.

(c) $\Longleftrightarrow$ (d) $\Longleftrightarrow$ (e): folgen jeweils aus Korollar 6.46.

(c) $\Longleftrightarrow$ (f): Es ist $\mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f^*) = A^t$ nach Satz 5.26, denn $\mathscr{B}^* = \mathscr{B}$. Die Matrix von f^*f ist A^tA und somit $A^tA = \mathbf{1} \iff f^*f = \mathrm{id}_V$.

(f) $\Longrightarrow$ (a): Für alle v, w gilt

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

also ist f skalarprodukterhaltend und nach Satz 9.12 eine Isometrie.

(g) $\Longrightarrow$ (f) und (h): trivial.

(f) oder (h) $\Longrightarrow$ (g): Wegen $f^*f = \mathrm{id}_V$ (oder $ff^* = \mathrm{id}_V$) ist f surjektiv (oder injektiv) und als Endomorphismus eines endlichdimensionalen Vektorraums auch bijektiv. Dann muß aber das Linksinverse (oder Rechtsinverse) schon Inverses sein: $f^{-1} = f^*$. □

Korollar 9.18. *Die Isometrien des $\mathbb{R}^n$ bezüglich des Standardskalarprodukts sind genau die Matrixmultiplikationen mit reellen orthogonalen Matrizen.*

Beweis. Das folgt sofort aus Satz 9.17 (a) $\Longleftrightarrow$ (c) im Fall $V = \mathbb{R}^n$ mit dem Standardskalarprodukt. □

Korollar 9.19. *Für $A \in \mathrm{O}(n)$ und $v \in \mathbb{R}^n$ gilt $\|Av\| = \|v\|$.*

Beweis. Sofort aus Korollar 9.18 und den Eigenschaften einer Isometrie. □

Korollar 9.20. *Jeder reelle Eigenwert einer orthogonalen Matrix $A \in \mathrm{O}(n)$ ist 1 oder -1.*

Beweis. Sei λ ein (reeller) Eigenwert von A, und sei v ein zugehöriger Eigenvektor. Dann gilt

$$\|v\| = \|Av\| = \|\lambda v\| = |\lambda| \cdot \|v\|,$$

und daher $\lambda = \pm 1$, weil $\|v\| \neq 0$. □

Proposition 9.21. *Jeder reelle Eigenwert einer Isometrie ist 1 oder -1.*

Beweis. Vermöge der Übersetzung von linearen Abbildungen in Matrizen folgt die Aussage für eine Isometrie sofort aus der für eine orthogonale Matrix, siehe Korollar 9.20. □

Korollar 9.22. *Ist f Isometrie eines euklidischen Vektorraums, dann gilt* $\det(f) = \pm 1$.

Beweis. Die Matrix $A = \mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f)$ bezüglich einer ONB $\mathscr{B}$ ist orthogonal. Daraus folgt

$$\det(f)^2 = \det(A)^2 = \det(A^t)\det(A) = \det(A^tA) = \det(\mathbf{1}) = 1.$$ □

Bemerkung 9.23. Korollar 9.22 bedeutet insbesondere, daß sich das Volumen eines Parallelotops unter einer Isometrie wegen Satz 8.46 nicht ändert. Das soll auch so sein, schließlich ist eine Isometrie eine Bewegung, die sämtliche metrische Eigenschaften erhalten soll.

Übungsaufgaben zu §9

Übungsaufgabe 9.1. Sei V ein euklidischer Vektorraum. Die Spiegelung am Orthogonalraum zu $v \in V$, $v \neq 0$ hat die Formel

$$S_v(w) = w - 2\frac{\langle v, w\rangle}{\langle v, v\rangle} v.$$

Berechnen Sie direkt mit dieser Formel, daß S_v eine Isometrie ist.

Übungsaufgabe 9.2. Wir betrachen im $\mathbb{R}^3$ den Würfel mit den Ecken $(\pm 1, \pm 1, \pm 1)$. Die Raumdiagonalen, das sind die Geraden durch gegenüberliegende Ecken des Würfels treffen sich im Punkt $0 \in \mathbb{R}^3$. Bestimmen Sie den Winkel zwischen zwei dieser Raumdiagonalen. Hängt das Ergebnis von der Wahl der Raumdiagonalen ab?

Übungsaufgabe 9.3. Auf $\mathbb{R}^4$ betrachten wir die Bilinearform der Signatur $(3, 1)$, welche für $x, y \in \mathbb{R}^4$ durch

$$\eta(x, y) = x^t \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix} y$$

gegeben ist. Auf der Teilmenge

$$V = \{x = (x_1, x_2, x_3, t) \in \mathbb{R}^4 \ ; \ t > 0, \ \eta(x, x) < 0\}$$

definieren wir eine „Norm“ durch

$$\|x\|_\eta = \sqrt{-\eta(x, x)}.$$

Zeigen Sie:
(1) Wenn $x, y \in V$, dann auch $x + y \in V$.
(2) Für $\|-\|_\eta$ gilt für Vektoren aus V nicht die Dreiecksungleichung, sondern die „falsche“ Dreiecksungleichung: für alle $x, y \in V$ gilt

$$\|x + y\|_\eta \geq \|x\|_\eta + \|y\|_\eta.$$

Anmerkung: Wenn man die Lichtgeschwindigkeit auf $c = 1$ normiert (kann man machen in den richtigen Einheiten), dann ist η die Minkowski–Metrik auf der 4-dimensionalen Raumzeit. Die Menge V besteht dann aus den aus physikalischer Sicht von 0 aus erreichbaren Punkten, wenn man dem Postulat folgt, daß nichts schneller als Licht unterwegs sein darf. Der Wert $\|x\|_\eta$ ist die nach der speziellen Relativitätstheorie vom mitfliegenden Betrachter erlebte verstrichene Eigenzeit auf der direkten geradlinigen Reise von 0 nach x. Die falsche Dreiecksungleichung beschreibt, daß die Reise auf gerader Linie von 0 nach $x + y$ mehr Eigenzeit in Anspruch nimmt als die Reise von 0 nach x und dann von x nach $x + y$.

Mit dieser Beobachtung kann man das Zwillingsparadoxon der speziellen Relativitätstheorie erklären: Von zwei Zwillingsschwestern geht Antonia auf eine lange schnelle Reise zu einem andern Stern, und Barbara bleibt auf der Erde. Nach der Rückkehr Antonias ist diese plötzlich jünger als Barbara, und zwar, je schneller sie unterwegs war, desto jünger ist Antonia.

Antonias Reise verläuft idealisiert von 0 nach $x = (x_1, x_2, x_3, t)$ und dann weiter nach

$$(2t, 0, 0, 0) = x + y$$

mit $y = (-x_1, -x_2, -x_3, t)$, während Barbaras Reise direkt von 0 nach $x + y$ verläuft.

Übungsaufgabe 9.4. Zeigen Sie, daß die Matrixmultiplikation mit

$$A = \begin{pmatrix} -\frac{4}{5} & 0 & \frac{3}{5} \\ \frac{9}{25} & -\frac{4}{5} & \frac{12}{25} \\ \frac{12}{25} & \frac{3}{5} & \frac{16}{25} \end{pmatrix}$$

eine Isometrie des $\mathbb{R}^3$ (mit Standardskalarprodukt) beschreibt. Bestimmen Sie die Isometrienormalform von A.

Übungsaufgabe 9.5. Sei $n \in \mathbb{N}$ eine ungerade natürliche Zahl und $A \in \mathrm{O}_n(\mathbb{R})$ eine orthogonale Matrix. Zeigen Sie, daß A einen reellen Eigenvektor hat.

Übungsaufgabe 9.6. Zeigen Sie, daß jede orthogonale Matrix $A \in \mathrm{SO}_3(\mathbb{R})$ eine Drehung um eine Drehachse $\mathbb{R}v$ ist, d.h., es gibt einen Eigenvektor v zum Eigenwert 1.

Übungsaufgabe 9.7. In dieser Aufgabe wollen wir eine elegante Methode zur Beschreibung von Drehungen des $\mathbb{R}^3$ behandeln, die in 3D-Computergraphik angewandt wird. Diese Aufgabe benutzt Begriffe aus der Vorlesung Grundlagen der Algebra (und fällt auch ansonsten offensichtlich aus dem Rahmen, ist aber zu interessant, um nicht gestellt zu werden).

Die Hamiltonschen Quaternionen kann man definieren als die Menge $\mathbb{H} \subseteq \mathrm{M}_2(\mathbb{C})$ der Matrizen

$$\begin{pmatrix} z & -\bar{w} \\ w & \bar{z} \end{pmatrix}$$

mit $z, w \in \mathbb{C}$ beliebig. Zeigen Sie die folgenden Aussagen.

(a) $\mathbb{H}$ ist ein Unterring des Matrizenrings $\mathrm{M}_2(\mathbb{C})$.

(b) $\mathbb{H}$ ist ein $\mathbb{R}$-Vektorraum der Dimension 4 mit Basis

$$\mathbf{1} := \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}, \quad i := \begin{pmatrix} i & \\ & -i \end{pmatrix}, \quad j := \begin{pmatrix} & 1 \\ -1 & \end{pmatrix}, \quad k := \begin{pmatrix} & i \\ i & \end{pmatrix}.$$

(c) Wir identifizieren die lineare Hülle von $\mathbf{1} \in \mathbb{H}$ mit $\mathbb{R}$ durch $\lambda \mapsto \lambda \cdot \mathbf{1}$ und setzen

$$\mathbb{H}_+ = \langle \mathbf{1} \rangle = \mathbb{R}$$
$$\mathbb{H}_- = \langle i, j, k \rangle$$

mit linearen Hüllen als $\mathbb{R}$-Untervektorräume. Dann ist $\mathbb{H} = \mathbb{H}_+ \oplus \mathbb{H}_-$ die Eigenraumzerlegung bezüglich der Involution von $\mathbb{H}$, genannt die **Konjugation**, welche induziert wird von der Involution auf $\mathrm{M}_2(\mathbb{C})$, die sowohl die Matrix transponiert als auch die Einträge komplex konjugiert. Wir schreiben $\overline{\alpha}$ für das Konjugierte zu $\alpha \in \mathbb{H}$.

(d) Sei $\pi_+ : \mathbb{H} \to \mathbb{H}$ die Projektion auf den Summanden $\mathbb{H}_+ = \mathbb{R}$. Dann definiert

$$(\alpha, \beta) = \pi_+(\alpha\overline{\beta})$$

ein Skalarprodukt auf $\mathbb{H}$. Bestimmen Sie die Matrix des Skalarprodukts bezüglich der Basis $1, i, j, k$.

(e) Wir bezeichnen die Menge der Norm-1-Elemente

$$\{\alpha \in \mathbb{H} \mid \alpha\overline{\alpha} = 1\} \subset \mathbb{H}$$

als 3-Sphäre S^3 (erklären Sie dies). Zeigen Sie, daß die Multiplikation von Quaternionen aus S^3 eine Gruppe macht und daß ein Quaternion α zu S^3 gehört, genau dann wenn $\alpha^{-1} = \overline{\alpha}$.

(f) Zu $\alpha \in S^3$ definieren wir $\varphi_\alpha : \mathbb{H} \to \mathbb{H}$ mittels der Quaternionenmultiplikation wie folgt:

$$\varphi_\alpha(x) = \alpha(x)\alpha^{-1}.$$

Zeigen Sie, daß φ_α zu $\alpha \in S^3$ die Zerlegung $\mathbb{H} = \mathbb{H}_+ \oplus \mathbb{H}_-$ respektiert und φ_α orthogonal ist bezüglich des Skalarprodukts aus (d).

(g) Identifiziert man $\mathbb{H}_-$ mit $\mathbb{R}^3$ durch die Basis i, j, k, so liefert die Konstruktion aus (f) einen Gruppenhomomorphismus

$$\varphi : S^3 \to \mathrm{SO}(3), \quad \alpha \mapsto \varphi_\alpha|_{\mathbb{H}_-},$$

mit Kern $\ker(\varphi) = \{\pm 1\}$.

(h) Jedes Quaternion $\alpha \in S^3$ läßt sich schreiben als

$$\alpha = \cos\theta + \sin\theta\big(v_1 i + v_2 j + v_3 k\big),$$

wobei $0 \le \theta < \pi$ und $v = (v_1, v_2, v_3) \in \mathbb{R}^3$ ein Einheitsvektor ist. Wie eindeutig ist die Korrespondenz $\alpha \leftrightarrow (\theta, v)$ und kommen alle Paare (θ, v) mit obigen Eigenschaften vor?

(i) Sei $\alpha = \cos\theta + \sin\theta\big(v_1 i + v_2 j + v_3 k\big) \in S^3$. Zeigen Sie, daß die orthogonale Abbildung $\varphi(\alpha)$ des $\mathbb{R}^3$ die Drehung um die Achse $\mathbb{R}v$ mit dem Drehwinkel 2θ ist. Ist der Gruppenhomomorphismus φ aus (g) surjektiv?

Übungsaufgabe 9.8. Sei $(V, \langle\ ,\ \rangle)$ ein euklidischer Vektorraum. Zeigen Sie, daß die Menge der Bewegungen $V \to V$ eine Gruppe bilden.

Zeigen Sie weiter, daß die Gruppe der Bewegungen des $\mathbb{R}^n$ mit Standardskalarprodukt isomorph ist zum semidirekten Produkt

$$\mathrm{O}(n) \ltimes \mathbb{R}^n.$$

Teil 4. Spektraltheorie

10. Spektraltheorie selbstadjungierter Endomorphismen

Spektralsätze geben über Eigenräume und Eigenwerte von Endomorphismen von Vektorräumen Auskunft.

10.1. **Normale Abbildungen.** Sei K ein beliebiger Körper und V ein endlichdimensionaler K-Vektorraum mit perfekter, symmetrischer Bilinearform $\langle\ ,\ \rangle$. Die lineare Abbildung

$$f \mapsto f^*$$

ist eine Involution auf $\mathrm{End}_K(V) = \mathrm{Hom}_K(V,V)$, also $(f^*)^* = f$, siehe Abschnitt §5.4. Wir betonen, daß die adjungierte Abbildung von der gewählten perfekten Bilinearform auf V abhängt.

Definition 10.1. Sei $(V, \langle\ ,\ \rangle)$ eine endlich-dimensionaler K-Vektorraum mit perfekter, symmetrischer Bilinearform.

(1) Ein **normaler Endomorphismus** ist ein $f \in \mathrm{End}_K(V)$, der mit f^* kommutiert:
$$ff^* = f^*f.$$
(2) Ein **selbstadjungierter Endomorphismus** ist ein $f \in \mathrm{End}_K(V)$ mit
$$f = f^*.$$

Der Begriff selbstadjungiert kam bereits in Proposition 6.42 vor, wo bewiesen wurde, daß dies genau die Endomorphismen mit bezüglich ONBs symmetrischer Darstellungsmatrix sind.

Beispiel 10.2. (1) Da ein Endomorphimus stets mit sich selbst kommutiert, sind selbstadjungierte Endomorphismen normal.

(2) Sei $(V, \langle\ ,\ \rangle)$ ein K-Vektorraum mit symmetrischer anisotroper Bilinearform, und sei $U \subseteq V$ ein Unterraum. Die orthogonale Projektion $p_U : V \to V$ auf U ist selbstadjungiert, siehe Proposition 6.35. Für alle $v_i = u_i + w_i$, $i = 1, 2$ mit $u_i \in U$ und $w_i \in U^\perp$ gilt $p_U(v_i) = u_i$ und
$$\langle v_1, p_U(v_2)\rangle = \langle u_1 + w_1, u_2\rangle = \langle u_1, u_2\rangle = \langle u_1, u_2 + w_2\rangle = \langle p_U(v_1), v_2\rangle.$$
(3) Beispiel 5.28 zeigt, daß die Drehung $R_\varphi \in \mathrm{End}(\mathbb{R}^2)$ ein normaler Endomorphismus ist:
$$R_\varphi R_\varphi^* = R_\varphi R_{-\varphi} = \mathbf{1} = R_{-\varphi} R_\varphi = R_\varphi^* R_\varphi.$$
Insbesondere ist nicht jeder normale Endomorphismus selbstadjungiert.

Das Beispiel der Drehungen verallgemeinert sich auf Isometrien.

Proposition 10.3. *Isometrien sind normale Endomorphismen.*

Beweis. Satz 9.17 zeigt $f^* = f^{-1}$, somit $ff^* = \mathrm{id} = f^*f$. □

Satz 10.4. *Sei $(V, \langle\ ,\ \rangle)$ ein endlich-dimensionaler K-Vektorraum mit einer perfekten symmetrischen Bilinearform, und sei $f : V \to V$ ein Endomorphismus. Dann ist f normal genau dann, wenn für alle $v, w \in V$ gilt*
$$\langle f(v), f(w)\rangle = \langle f^*(v), f^*(w)\rangle. \tag{10.1}$$

Beweis. Sei f normal. Dann ist für alle $v, w \in V$
$$\langle f^*(v), f^*(w)\rangle = \langle v, f(f^*(w))\rangle = \langle v, f^*(f(w))\rangle = \langle f(v), f(w)\rangle.$$
Wir nehmen nun (10.1) an. Dann gilt für alle $v, w \in V$ gilt
$$\langle v, f^*(f(w))\rangle = \langle f(v), f(w)\rangle = \langle f^*(v), f^*(w)\rangle = \langle v, f(f^*(w))\rangle.$$
Nach der Eindeutigkeit aus Korollar 5.6 ist dies äquivalent zu $f^*(f(w)) = f(f^*(w))$ für alle $w \in V$ und damit dazu, daß f normal ist. □

10.2. **Eigenwerte und adjungierte Abbildungen.**

Notation 10.5. Sei $f : V \to V$ ein Endomorphismus des K-Vektorraums V, und sei $\lambda \in K$. Der Eigenraum von f zum Eigenwert λ ist der Untervektorraum von V

$$V_\lambda(f) = \{v \in V \ ; \ f(v) = \lambda v\} = \ker(f - \lambda \operatorname{id}_V).$$

Proposition 10.6. *Sei $(V, \langle\ ,\ \rangle)$ ein endlich-dimensionaler K-Vektorraum mit einer perfekten symmetrischen Bilinearform $\langle\ ,\ \rangle$, und sei $f \in \operatorname{End}_K(V)$. Wenn $\lambda \neq \mu$, dann sind $V_\lambda(f)$ und $V_\mu(f^*)$ orthogonal zueinander.*

Beweis. Sei $v \in V_\lambda(f)$ und $w \in V_\mu(f^*)$. Dann gilt

$$\lambda\langle v, w\rangle = \langle \lambda v, w\rangle = \langle f(v), w\rangle = \langle v, f^*(w)\rangle = \langle v, \mu w\rangle = \mu\langle v, w\rangle,$$

so daß $(\lambda - \mu)\langle v, w\rangle = 0$, und wenn $\lambda \neq \mu$ folgt $v \perp w$. □

Korollar 10.7. *Sei $(V, \langle\ ,\ \rangle)$ ein endlich-dimensionaler K-Vektorraum mit einer perfekten symmetrischen Bilinearform $\langle\ ,\ \rangle$, und sei $f \in \operatorname{End}_K(V)$ selbstadjungiert. Wenn $\lambda \neq \mu$, dann sind $V_\lambda(f)$ und $V_\mu(f)$ orthogonal zueinander.*

Beweis. Das folgt sofort wegen $f = f^*$ aus Proposition 10.6. □

Beispiel 10.8. Sei V ein euklidischer Vektorraum, und sei $p_U : V \to V$ die orthogonale Projektion auf einen Unterraum $U \subseteq V$. Dann hat $p_U = p_U^*$ die Eigenwerte 0 und 1 und

$$U = \operatorname{im}(p_U) = V_1(p_U) \quad \text{ist orthogonal zu} \quad V_0(p_U^*) = V_0(p_U) = \ker(p_U) = U^\perp.$$

Lemma 10.9. *Sei $(V, \langle\ ,\ \rangle)$ ein endlich-dimensionaler K-Vektorraum mit einer perfekten symmetrischen Bilinearform, und seien $f, g : V \to V$ normale Endomorphismen, die kommutieren: $fg = gf$. Dann ist der Endomorphismus*

$$f + g^* \in \operatorname{End}_K(V)$$

auch normal.

Beweis. Es kommutieren auch f^* und g^* weil $f^*g^* = (gf)^* = (fg)^* = g^*f^*$. Damit folgt

$$\begin{aligned}(f + g^*)^*(f + g^*) &= (f^* + g)(f + g^*) = f^*f + gf + f^*g^* + gg^* \\ &= ff^* + fg + g^*f^* + g^*g = (f + g^*)(f^* + g) = (f + g^*)(f + g^*)^*.\end{aligned}$$ □

Korollar 10.10. *Sei $(V, \langle\ ,\ \rangle)$ ein endlich-dimensionaler K-Vektorraum mit einer perfekten symmetrischen Bilinearform, und sei $f : V \to V$ ein normaler Endomorphismus. Dann ist für jedes $\lambda \in K$ der Endomorphismus*

$$f - \lambda \operatorname{id}_V \in \operatorname{End}_K(V)$$

auch normal.

Beweis. Weil $(-\lambda \operatorname{id}_V)^* = -\lambda \operatorname{id}_V^* = -\lambda \operatorname{id}_V$ selbstadjungiert (also normal) ist und mit f kommutiert, folgt das Korollar sofort aus Lemma 10.9. □

Wir erinnern daran, daß anisotrope Bilinearformen perfekt sind.

Proposition 10.11. *Sei $(V, \langle\ ,\ \rangle)$ ein endlich-dimensionaler K-Vektorraum mit einer anisotropen symmetrischen Bilinearform. Sei $f : V \to V$ ein normaler Endomorphismus. Dann ist*

$$V_\lambda(f) = V_\lambda(f^*).$$

Beweis. Es gilt wegen der vorausgesetzzten Anisotropie

$$f(v) = \lambda v \iff f(v) - \lambda v = 0 \iff \langle (f - \lambda \operatorname{id}_V)(v), (f - \lambda \operatorname{id}_V)(v)\rangle = 0.$$

Mit f ist nun nach Korollar 10.10 auch $f - \lambda \operatorname{id}_V$ normal, und wegen $(f - \lambda \operatorname{id}_V)^* = f^* - \lambda \operatorname{id}_V$ folgt daher aus Satz 10.4

$$\langle (f - \lambda \operatorname{id}_V)(v), (f - \lambda \operatorname{id}_V)(v)\rangle = \langle (f^* - \lambda \operatorname{id}_V)(v), (f^* - \lambda \operatorname{id}_V)(v)\rangle.$$

Jetzt argumentieren wir mit f^* anstelle von f rückwärts und finden $f^*(v) = \lambda v$. $\square$

Notation 10.12. Sei $f \in \mathrm{End}_K(V)$ ein Endomorphimus eines endlichdimensionalen K-Vektorraums. Das **charakteristische Polynom** von f ist

$$\chi_f(X) = \det(X \cdot \mathrm{id}_V - f) \in K[X].$$

Theorem 10.13 (Spektralsatz für zerfallende normale Operatoren)**.** *Sei* $(V, \langle\ ,\ \rangle)$ *ein endlichdimensionaler* K*-Vektorraum mit symmetrischer anisotroper Bilinearform. Sei* $f \in \mathrm{End}_K(V)$ *ein Endomorphismus. Dann sind äquivalent:*

(a) *V besitzt eine Orthogonalbasis aus Eigenvektoren für f.*
(b) *$V = V_{\lambda_1}(f) \oplus^\perp \ldots \oplus^\perp V_{\lambda_r}(f)$ für die Eigenwerte $\lambda_1, \ldots, \lambda_r$ von f.*
(c) *f ist normal und $\chi_f(X)$ zerfällt in $K[X]$ in Linearfaktoren.*
(d) *f ist selbstadjungiert und $\chi_f(X)$ zerfällt in $K[X]$ in Linearfaktoren.*

Beweis. (b) $\Longrightarrow$ (a): Wir wählen für jeden Eigenraum $V_{\lambda_i}(f)$ eine Orthogonalbasis. Diese fügen sich zusammen zu einer Orthogonalbasis von V, da die Eigenräume paarweise zueinander orthogonal sind. Die resultierende Orthogonalbasis besteht aus Eigenvektoren zu f.

(a) $\Longrightarrow$ (d): Sei $\mathscr{B} = (b_1, \ldots, b_n)$ eine Orthogonalbasis aus Eigenvektoren für f. Sei λ_i der Eigenwert zu b_i. Dann sind für beliebige $v = \sum_{i=1}^n x_i b_i$ und $w = \sum_{i=1}^n y_i b_i$

$$\begin{aligned}\langle f(v), w\rangle &= \sum_{i=1}^n x_i \langle f(b_i), w\rangle = \sum_{i=1}^n x_i \langle \lambda_i b_i, w\rangle = \sum_{i=1}^n \sum_{j=1}^n x_i y_j \langle \lambda_i b_i, b_j\rangle = \sum_{i=1}^n x_i y_i \langle \lambda_i b_i, b_i\rangle \\ &= \sum_{i=1}^n x_i y_i \langle b_i, \lambda_i b_i\rangle = \sum_{i=1}^n \sum_{j=1}^n x_j y_i \langle b_j, \lambda_i b_i\rangle = \sum_{i=1}^n y_i \langle v, \lambda_i b_i\rangle = \sum_{i=1}^n y_i \langle v, f(b_i)\rangle \\ &= \langle v, f(w)\rangle.\end{aligned}$$

Hieraus folgt $f = f^*$. Außerdem ist $A = \mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f)$ eine Diagonalmatrix mit Einträgen $\lambda_1, \ldots, \lambda_n$. Somit zerfällt $\chi_f(X) = \chi_A(X) = \prod_{i=1}^n (X - \lambda_i)$ als charakteristisches Polynom einer Diagonalmatrix in Linearfaktoren.

(d) $\Longrightarrow$ (c): Das ist trivial.

(c) $\Longrightarrow$ (b): Wir beweisen diesen Schritt per vollständiger Induktion nach $n = \dim_K(V)$. Für $n = 0$ ist nichts zu zeigen. Wir nehmen nun an, daß Theorem 10.13 für Dimension $< n$ richtig ist. Da $\chi_f(X)$ in Linearfaktoren zerfällt, gibt es einen Eigenwert λ aus K. Da $\langle\ ,\ \rangle$ anisotrop ist, ist $W = V_\lambda(f)^\perp$ das orthogonale Komplement zu $V_\lambda(f)$ und

$$V = V_\lambda(f) \oplus^\perp W.$$

Weil f normal ist und damit $V_\lambda(f) = V_\lambda(f^*)$ nach Proposition 10.11, gilt für all $w \in W$ und $v \in V_\lambda(f)$

$$\langle v, f(w)\rangle = \langle f^*(v), w\rangle = \langle \lambda v, w\rangle = \lambda\langle v, w\rangle = 0$$

und

$$\langle v, f^*(w)\rangle = \langle f(v), w\rangle = \langle \lambda v, w\rangle = \lambda\langle v, w\rangle = 0.$$

Es folgt, daß W ein f-stabiler und f^*-stabiler Unterraum ist. Sei $h = f|_W$ und $g = f^*|_W$. Dann ist für alle $w_1, w_2 \in W$

$$\langle h(w_1), w_2\rangle = \langle f(w_1), w_2\rangle = \langle w_1, f^*(w_2)\rangle = \langle w_1, g(w_2)\rangle,$$

und das heißt

$$(f|_W)^* = h^* = g = f^*|_W.$$

Damit ist $(W, \langle\ ,\ \rangle|_{W\times W})$ ein K-Vektorraum mit anisotroper symmetrischer Bilinearform, und $f|_W \in \mathrm{End}_K(W)$ ein normaler Endomorphismus:

$$f|_W \circ (f|_W)^* = f|_W \circ f^*|_W = (f \circ f^*)|_W = (f^* \circ f)|_W = f^*|_W \circ f|_W = (f|_W)^* \circ f|_W.$$

In einer zur Zerlegung $V = V = V_\lambda(f) \oplus^\perp W$ angepaßten Basis nimmt f die Blockdiagonalform

$$\begin{pmatrix} \lambda \cdot \mathbf{1} & 0 \\ 0 & f|_W \end{pmatrix}$$

an, so daß

$$\chi_f(X) = \chi_{f|_W}(X) \cdot (X - \lambda)^{\dim(V_\lambda(f))}.$$

Es folgt nun aus der Eindeutigkeit der Zerlegung in irreduzible Faktoren im Polynomring $K[X]$, siehe Skript zur Vorlesung *Grundlagen der Algebra*, daß auch $f|_W$ ein in Linearfaktoren zerfallendes charakteristisches Polynom besitzt. Daher können wir für W und $f|_W$ per Induktion auf Eigenschaft (b) schließen, denn $\dim(W) < \dim(V)$.

Konkret: wir setzen für $\mu \in K$ für den Eigenraum von $f|_W$ zum Eigenwert μ

$$W_\mu(f) := V_\mu(f|_W) = \{w \in W \ ; \ f(w) = \mu w\} = W \cap V_\mu(f). \tag{10.2}$$

Nach Induktionsvoraussetzung sind für die Eigenwerte $\mu_1, \ldots, \mu_s$ von $f|_W$

$$W = W_{\mu_1}(f) \oplus^\perp \ldots \oplus^\perp W_{\mu_s}(f)$$

die orthogonale Summe der Eigenräume von $f|_W$. Für $\lambda \neq \mu$ ist nach Proposition 10.6 und Proposition 10.11

$$V_\mu(f) \subseteq V_\lambda(f)^\perp \subseteq \langle v \rangle^\perp = W$$

und damit $V_\mu(f) = W_\mu(f)$. Für $\lambda = \mu$ gilt

$$W_\lambda(f) = V_\lambda(f) \cap V_\lambda(f)^\perp = (0).$$

Wegen

$$\chi_f(X) = (X - \lambda)^{\dim(V_\lambda(f))} \chi_{f|_W}(X)$$

sind die Eigenwerte von f gerade $\lambda, \mu_1, \mu_2, \ldots, \mu_s$. Dann gilt:

$$\begin{aligned} V = V_\lambda(f) \oplus^\perp W &= V_\lambda(f) \oplus^\perp (W_{\mu_1}(f) \oplus^\perp \ldots \oplus^\perp W_{\mu_s}(f)) \\ &= V_\lambda(f) \oplus^\perp V_{\mu_1}(f) \oplus^\perp \ldots \oplus^\perp V_{\mu_s}(f). \end{aligned}$$ □

Bemerkung 10.14. Theorem 10.13 gilt insbesondere für $K = \mathbb{R}$ und einen euklidischen Vektorraum $(V, \langle \ , \ \rangle)$. Ein Skalarprodukt ist symmetrisch, positiv definit und daher anisotrop. Beispielsweise kann man $V = \mathbb{R}^n$ mit dem Standardskalarprodukt wählen.

Im Fall $K = \mathbb{R}$ und einem euklidischen Vektorraum $(V, \langle \ , \ \rangle)$ kann man überdies in der Orthogonalbasis die Basisvektoren normieren, so daß man speziell eine ONB erhält.

Bemerkung 10.15. Drehungen sind normal, aber nicht selbstadjungiert. Auf die Voraussetzung zerfallender charakteristischer Polynome in Theorem 10.13(c)-(d) kann man nicht verzichteten.

Theorem 10.16 (Der Spektralsatz für Matrizen)**.** *Sei $A \in \mathrm{M}_n(K)$ ein quadratische Matrix und das Standardskalarprodukt auf K^n sei anisotrop. Dann sind äquivalent:*

(a) *K^n besitzt eine Orthogonalbasis aus Eigenvektoren für A.*
(b) *$V = V_{\lambda_1}(A) \oplus^\perp \ldots \oplus^\perp V_{\lambda_r}(A)$ für die Eigenwerte $\lambda_1, \ldots, \lambda_r$ von A.*
(c) *A ist normal und $\chi_A(X)$ zerfällt in $K[X]$ in Linearfaktoren.*
(d) *$A = A^t$ ist symmetrisch und $\chi_A(X)$ zerfällt in $K[X]$ in Linearfaktoren.*

Beweis. Die Aussagen (a) - (d) sind die entsprechenden Aussagen des Theorems 10.13 für $V = K^n$ mit dem Standardskalarprodukt und $f = L_A$ der Matrixmultiplikation mit A. Insbesondere ist die adjungierte Abbildungen durch Multiplikation mit der transponierten A^t gegeben. □

Bemerkung 10.17. Im Fall $K = \mathbb{R}$ und $V = \mathbb{R}^n$ mit Standardskalarprodukt zeigen wir im Beweis der Hauptachsentransformation, Theorem 12.1, daß die Bedingung an $\chi_A(X)$ in ein Produkt von Linearfaktoren zu zerfallen für eine symmetrische Matrix $A = A^t \in \mathrm{M}_n(\mathbb{R})$ automatisch erfüllt ist.

Übungsaufgaben zu §10

Übungsaufgabe 10.1. Sei $(V, \langle\ ,\ \rangle)$ ein endlich-dimensionaler Vektorraum mit einer perfekten symmetrischen Bilinearform. Sei $f : V \to V$ ein normaler Endomorphismus und $P(X) \in K[X]$ ein Polynom. Zeigen Sie, daß auch $P(f) : V \to V$ ein normaler Endomorphismus ist.

Übungsaufgabe 10.2. Sei $(V, \langle\ ,\ \rangle)$ ein endlich-dimensionaler Vektorraum mit einer anisotropen symmetrischen Bilinearform. Zu einem Endomorphismus $f : V \to V$ definieren wir die Bilinearform $\langle\ ,\ \rangle_f : V \times V \to K$ durch

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

für alle $v, w \in V$. Zeigen Sie die folgenden Aussagen.

(1) $\langle\ ,\ \rangle_f$ ist Bilinearform.
(2) $\langle\ ,\ \rangle_f$ ist symmetrisch genau dann, wenn f selbstadjungiert ist.
(3) Eine bezüglich $\langle\ ,\ \rangle$ orthogonale Basis aus Eigenvektoren von f ist dasselbe wie eine Basis die bezüglich $\langle\ ,\ \rangle$ und $\langle\ ,\ \rangle_f$ orthogonal ist.

11. Die Isometrie-Normalform

Eine Isometrie eines euklidischen Vektorraums ist ein normaler Endomorphismus. Der Spektralsatz, Theorem 10.13, greift aber nur für solche normalen Endomorphismen, deren charakteristisches Polynom über $\mathbb{R}$ in lineare Faktoren zerfällt. Da als Eigenwerte von Isometrien nur 1 und -1 in Frage kommen, gibt es für diagonalisierbare Isometrien nach Aussage des Spektralsatzes (nachdem man die Vektoren der Orthogonalbasis normiert hat) eine ONB $\mathscr{B} = (b_1, \ldots, b_n)$ für die $f(b_i) = \pm b_i$ gilt. Nach eventueller Umordnung gibt es dann $r, s \in \mathbb{N}_0$ mit $n = r + s$ und

$$\mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f) = \begin{pmatrix} \mathbf{1}_r & 0 \\ 0 & -\mathbf{1}_s \end{pmatrix},$$

wobei $\mathbf{1}_r$ und $\mathbf{1}_s$ die Einheitsmatrizen der entsprechenden Größe r und s sind. Dies ist ein Produkt der Spiegelungen $S_i = S_{b_i}$ an den Hyperebenen $\langle b_i\rangle^{\perp}$ für $i = r+1, \ldots, r+s = n$, wie man leicht nachrechnet. Dies ist eine sehr eingeschränkte Klasse von Isometrien, Produkte von Spiegelungen an Hyperebenen mit paarweise zueinander orthogonalen Normalenvektoren.

11.1. **Komplexifizierung reeller Vektorräume.** Sei V ein $\mathbb{R}$-Vektorraum. Die $\mathbb{R}$-Basis

$$1, i \in \mathbb{C}$$

führt zu einer inneren direkten Summe als $\mathbb{R}$-Vektorräume

$$V_{\mathbb{C}} := \mathbb{C} \otimes_{\mathbb{R}} V = V \oplus iV,$$

wobei wir V mit dem Bild der linearen Abbildung

$$V \to V_{\mathbb{C}}, \quad v \mapsto 1 \otimes v$$

identifizieren und mit iV das Bild der linearen Abbildung

$$V \to V_{\mathbb{C}}, \quad v \mapsto i \otimes v$$

bezeichnen. Jeder Vektor $w \in V_{\mathbb{C}}$ läßt sich somit eindeutig als

$$w = u + iv$$

mit $u, v \in V$ schreiben. Es operiert $\mathbb{C}$ auf $\mathbb{R}$-lineare Weise auf $V_{\mathbb{C}}$ wie folgt. Zu $w = u + iv$ mit $u, v \in V$ und $z = x + iy \in \mathbb{C}$ ist

$$zw = (x + iy)(u + iv) := (xu - yv) + i(xv + yu) \in V_{\mathbb{C}}.$$

Insbesondere gilt damit für $z, \zeta \in \mathbb{C}$ und $v \in V$

$$i \cdot v = iv$$
$$z(\zeta \otimes v) = z\zeta \otimes v.$$

Man rechnet leicht nach, daß damit $V_{\mathbb{C}}$ ein $\mathbb{C}$-Vektorraum wird. Eine Basis von $V_{\mathbb{C}}$ als $\mathbb{C}$-Vektorraum erhält man aus einer $\mathbb{R}$-Basis $(b_1, \ldots, b_n)$ von V, indem man das Tupel mittels $b_i = 1 \otimes b_i$ aufgefaßt in $V \subseteq V_{\mathbb{C}}$. Unter Mißbrauch der Notation schreiben wir diese komplexifizierte Basis weiterhin als

$$\mathscr{B} = (b_1 = 1 \otimes b_1, \ldots, b_n = 1 \otimes b_n).$$

Der Übergang von $\mathbb{R}$-Vektorräumen V zu $\mathbb{C}$-Vektorräumen $\mathbb{C}$ nennt man **Komplexifizierung**.

$$\{\mathbb{R}\text{-Vektorräume}\} \xrightarrow{\mathbb{C}\otimes -} \{\mathbb{C}\text{-Vektorräume}\}$$

Dies ist das einfachste Beispiel eines Basiswechsels (gemeint ist nicht die Basis eines Vektorraums, sondern der Bereich aus dem die Koordinaten stammen), wie man ihn in der kommutativen Algebra oder algebraischen Geometrie antrifft.

Beispiel 11.1. Die Komplexifizierung des Standard-$\mathbb{R}$-Vektorraums ist

$$(\mathbb{R}^n)_{\mathbb{C}} = \mathbb{C}^n.$$

Im allgemeinen Fall macht die Komplexifizierung dasselbe, nur diesmal koordinatenfrei.

Nach Vektorräumen komplexifizieren wir nun lineare Abbildungen. Sei $f : V \to W$ eine $\mathbb{R}$-lineare Abbildung. Die $\mathbb{R}$-bilineare Abbildung

$$\mathbb{C} \times V \to W_{\mathbb{C}}, \quad (z, v) \mapsto zf(v)$$

liefert über die universelle Eigenschaft des Tensorprodukts eine zunächst $\mathbb{R}$-lineare Abbildung

$$f_{\mathbb{C}} : V_{\mathbb{C}} \to W_{\mathbb{C}}, \quad f_{\mathbb{C}}(u + iv) = f(u) + if(v) \qquad \forall u, v \in V.$$

Aus dieser Formel ist unmittelbar klar, daß $f_{\mathbb{C}}$ sogar $\mathbb{C}$-linear ist. Für alle $z = a+bi \in \mathbb{C}$, $a, b \in \mathbb{R}$ und $w = u + iv \in V_{\mathbb{C}}$, $u, v \in V$ gilt

$$\begin{aligned} zf_{\mathbb{C}}(w) &= (a + bi)(f(u) + if(v)) = \big(af(u) - bf(v)\big) + i\big(af(v) + bf(u)\big) \\ &= f(au - bv) + if(av + bu) = f_{\mathbb{C}}(zw). \end{aligned}$$

Besonders einfach wird die Matrixbeschreibung von $f_{\mathbb{C}}$ bezüglich komplexifizierter Basen der $\mathbb{R}$-Vektorräume. Sei $\mathscr{B}$ eine Basis von V und $\mathscr{C}$ eine Basis von W, und bezeichnen wir mit derselben Notation die komplexifizierten Basen von $V_{\mathbb{C}}$ und $W_{\mathbb{C}}$, dann ist

$$\mathrm{M}^{\mathscr{B}}_{\mathscr{C}}(f_{\mathbb{C}}) = \mathrm{M}^{\mathscr{B}}_{\mathscr{C}}(f) \in \mathrm{M}_{n\times m}(\mathbb{R}) \subseteq \mathrm{M}_{n\times m}(\mathbb{C}),$$

wie man sofort aus der Definition der Darstellungsmatrizen nachrechnet. Sei $A = (a_{ij})$ die Darstellungsmatrix, also

$$f(b_j) = \sum_i a_{ij} c_i,$$

dann gilt

$$f_{\mathbb{C}}(b_j) = f_{\mathbb{C}}(b_j + i \cdot 0) = f(b_j) + if(0) = f(b_j) = \sum_i a_{ij} c_i.$$

11.2. **Der Spektralsatz für Isometrien.** Um den Spektralsatz auch für die anderen Isometrien zu bekommen, muß man die Aussage abschwächen und Drehkästchen erlauben: die nun zu behandelnde Isometrie–Normalform. Wie üblich bezeichnen wir mit $\mathbf{1}_m$ die Einheitsmatrix der Größe m und mit

$$D_\varphi = \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}$$

die Drehmatrix zum Winkel φ.

Theorem 11.2 (Isometrie–Normalform)**.** *Sei $f : V \to V$ eine Isometrie eines euklidischen Vektorraums V der Dimension n. Dann gibt es eine ONB $\mathscr{B}$ von V und eindeutige $r, s, t \in \mathbb{N}_0$ mit $n = r + s + 2t$ und eindeutig bestimmte Winkel $0 < \varphi_1 \leq \ldots \leq \varphi_t < \pi$, so daß die f darstellende Matrix die folgende Isometrie–Normalform annimmt:*

$$\mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f) = \begin{pmatrix} \mathbf{1}_r & & & & \\ & -\mathbf{1}_s & & & \\ & & D_{\varphi_1} & & \\ & & & \ddots & \\ & & & & D_{\varphi_t} \end{pmatrix}.$$

Korollar 11.3 (Isometrie–Normalform orthogonaler Matrizen)**.** *Sei $A \in \mathrm{O}(n)$ eine orthogonale Matrix. Dann gibt es eine orthogonale Matrix $S \in \mathrm{O}(n)$ und eindeutige $r, s, t \in \mathbb{N}_0$ mit $n = r + s + 2t$ und eindeutig bestimmte Winkel $0 < \varphi_1 \leq \ldots \leq \varphi_t < \pi$, so daß die konjugierte Matrix die folgende Isometrie–Normalform annimmt:*

$$SAS^{-1} = \begin{pmatrix} \mathbf{1}_r & & & & \\ & -\mathbf{1}_s & & & \\ & & D_{\varphi_1} & & \\ & & & \ddots & \\ & & & & D_{\varphi_t} \end{pmatrix}.$$

Beweis. Das ist nur die Matrixversion von Theorem 11.2, denn orthogonale Matrizen gehören nach Satz 9.17 zu Isometrien. Der Basiswechsel von der Standardbasis, einer ONB, in die ONB $\mathscr{B}$, die das Theorem 11.2 liefert, wird nach Satz 6.45 durch eine orthogonale Basiswechselmatrix vollzogen: die Matrix S ist orthogonal. □

Korollar 11.4. *Sei f wie in Theorem 11.2 bzw. A wie in Korollar 11.3. Dann ist das charakteristische Polynom*

$$\chi_f(X) = \chi_A(X) = (X-1)^r(X+1)^s \prod_{i=1}^{t}(X^2 - 2\cos(\varphi_i)X + 1)$$

und das Minimalpolynom enthält jeden irreduziblen Faktor des charakteristischen Polynoms genau einmal.

Beweis. Dies liest man sofort aus der Blockdiagonalform der Isometrie-Normalform ab. □

Bemerkung 11.5. Für $0 < \varphi < \pi$ ist $-1 < \cos(\varphi) < 1$, und das Polynom

$$X^2 - 2\cos(\varphi) + 1 \in \mathbb{R}[X]$$

ist irreduzibel. Die beiden komplexen Nullstellen $\cos(\varphi) \pm i\sin(\varphi)$ liegen wegen $\sin(\varphi) \neq 0$ nicht in $\mathbb{R}$.

Korollar 11.6. *Eine Isometrie des $\mathbb{R}^2$ ist eine Drehung oder eine Spiegelung.*

Beweis. Das haben wir bereits in Satz 9.7 bewiesen □

Korollar 11.7. *Jede orientierungserhaltende Isometrie des $\mathbb{R}^3$ ist eine Drehung.*

Beweis. Weil $\det(D_\varphi) = 1$ führt die Isometrienormalform mit Parametern wie im Theorem 11.2 zu $\det = (-1)^s$. Also ist $s = 0$ und $r = 1$ oder $r = 3$. In jedem Fall bleibt eine Drehung um den Winkel φ in einer Ebene $E \subseteq \mathbb{R}^3$ (bei $r = 3$ eben um den Winkel $\varphi = 0$). Orthogonal dazu in $E^\perp$ operiert die Isometrie durch den Eigenwert 1, also durch $\mathrm{id}_{E^\perp}$. □

Korollar 11.8. *(1) Jede orthogonale Matrix ist Produkt von Spiegelungen.*
(2) Die orthogonale Gruppe $\mathrm{O}(n)$ wird durch Spiegelungen erzeugt.
(3) Jede Isometrie des $\mathbb{R}^n$ ist ein Produkt von Spiegelungen.

Beweis. (1) Nach Theorem 11.2 reicht es, die Isometrie-Normalform als Produkt von Spiegelungen zu schreiben. Außerdem reicht es aus, jedes Drehkästchen einzeln als Produkt von zwei Spiegelungen zu schreiben und dann aufzumultiplizieren. Dies gelingt nach Korollar 9.9.

Die Aussagen (2) und (3) sind triviale Folgen von Aussage (1). □

Beweis von Theorem 11.2. Die Eindeutigkeit folgt sofort aus Korollar 11.4, in dessen Beweis die Eindeutigkeitsaussage nicht verwendet wird. Es ist r (bzw. s) die Vielfachheit der Nullstelle 1 (bzw. -1) von $\chi_f(X)$ und die φ_i sind eindeutig durch die quadratischen irreduziblen Faktoren von $\chi_f(X)$ bestimmt, weil für $0 < \varphi < \pi$ der Cosinus streng monoton fallend in den Bereich in $1 > \cos(\varphi) > -1$ abbildet. Hier benötigen wir noch die Überlegung, daß

$$X^2 - 2cX + 1 \in \mathbb{R}[X]$$

für $|c| < 1$ irreduzibel ist, weil die Nullstellen $c \pm i\sqrt{1-c^2}$ in $\mathbb{C}$ aber nicht in $\mathbb{R}$ liegen.

Der Beweis der Existenz läuft parallel zum Beweis des Spektralsatzes für zerfallende normale Operatoren, Theorem 10.13. Wir arbeiten per Induktion nach $n = \dim(V)$. Der Induktionsanfang $n = 0$ ist trivial. Wir überlegen uns die Fälle $n = 1$ und $n = 2$ ebenfalls, da der Induktionsschritt wesentlich darauf basiert.

Schritt 1: Dimension 1. Für $n = 1$ haben die Isometrien einen Eigenvektor und damit nach Proposition 9.21 den Eigenwert ± 1. Damit gilt die Isometrienormalform.

Schritt 2: Dimension 2. Das ist gerade Satz 9.7.

Schritt 3: Ein invarianter Summand. Wir nehmen an, daß $n > 0$ ist und die Aussage von Theorem 11.2 für Isometrien von euklidischen Vektorräumen der Dimension $< n$ gilt. Sei

$$f_{\mathbb{C}} : V_{\mathbb{C}} \to V_{\mathbb{C}}$$

die Komplexifizierung von f. Weil in $\mathbb{C}$ nach dem Fundamentalsatz der Algebra (Theorem **??**) jedes Polynom eine Nullstelle besitzt, hat $f_{\mathbb{C}}$ einen Eigenwert $\lambda = x + iy$ mit einem Eigenvektor $0 \neq u + iv \in V_{\mathbb{C}}$, d.h.,

$$f(u) + if(v) = f(u+iv) = \lambda(u+iv) = (xu - yv) + i(xv + yu).$$

Die lineare Hülle $U = \langle u, v\rangle_{\mathbb{R}} \subseteq V$ hat Dimension 1 oder 2, und sie ist f-invariant (Koeffizientenvergleich):

$$f(u) = xu - yv \in U \quad \text{und} \quad f(v) = xv + yu \in U.$$

Außerdem ist U auch $f^* = f^{-1}$ invariant, denn f ist bijektiv (die reellen Eigenwerte sind 1 oder -1; die 0 tritt nicht als Eigenwert auf, siehe Proposition 9.21).

Schritt 4: Ein invariante orthogonale Zerlegung. Wir setzen $W = U^{\perp}$ und betrachten die orthogonale Zerlegung

$$V = U \oplus^{\perp} W.$$

Dann ist W auch f-invariant: zu $w \in W$ und $u \in U$ folgt $f^*(u) \in U$ und somit

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

Damit gilt $f(w) \in U^{\perp} = W$.

Schritt 5: Zusammenfügen. Die Einschränkungen $f|_U$ und $f|_W$ von f auf U und W sind weiterhin Isometrien, jetzt bezüglich der Einschränkung des Skalarprodukts von V auf U bzw. auf W (was ein Skalarprodukt bleibt!). Per Induktionsannahme hat $f|_W$ die verlangte Form, und $f|_U$ wurde als Induktionsanfang behandelt. Sei $\mathscr{C}$ (bzw. $\mathscr{D}$) eine Basis von U (bzw. W), so daß $f|_U$ (bzw. $f|_W$ in Isometrie-Normalform erscheint. Sei $\mathscr{B}$ die Basis von V, die durch zusammenfügen von $\mathscr{C}$ und $\mathscr{D}$ entsteht. Dann hat f wegen

$$\mathrm{M}_{\mathscr{B}}^{\mathscr{B}}(f) = \begin{pmatrix} \mathrm{M}_{\mathscr{C}}^{\mathscr{C}}(f|_U) & \\ & \mathrm{M}_{\mathscr{D}}^{\mathscr{D}}(f|_W) \end{pmatrix}$$

bezüglich $\mathscr{B}$ eine Matrix mit den erlaubten Block-Diagonaleinträgen. Eventuelles umsortieren der gefundenen Basis bring die Kästchen der Matrix in die gewünschte Reihenfolge der Isometrie-Normalform. $\square$

Bemerkung 11.9. Sei V ein euklidischer Vektorraum und $f \in \mathrm{End}_{\mathbb{R}}(V)$. Sei $A \in \mathrm{M}_n(\mathbb{R})$ die Matrix von f bezüglich einer ONB von V. (Oder man startet gleich mit einem A; das ist der Fall $f =$ Multiplikation mit A auf $V = \mathbb{R}^n$ mit Standardskalarprodukt.)

Es ist dann f eine Isometrie genau dann, wenn A orthogonale Matrix ist, und das testet man durch den Nachweis von

$$A^t A = \mathbf{1}.$$

Sei dies erfüllt, dann findet man die Isometrie-Normalform durch Faktorisieren des charakteristischen Polynoms. Nach Korollar 11.4 treten die Nullstelle ± 1 mit der Vielfachheit auf, in der ± 1 als Diagonaleintrag in der Isometrie-Normalform auftritt. Und jeder quadratische Faktor ist von der Form

$$X^2 - 2\cos(\varphi)X + 1,$$

woraus der eindeutige Drehwinkel $\varphi \in (0, \pi)$ des zugehörigen Kästchens bestimmt werden kann. Die Winkel 0 und π treten nicht auf, da dann der quadratische Faktor nicht irreduzibel ist.

Die geeignete Basis $\mathscr{B}$ zu finden, in der f bzw. A Isometrie-Normalform annimmt, ist aufwendiger, aber auch aus dem Beweis von Theorem 11.2 herauszulesen. Man bestimme Basen zu den Eigenräumen zum Eigenwert ± 1 und auch zu den Eigenwerten $\lambda \in \mathbb{C} \setminus \mathbb{R}$ der Komplexifizierung $f_{\mathbb{C}}$. Man bekommt dadurch f-invariante Unterräume $U \subseteq V$ der Dimension 1 (reeller Eigenwert) oder 2 (komplexer, nicht-reeller Eigenwert). Im Fall $\dim(U) = 2$ muß man noch eine ONB von U bestimmen.

Am besten arbeitet man jeweils mit dem orthogonalen Komplement des bereits bestimmten Teils der Basis $\mathscr{B}$. Dadurch verringert sich die Dimension, und nach endlich vielen Schritten ist man fertig.

12. Quadriken und die Hauptachsentransformation

12.1. **Die Hauptachsentransformation.** Ziel ist der folgende geometrische Satz und seine drei unmittelbaren Korollare. Diese vier Aussagen sind relativ schnell ineinander übersetzt und somit äquivalent. Es reicht eine davon zu beweisen.

Theorem 12.1 (Hauptachsentransformation)**.** *Sei $A \in \mathrm{M}_n(\mathbb{R})$ eine symmetrische Matrix. Dann gibt es $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ und eine orthogonale Matrix $S \in \mathrm{O}(n)$, so daß*

$$S^t A S = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Bemerkung 12.2. Sei $\langle\ ,\ \rangle_A$ die symmetrische Bilinearform auf $\mathbb{R}^n$ mit Gram'scher Matrix A. Dann besagt Theorem 12.1, daß es einen orthogonalen Basiswechsel mittels $S \in \mathrm{O}(n)$ gibt in eine ONB $\mathscr{B}$ von $\mathbb{R}^n$, die auch für $\langle\ ,\ \rangle_A$ orthogonal ist.

Korollar 12.3. *Sei V ein euklidischer Vektorraum und $f : V \times V \to \mathbb{R}$ eine symmetrische Bilinearform. Dann gibt es eine ONB, die bezüglich f eine Orthogonalbasis ist.*

Beweis. Wir wählen eine erste ONB $\mathscr{B}$ beliebig. Die Gram'sche Matrix $A = \mathrm{M}^{\mathscr{B},\mathscr{B}}(f)$ ist eine symmetrische reelle Matrix, weil f symmetrisch ist. Nach Theorem 12.1 gibt es eine orthogonale Matrix S, so daß $S^t A S$ Diagonalgestalt hat. Dieses S ist die Basiswechselmatrix in eine Orthogonalbasis $\mathscr{C}$ bezüglich f. Als orthogonale Matrix wechselt S wieder in eine ONB nach Satz 6.45. $\square$

Wir übersetzen nun die Aussagen über eine zweite Bilinearform in Aussagen über einen Endomorphismus. Dabei benutzen wir wesentlich, daß für eine orthogonale Matrix S die Formel $S^{-1} = S^t$ gilt, um via

$$S^t AS = S^{-1}AS$$

die Basiswechselformel für die Darstellungsmatrix eines Endomorphismus in die Basiswechselformel für eine Gram'sche Matrix zu übersetzen.

Korollar 12.4. *Sei $A \in \mathrm{M}_n(\mathbb{R})$ eine symmetrische Matrix. Dann ist A durch eine orthogonale Matrix diagonalisierbar: es gibt $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ und eine orthogonale Matrix $S \in \mathrm{O}(n)$, so daß*

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Beweis. Das ist wegen $S^{-1} = S^t$ für alle $S \in \mathrm{O}(n)$ genau Theorem 12.1. □

Korollar 12.5. *Sei V ein euklidischer Vektorraum und $f : V \to V$ ein selbstadjungierter Endomorphismus. Dann gibt es eine ONB aus Eigenvektoren von f.*

Beweis. Wir wählen eine erste ONB $\mathscr{B}$ beliebig. Nach Propositon 6.42 und wegen $f = f^*$ ist dann die darstellende Matrix

$$A = \mathrm{M}^{\mathscr{B}}_{\mathscr{B}}(f) = \mathrm{M}^{\mathscr{B}}_{\mathscr{B}}(f^*)^t = \mathrm{M}^{\mathscr{B}}_{\mathscr{B}}(f)^t = A^t$$

symmetrisch. Nach Theorem 12.1 gibt es eine orthogonale Matrix S, so daß $S^{-1}AS = S^t AS$ Diagonalgestalt hat. Dieses S ist die Basiswechselmatrix in eine Basis $\mathscr{C}$ von Eigenvektoren des Endomorphismus f. Als orthogonale Matrix wechselt S wieder in eine ONB nach Satz 6.45. □

Bemerkung 12.6. (1) Der Diagonalformensatz Theorem 6.12 diagonalisiert symmetrische Bilinearformen über einem beliebigen Körper mit $2 \in K^\times$. Der auftretende Basiswechsel ist aus $\mathrm{GL}_n(K)$.

Im Vergleich dazu studiert die Hauptachsentransformation, Theorem 12.1, vor dem Hintergrund eines euklidischen Vektorraums eine weitere Bilinearform $\langle\ ,\ \rangle_A$, die mittels bezüglich der euklidischen Struktur orthogonalem Basiswechsel (im Fall von $\mathbb{R}^n$ aus $\mathrm{O}(n)$) in eine Diagonalform gebracht wird.

Diagonalformensatz		**Hauptachsentransformation**
eine symmetrische Bilinearform	$\longleftrightarrow$	zwei symmetrische Bilinearformen eine davon Skalarprodukt
$2 \in K^\times$		über $\mathbb{R}$
Basiswechsel aus $\mathrm{GL}_n(K)$		Basiswechsel aus $\mathrm{O}_n(\mathbb{R})$

(2) Betrachten wir eine Matrix $A \in \mathrm{M}_n(\mathbb{R})$ als Endomorphsimus von $\mathbb{R}^n$, oder besser als Matrix bezüglich einer Basis $\mathscr{B}$ eines Endomorphismus $f \in \mathrm{End}_{\mathbb{R}}(V)$ für einen $\mathbb{R}$-Verktorraum V, dann hat der Basiswechsel zu einer Basis $\mathscr{C}$ den Effekt

$$A = \mathrm{M}^{\mathscr{B}}_{\mathscr{B}}(f) \rightsquigarrow S^{-1}AS = \mathrm{M}^{\mathscr{C}}_{\mathscr{C}}(f)$$

mit $S = \mathrm{M}^{\mathscr{C}}_{\mathscr{B}}(\mathrm{id}_V)$. Faßt man hingegen A als Gram'sche Matrix einer Bilinearform $\langle\ ,\ \rangle$ auf V zur Basis $\mathscr{B}$ auf, so hat der Basiswechsel zu $\mathscr{C}$ den Effekt

$$A = \mathrm{M}^{\mathscr{B},\mathscr{B}}(\langle\ ,\ \rangle) \rightsquigarrow S^t AS = \mathrm{M}^{\mathscr{C},\mathscr{C}}(\langle\ ,\ \rangle).$$

Unter einem orthogonalen Basiswechsel, und nur für S orthogonal, stimmen wegen

$$S^{-1} = S^t \in \mathrm{O}(n)$$

die beiden Transformationen überein! Ein Normalformensatz für einen Endomorphismus übersetzt sich in einen Normalformensatz für eine Bilinearform und umgekehrt unter der Voraussetzung, daß von einem orthogonalen Basiswechsel die Rede ist.

Korollar 12.7. *Das charakteristisches Polynom einer reellen symmetrischen Matrix zerfällt in* $\mathbb{R}[X]$ *in Linearfaktoren.*

Beweis. Sei A symmetrisch und S wie in Theorem 12.1. Die Matrizen A und $S^{-1}AS$ haben dasselbe charakteristische Polynom, und für Diagonalmatrizen zerfällt es in Linearfaktoren. □

Korollar 12.8. *Eine reelle symmetrische Matrix hat nur reelle Eigenwerte.*

Beweis. Das folgt sofort aus Korollar 12.7 und dem Zusammenhang zwischen Eigenwerten und Nullstellen des charakteristischen Polynoms. □

Korollar 12.9. *Sei V ein euklidischer Vektorraum. Das charakteristisches Polynom eines selbstadjungierten Endomorphismus $f : V \to V$ zerfällt in $\mathbb{R}[X]$ in Linearfaktoren.*

Beweis. Das folgt nach Einführung von Koordinaten sofort aus Korollar 12.7. □

Korollar 12.10. *Sei V ein euklidischer Vektorraum. Ein selbstadjungierter Endomorphismus $f : V \to V$ hat nur reelle Eigenwerte.*

Beweis. Das folgt sofort aus Korollar 12.9 und dem Zusammenhang zwischen Eigenwerten und Nullstellen des charakteristischen Polynoms. □

12.2. **Beweis der Hauptachsentransformation.** Wir beweisen nun das Theorem 12.1 über die Hauptachsentransformation.

Beweis. Eine symmetrische Matrix $A \in \mathrm{M}_n(\mathbb{R})$ kommutiert mit $A^t = A$ und vermittelt durch Matrixmultiplikation daher einen normalen Endomorphismus von $\mathbb{R}^n$ ausgestattet mit dem Standardskalarprodukt. Das Standardskalarprodukt ist anisotrop, so daß zur Anwendung des Spektralsatzes für zerfallende normale Operatoren, Theorem 10.13 nur noch die Aussage von Korollar 12.7 fehlt. Wir haben zwar dieses Korollar aus dem Theorem der Hauptachsentransformation abgeleitet, stellen aber nun fest, daß nach der Vorarbeit aus Theorem 10.13 genau dieses Korollar 12.7 der entscheidende zu beweisende Schritt ist. Wir müssen daher für das Korollar einen alternativen Beweis finden. Weil wir per Induktion arbeiten werden, reicht die Aussage des Korollars 12.8.

Der Beweis von Korollar 12.8 benutzt ein wenig Analysis. Sei $A \in \mathrm{M}_n(\mathbb{R})$ symmetrisch. Die $(n-1)$-Sphäre

$$\mathrm{S}^{n-1} = \{v \in \mathbb{R}^n \ ; \ \|v\| = 1\}$$

ist abgeschlossen und beschränkt, also kompakt und daher nimmt die stetige Funktion

$$\begin{aligned} f : \mathrm{S}^{n-1} &\to \mathbb{R} \\ v &\mapsto \langle v, Av\rangle \end{aligned}$$

auf S^{n-1} sein Supremum als Maximum an. Das ist ein Fakt aus der Analysis.

Sei also $v_0 \in \mathrm{S}^{n-1}$ normiert und für alle $w \in \mathrm{S}^{n-1}$

$$\langle v_0, Av_0\rangle \geq \langle w, Aw\rangle.$$

Wir behaupten, daß v_0 ein Eigenvektor zum Eigenwert $\lambda_0 = \langle v_0, Av_0\rangle$ ist. Dazu nutzen wir aus, daß für alle $w \in \mathbb{R}^n$ die Funktion

$$F_w(t) = \frac{\langle v_0 + tw, A(v_0 + tw)\rangle}{\|v_0 + tw\|^2} = f(\frac{1}{\|v_0 + tw\|}(v_0 + tw))$$

in einer Umgebung von 0 stetig differenzierbar ist und in $t = 0$ ein Maximum hat. Die Ableitung bei $t = 0$ ist demnach

$$\begin{aligned} 0 = F_w'(0) &= \frac{d}{dt}\Big|_{t=0} \frac{\langle v_0, Av_0\rangle + t\big(\langle w, Av_0\rangle + \langle v_0, Aw\rangle\big) + t^2\langle w, Aw\rangle}{1 + t\big(\langle w, v_0\rangle + \langle v_0, w\rangle\big) + t^2\langle w, w\rangle} \\ &= \langle w, Av_0\rangle + \langle v_0, Aw\rangle - \langle v_0, Av_0\rangle \cdot \big(\langle w, v_0\rangle + \langle v_0, w\rangle\big) \\ &= \langle w, Av_0\rangle + \langle Av_0, w\rangle - \lambda_0 \cdot \big(2\langle w, v_0\rangle\big) \qquad (A \text{ und } \langle\ ,\ \rangle \text{ symmetrisch}) \\ &= 2 \cdot \big(\langle w, Av_0\rangle - \lambda_0 \cdot \langle w, v_0\rangle\big) = 2\langle w, Av_0 - \lambda_0 v_0\rangle. \end{aligned}$$

Es gilt also für alle $w \in \mathbb{R}^n$

$$\langle w, Av_0 - \lambda_0 v_0\rangle = 0.$$

Da das Standardskalarprodukt nichtausgeartet ist, folgt

$$Av_0 - \lambda_0 v_0 = 0$$

und damit die Behauptung, daß v_0 Eigenvektor zum Eigenwert λ_0 ist. Damit ist gezeigt, daß jede relle symmetrische Matrix einen reellen Eigenwert hat.

Wir zeigen nun Theorem 12.1 per Induktion nach der Dimension n. Wie im Beweis des Spektralsatzes Theorem 10.13 gehen wir nun zum orthogonalen Komplement des Eigenvektors v_0 über. Wir ergänzen v_0 mittels Satz 8.18 zu einer ONB und betrachten die Matrix T, deren Spalten genau aus dieser ONB bestehen, die erste aus v_0. Dann erhalten wir wegen $Av_0 = \lambda_0 v_0$ eine Blockform

$$A' = T^{-1}AT = T^tAT = \left(\begin{array}{c|ccc} \lambda_0 & \beta_2 & \cdots & \beta_n \\ \hline 0 & & & \\ \vdots & & B & \\ 0 & & & \end{array}\right)$$

mit $B \in \mathrm{M}_{n-1}(\mathbb{R})$ und $\beta_2, \ldots, \beta_n \in \mathbb{R}$. Es ist A' symmetrisch:

$$A'^t = (T^tAT)^t = T^tA^t(T^t)^t = T^tAT = A'.$$

Daher gilt $B^t = B$ und $\beta_2 = \ldots = \beta_n = 0$. Nach Induktionsvoraussetzung gibt es für B eine orthogonale Matrix R, so daß R^tBR diagonal ist. Dann ist auch

$$S = T\begin{pmatrix} 1 & \\ & R \end{pmatrix}$$

orthogonal und

$$S^tAS = \begin{pmatrix} 1 & \\ & R^t \end{pmatrix}(T^tAT)\begin{pmatrix} 1 & \\ & R \end{pmatrix} = \begin{pmatrix} 1 & \\ & R^t \end{pmatrix}\begin{pmatrix} \lambda_0 & \\ & B \end{pmatrix}\begin{pmatrix} 1 & \\ & R \end{pmatrix} = \begin{pmatrix} \lambda_0 & \\ & R^tBR \end{pmatrix}$$

hat Diagonalgestalt. □

12.3. **Quadratische Formen.**

Definition 12.11. Eine **quadratische Form** auf einem K-Vektorraum V ist eine Abbildung

$$q : V \to K$$

mit den folgenden beiden Eigenschaften:
(i) $q(\lambda v) = \lambda^2 q(v)$ für alle $\lambda \in K$ und $v \in V$.
(ii) Die Abbildung $f : V \times V \to K$ definiert eine Bilinearform durch

$$f(v, w) = q(v + w) - q(v) - q(w).$$

Proposition 12.12. *Sei K ein Körper mit $2 \in K^\times$. Sei V ein K-Vektorraum. Dann sind die folgenden Konstruktionen zueinander inverse Bijektionen zwischen symmetrischen Bilinearformen und quadratischen Formen:*

(1) Zu einer symmetrischen Bilinearform $f : V \times V \to K$ *haben wir eine quadratische Form*
$$q(v) = f(v,v) \quad \textit{für alle } v \in V.$$
(2) Zu einer quadratischen Form $q : V \to K$ *haben wir eine symmetrische Bilinearform*
$$f(v,w) = \frac{1}{2}\big(q(v+w) - q(v) - q(w)\big) \quad \textit{für alle } v, w \in V.$$

Beweis. Wenn f eine Bilinearform ist, dann setzen wir $q(v) = f(v,v)$. Für alle $\lambda \in K$ gilt dann
$$q(\lambda v) = f(\lambda v, \lambda v) = \lambda^2 f(v,v) = \lambda^2 q(v),$$
und für alle $v, w \in V$ nach der Polarisationsformel, Satz 5.16,
$$\frac{1}{2}\big(q(v+w) - q(v) - q(w)\big) = \frac{1}{2}\big(f(v+w, v+w) - f(v,v) - f(w,w)\big) = f(v,w).$$
Also ist $q : V \to K$ eine quadratische Form.

Sei umgekehrt $q : V \to K$ eine quadratische Form. Dann ist f wie in (2) per Definition eine symmetrische Bilinearform.

Bleibt zu zeigen, daß die beiden Konstruktionen invers zueinander sind. Das folgt, wenn wir mit einer symmetrischen Bilinearform f starten, wie oben bereits gerechnet aus der Polarisationsformel, Satz 5.16. Zum andern, wenn wir mit einer quadratischen Form q starten, dann gilt für alle $v \in V$ für das zugehörige $f(v,w) = q(v+w) - q(v) - q(w)$:
$$f(v,v) = \frac{1}{2}\big(q(2v) - 2q(v)\big) = \frac{1}{2}\big(4q(v) - 2q(v)\big) = q(v).$$ □

Eine quadratische Form auf V wird in Koordinaten durch eine quadratische Form in $\dim_K(V)$-vielen Variablen beschrieben.

Definition 12.13. Eine **quadratische Form in** n **Variablen** $X_1, \ldots, X_n$ mit Koeffizienten aus einem Körper K ist ein homogenes Polynom vom Grad 2, also ein Ausdruck der Form
$$Q(X_1, \ldots, X_n) = \sum_{1 \le i \le j \le n} a_{ij} X_i X_j.$$
mit $a_{ij} \in K$ für alle $1 \le i \le j \le n$.

Sei $q : V \to K$ eine quadratische Form und $f(v,w)$ die zugehörigen symmetrische Bilinearform nach Proposition 12.12. In Koordinaten bezüglich einer Basis $\mathscr{B} = (b_1, \ldots, b_n)$ von V und der Gram'schen Matrix $A = \mathrm{M}^{\mathscr{B},\mathscr{B}}(f) = (a_{ij})$ berechnet sich bei $v = \sum_{i=1}^n x_i b_i$
$$q(v) = f(v,v) = \sum_{i=1}^n \sum_{j=1}^n x_i x_j f(b_i, b_j) = \sum_{i,j} a_{ij} x_i x_j.$$
Dies ist die Auswertung in $x_1, \ldots, x_n$ der quadratischen Form in n Variablen
$$Q(X_1, \ldots, X_n) = \langle \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}, \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \rangle_A = (X_1, \ldots, X_n) A \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \sum_{i,j} a_{ij} X_i X_j.$$

Hat man umgekehrt eine quadratische Form in n Variablen
$$Q(X_1, \ldots, X_n) = \sum_{1 \le i \le j \le n} a_{ij} X_i X_j.$$
gegeben, dann definieren wir zunächst $a_{ij} = 0$, wenn $i > j$ und setzen dann
$$A = \left(\frac{a_{ij} + a_{ji}}{2}\right)_{1 \le i,j \le n} \in \mathrm{M}_n(K).$$

Auf der Diagonale passiert nichts, aber die Einträge a_{ij} werden jeweils zur Hälte auf die ij und die ji-Position verteilt. Dann ist

$$A = A^t$$

symmetrisch und die Quadratische Form $q : K^n \to K$, die nach Propositon 12.12 der symmetischen Bilinearform $f(x,y) = x^t A y$ zugeordnet ist, hat bei $x = (x_1, \ldots, x_n)^t$ den Wert:

$$q(x) = x^t A x = \sum_{ij} x_i x_j (\frac{a_{ij} + a_{ji}}{2}) = Q(x_1, \ldots, x_n).$$

Quadratische Formen in n Variablen sind also nichts anderes als die Formel für die Koordinatenbeschreibung einer quadratischen Form auf einem n-dimensionalen K-Vektorraum. Und dies ist via Propositon 12.12 beschrieben durch eine symmetrische Bilinearform, also in Koordinaten wieder über eine symmetrische Matrix. Diese Übersetzungen benötigen $2 \in K^\times$.

12.4. **Diagonalgestalt für quadratische Formen.** Der Diagonalformensatz Theorem 6.12 hat über Korollar 6.14 eine Anwendung auf quadratische Formen.

Proposition 12.14. *Sei $2 \in K^\times$. Sei $Q(X_1, \ldots, X_n) = \sum_{i \leq j} a_{ij} X_i X_j \in K[X_1, \ldots, X_n]$ eine quadratische Form in n Variablen. Dann gibt es eine lineare invertierbare Variablensubstitution*

$$U_i = \sum_{j=1}^{n} s_{ij} X_j$$

mit der Matrix $S = (s_{ij}) \in \mathrm{GL}_n(K)$, so daß q in den Variablen U_i Diagonalgestalt annimmt, d.h. es gibt $\lambda_i \in K$ mit

$$Q(X_1, \ldots, X_n) = \lambda_1 U_1^2 + \ldots + \lambda_n U_n^2.$$

Beweis. Sei A die symmetrische Matrix zur quadratischen Form q. Wie in Korollar 6.14 schreiben wir mit einer Diagonalmatrix $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \in \mathrm{M}_n(K)$ und $S \in \mathrm{GL}_n(K)$

$$A = S^t D S.$$

Wir setzen

$$\vec{U} := \begin{pmatrix} U_1 \\ \vdots \\ U_n \end{pmatrix} = S \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = S\vec{X},$$

und rechnen

$$Q(X_1, \ldots, X_n) = \vec{X}^t A \vec{X} = \vec{X}^t S^t D S \vec{X} = (S\vec{X})^t D (S\vec{X}) = \vec{U}^t D \vec{U} = \lambda_1 U_1^2 + \ldots + \lambda_n U_n^2. \quad \square$$

Bemerkung 12.15. Proposition 12.14 beschreibt das multidimensionale quadratische Ergänzen. Der Fall $n = 2$ läßt sich als eindimensionales quadratisches Ergänzen verstehen: sei

$$Q(X,Y) = aX^2 + bXY + cZ^2$$

gegeben, und sei der Einfachheit halber $a \neq 0$. Dann setzen wir formal $T = X/Y$ und betrachten zunächst das quadratische Polynom in T

$$Q(X,Y)/Y^2 = Q(X/Y, 1) = aT^2 + bT + c = a(T + \frac{b}{2a})^2 + \frac{4ac - b^2}{4a}.$$

Zurückübersetzt ergibt sich

$$Q(X,Y) = Q(T,1) \cdot Y^2 = a(X + \frac{b}{2a} Y)^2 + \frac{4ac - b^2}{4a} Y^2,$$

also in der Variablen $U = X + \frac{b}{2a} Y$ und $V = Y$ eine diagonale Form.

Bemerkung 12.16. Sei Q eine quadratische Form in n-Variablen $X_1, \ldots, X_n$ mit zugehöriger symmetrischer Matrix $A \in \mathrm{M}_n(K)$. Eine lineare Variablentransformation, die Q in Diagonalgestalt bringt, berechnet man wie folgt. Zur Vereinfachung nehmen wir an, daß A eine invertierbare Matrix ist.

Schritt 1: Man bestimmt eine Orthogonalbasis $\mathscr{B} = (b_1, \ldots, b_n)$ bezüglich $\langle\ ,\ \rangle_A$ auf K^n. Dazu startet man mit einer beliebigen Basis und wendet das Gram-Schmidt Verfahren aus Satz 6.20 an. Da man zu Beginn nicht weiß, daß $\langle\ ,\ \rangle_A$ anisotrop ist, hat man keine Garantie, daß das Gram-Schmidt Verfahren durchführbar ist. Die Chancen sind aber mehr als gut (falls $K = \mathbb{R}$ haben die Ausnahmen Maß 0), und wenn es doch Probleme gibt, verfährt man wie im Beweis von Theorem 6.12.

Schritt 2: Gesucht ist die Basiswechselmatrix $S = \mathrm{M}^{\mathscr{E}}_{\mathscr{B}}(\mathrm{id})$. Leichter ist allerdings

$$S^{-1} = \mathrm{M}^{\mathscr{B}}_{\mathscr{E}}(\mathrm{id}) = [b_1, \ldots, b_n],$$

die Matrix mit Spalten $b_1, \ldots, b_n$.

Schritt 3: Diese Matrix $T = S^{-1}$ muß nun invertiert werden. Aufgrund der speziellen Form geht das leichter als üblich. Da die Spalten paarweise orthogonal bezüglich $\langle\ ,\ \rangle_A$ sind, gilt

$$\begin{aligned} T^t A T &= T^t(Ab_1, \ldots, Ab_n) = (\langle b_i, Ab_j\rangle)_{1\le i,j\le n} \\ &= (\langle b_i, b_j\rangle_A)_{1\le i,j\le n} = \begin{pmatrix} \langle b_1, b_1\rangle_A & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \langle b_n, b_n\rangle_A \end{pmatrix} =: D. \end{aligned}$$

Da A als invertierbar angenommen wurde ist auch $D = T^t A T$ invertierbar. Damit ist

$$S = T^{-1} = D^{-1} T^t A$$

und anstelle von T ist nur die Diagonalmatrix D zu invertieren und zwei Matrixmultiplikationen sind auszuführen. Das ist viel einfacher.

Schritt 4: Die neuen Variablen sind nun

$$U_i = \sum_{j=1}^n S_{ij} X_j$$

und die quadratische Form nimmt die folgende Gestalt an:

$$Q(U_1, \ldots, U_n) = \sum_{i=1}^n \langle b_i, b_i\rangle_A \cdot U_i^2.$$

12.5. **Anwendung der Hauptachsentransformation auf Quadriken.** Eine quadratische Form $q(X_1, \ldots, X_n)$ ist nichts anderes als die zweite symmetrische Bilinearform $\langle v, v\rangle_A$ aus Korollar 12.2 geschrieben als formaler Ausdruck in Variablen $X_1, \ldots, X_n$ als Koeffizienten für den eingesetzen Vektor v, also ausgewertet für

$$v = X_1 e_1 + \ldots + X_n e_n.$$

Daher ergibt sich sofort die folgende Umformulierung.

Korollar 12.17. *Zu einer quadratischen Form auf* $\mathbb{R}^n$

$$q(X_1, \ldots, X_n)$$

gibt es eine ONB von $\mathbb{R}^n$ *mit Koordinaten* $U_1, \ldots, U_n$ *und* $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$*, so daß die quadratische Form die Gestalt*

$$\lambda_1 U_1^2 + \ldots + \lambda_n U_n^2$$

annimmt.

Bemerkung 12.18. Die folgenden Bilder illustrieren den Namen Hauptachsentransformation. Zu $r \in \mathbb{R}$ und einer quadratischen Form $q(X,Y) = X^t AY$ betrachten wir die Lösungsmenge

$$V(q(X,Y) - r) := \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \ ; \ q(x,y) - r = 0 \right\} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \ ; \ x^t Ay = r \right\}.$$

Eine solche wird **ebene Quadrik** oder Kegelschnitt genannt (dazu mehr in Abschnitt 12.6). Die Gestalt einer ebenen Quadrik hängt von der Signatur der zur quadratischen Form gehörenden symmetrischen Bilinearform bzw. deren Gram'scher Matrix A und vom Vorzeichen des Werts r ab. Die eingezeichneten Achsen sind die Hauptachsen, das sind die Koordinatenachsen der ONB bezüglich derer die quadratische Form eine Linearkombination von Quadraten der Koordinaten wird.

Beispiel 12.19. Wenn A positiv definit (Signatur $(2,0)$) und r positiv ist, so entsteht eine Ellipse.

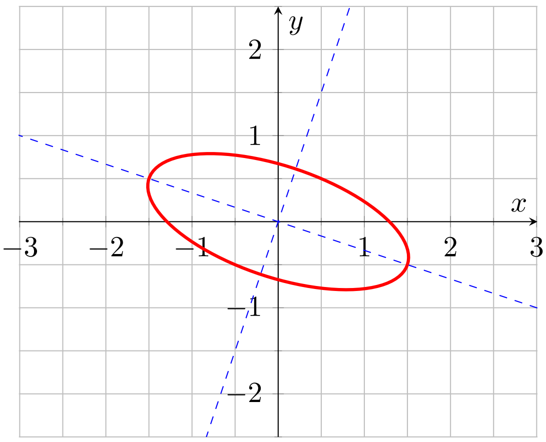


ABBILDUNG 14. Die Quadrik $3X^2 + 6XY + 11Y^2 = 5$.

Hier ist $r = 5$ und

$$A = \begin{pmatrix} 3 & 3 \\ 3 & 11 \end{pmatrix}$$

nach dem Hauptminorenkriterium positiv definit. Das charakteristische Polynom ist

$$(X-3)(X-11) - 9 = X^2 - 14X + 24 = (X-7)^2 - 25,$$

so daß die Eigenwerte 2 und 12 sind mit zugehörigen Eigenvektoren

$$\begin{pmatrix} 3 \\ -1 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

In diese orthogonalen Richtungen laufen die Symmetrieachsen (Hauptachsen) der Ellipse. Die orthogonale Transformationsmatrix S bekommt man als transponierte zu

$$S^t = S^{-1} = (\text{Spalten: Basis aus normierten Eigenvektoren von } A) = 1/\sqrt{10} \begin{pmatrix} 3 & 1 \\ -1 & 3 \end{pmatrix}$$

Die Matrix A wird durch die Transformation mit S diagonalisiert:

$$SAS^{-1} = \frac{1}{10} \begin{pmatrix} 3 & -1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 3 & 3 \\ 3 & 11 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ -1 & 3 \end{pmatrix} = \begin{pmatrix} 2 & \\ & 12 \end{pmatrix}.$$

Es sind X, Y die Koordinatenvariablen bezüglich der Standardbasis, und U, V diejenigen bezüglich der ONB aus Eigenvektoren

$$(1/\sqrt{10} \begin{pmatrix} 3 \\ -1 \end{pmatrix}, 1/\sqrt{10} \begin{pmatrix} 1 \\ 3 \end{pmatrix}),$$

dann gilt

$$\begin{pmatrix} X \\ Y \end{pmatrix} = U \cdot 1/\sqrt{10} \begin{pmatrix} 3 \\ -1 \end{pmatrix} + V \cdot 1/\sqrt{10} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = S^t \begin{pmatrix} U \\ V \end{pmatrix}$$

und transponiert $(X,Y) = (U,V)S$. Die Gleichung der Quadrik $(X,Y)A(X,Y)^t = 5$ wird zu:

$$5 = (X,Y)A(X,Y)^t = (U,V)SAS^t \begin{pmatrix} U \\ V \end{pmatrix} = (U,V) \begin{pmatrix} 2 & \\ & 12 \end{pmatrix} \begin{pmatrix} U \\ V \end{pmatrix} = 2U^2 + 12V^2.$$

Ganz allgemein im Fall positiv definiter Matrix A gilt: Sind λ, μ die Eigenwerte der Matrix A, dann sind $\lambda, \mu > 0$, und die „rotierte“ Gleichung lautet

$$\lambda U^2 + \mu V^2 = r.$$

Damit haben die Halbachsen der Ellipse die Länge $\sqrt{r/\lambda}$ und $\sqrt{r/\mu}$. Für $\lambda = \mu$ entsteht ein Kreis.

Beispiel 12.20. Wenn A indefinit ist (Signatur $(1,1)$) und $r \neq 0$, dann entsteht eine Hyperbel.

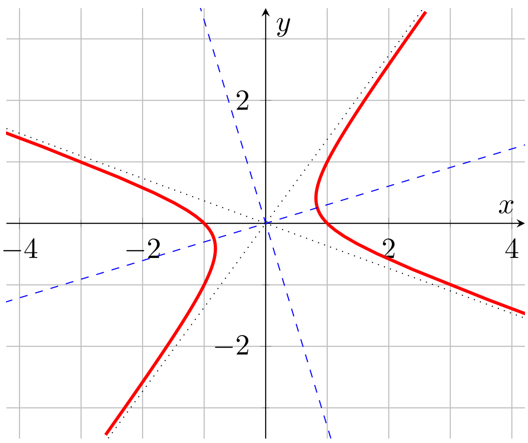


ABBILDUNG 15. Die Quadrik $X^2 + 2XY - 2Y^2 = 1$.

Hier ist $r = 1$ und

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

nach dem Hauptminorenkriterium negativ definit. Das charakteristische Polynom ist

$$(X-1)(X+2) - 1 = X^2 + X - 3,$$

so daß die Eigenwerte

$$\lambda_\pm = \frac{-1 \pm \sqrt{13}}{2}$$

und Eigenvektoren

$$v_+ = \begin{pmatrix} 3+\sqrt{13} \\ 2 \end{pmatrix} \approx \begin{pmatrix} 6,606 \\ 2 \end{pmatrix} \quad \text{und} \quad v_- = \begin{pmatrix} -2 \\ 3+\sqrt{13} \end{pmatrix} \approx \begin{pmatrix} 2 \\ 6.606 \end{pmatrix}$$

In diese orthogonalen Richtungen laufen die Symmetrieachsen (Hauptachsen) der Hyperbel.

Sind allgmein im negativ definiten Fall $\lambda, -\mu$ die Eigenwerte der Matrix A, dann ist oBdA $\lambda > 0 > -\mu$. Die rotierte Gleichung lautet

$$\lambda U^2 - \mu V^2 = r,$$

was sich als

$$(\sqrt{\lambda}U + \sqrt{\mu}V)(\sqrt{\lambda}U - \sqrt{\mu}V) = r$$

schreiben läßt und so als Hyperbel zu erkennen ist.

Beispiel 12.21. Wenn $A \neq 0$ den Eigenwert 0 hat, also nicht invertierbar ist (Signatur $(1,0,1)$ oder $(0,1,1)$) und $r \neq 0$, dann entstehen zwei parallele Geraden.

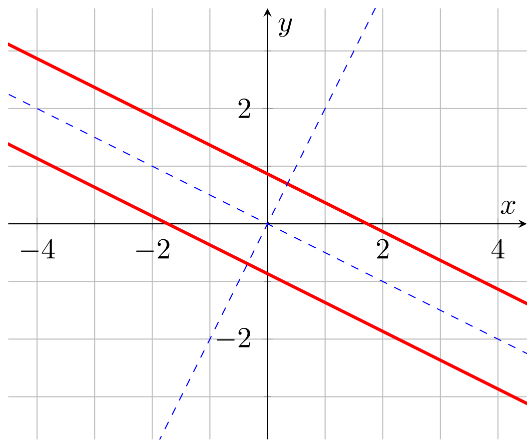


ABBILDUNG 16. Die Quadrik $X^2 + 4\,XY + 4\,Y^2 = 3$.

Im konkreten Beispiel sieht man

$$3 = X^2 + 4XY + 4Y^2 = (X + 2Y)^2$$

oder eben $X + 2Y = \pm\sqrt{3}$.

In einem geeigneten rotierten Koordinatensystem mit Koordinatenvariablen U, V lautet die Gleichung der Quadrik

$$\lambda U^2 = r,$$

was die beiden durch $U = \pm\sqrt{r/\lambda}$ definierten parallelen Geraden beschreibt.

Beispiel 12.22. Wir studieren die quadratische Form

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

und die Quadrik $q(x,y) = 25$:

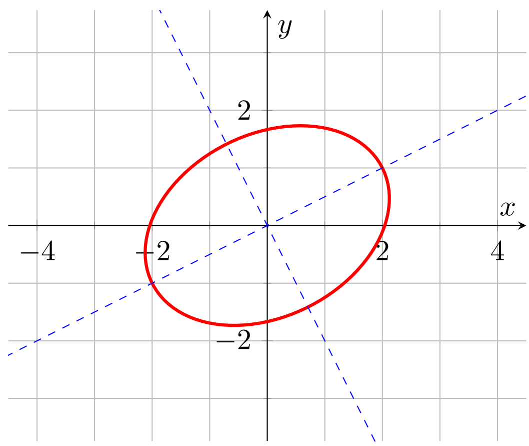


ABBILDUNG 17. Die Quadrik $6x^2 - 4xy + 9y^2 = 25$.

Die Eigenwerte und Eigenvektoren der Matrix

$$A = \begin{pmatrix} 6 & -2 \\ -2 & 9 \end{pmatrix}$$

sind abzulesen aus

$$A\begin{pmatrix}2\\1\end{pmatrix} = 5\begin{pmatrix}2\\1\end{pmatrix} \quad \text{und} \quad A\begin{pmatrix}-1\\2\end{pmatrix} = 10\begin{pmatrix}-1\\10\end{pmatrix}.$$

In Koordinaten u, v bezüglich der Eigenvektoren

$$\begin{pmatrix}x\\y\end{pmatrix} = u\begin{pmatrix}2\\1\end{pmatrix} + v\begin{pmatrix}-1\\2\end{pmatrix}$$

nimmt die quadratische Form Diagonalgestalt an:

$$q(u,v) = 25(u^2 + 2v^2).$$

Die Basiswechselmatrix ist

$$S = \begin{pmatrix}2 & -1\\1 & 2\end{pmatrix} \text{ und } S^tAS = \begin{pmatrix}25 & \\ & 50\end{pmatrix}.$$

Beispiel 12.23. Wir betrachten zu einer symmetrischen Matrix $A \in \mathrm{M}_3(\mathbb{R})$ und $r \in \mathbb{R}$ die Quadriken $\{v \in \mathbb{R}^3 \ ; \ v^tAv = r\}$ im $\mathbb{R}^3$. Wir nehmen an, daß A invertierbar ist. Die degenerierten Fälle mit einem nichttrivialen Radikal von $\langle\ ,\ \rangle_A$ führen zu Quadriken der Form $Q = Q_0 \times \mathbb{R}$ für eine ebene Quadrik $Q_0 \subseteq \mathbb{R}^2$.

Wenn man A durch $-A$ und r durch $-r$ ersetzt, ändert sich die Quadrik nicht. Daher können wir uns auf den Fall der Signatur $(3,0)$ und $(2,1)$ beschränken. Der Bequemlichkeit beschränken wir uns in Abbildung 18 im positiv definiten Fall auf Quadriken mit der Gleichung

$$x_1^2 + \frac{1}{4}x_2^2 + \frac{1}{9}x_3^2 = r,$$

und im indefiniten Fall weiter auf Quadriken mit der Gleichung

$$x_1^2 + x_2^2 \pm x_3^2 = r.$$

Beispiel 12.24. Im indefiniten Fall illustrieren wir den Übergang vom einschaligen zum zweischaligen Hyperboloid durch eine Art Daumenkino in Abbildung 19, in dem der Parameter r von $r = 4$ zu $r = -1$ variiert.

12.6. **Affine quadratische Formen.**

Definition 12.25. Eine **affine quadratische Form** auf $\mathbb{A}^n(K) = K^n$ ausgestattet mit dem Standardskalarprodukt $\langle\ ,\ \rangle$ ist eine Abbildung $q : K^n \to K$ der Form

$$q(x) = \langle x, Ax\rangle + \langle b, x\rangle + c$$

mit symmetischer Matrix $A \in \mathrm{M}_n(K)$, einem Vektor $b \in K^n$ und $c \in K$. Die zugehörige **Quadrik** ist die Lösungsmenge

$$V(q) = \{x \in K^n \ ; \ q(x) = 0\} \subseteq K^n.$$

Bemerkung 12.26. Der Kegel im $\mathbb{A}^3(\mathbb{R}) = \mathbb{R}^3$ mit Rotationssymmetrie um die z-Achse und $45°$-Winkel wird durch die Gleichung

$$x^2 + y^2 = z^2$$

beschrieben. Wir schneiden nun diesen Kegel mit der affinen Ebene, gegeben durch die Gleichung

$$ax + by + cz = d.$$

Wir nehmen der Einfachheit halber an, daß $c \neq 0$ und skalieren mit c^{-1}, so daß $c = 1$ wird. Dann gilt

$$z = d - ax - by.$$

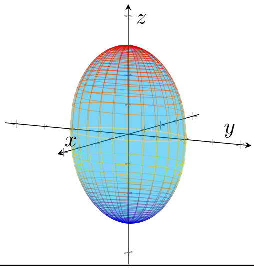
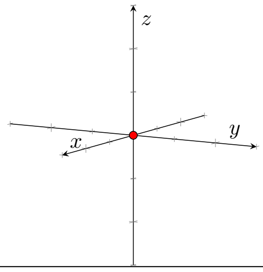
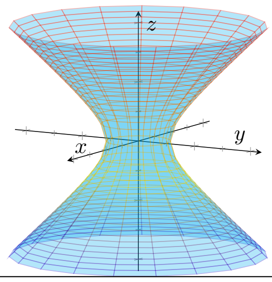
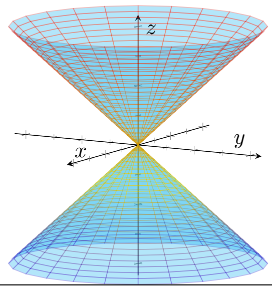
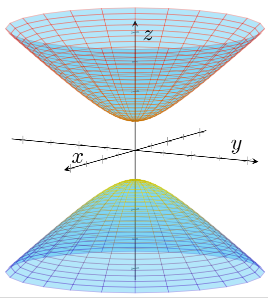
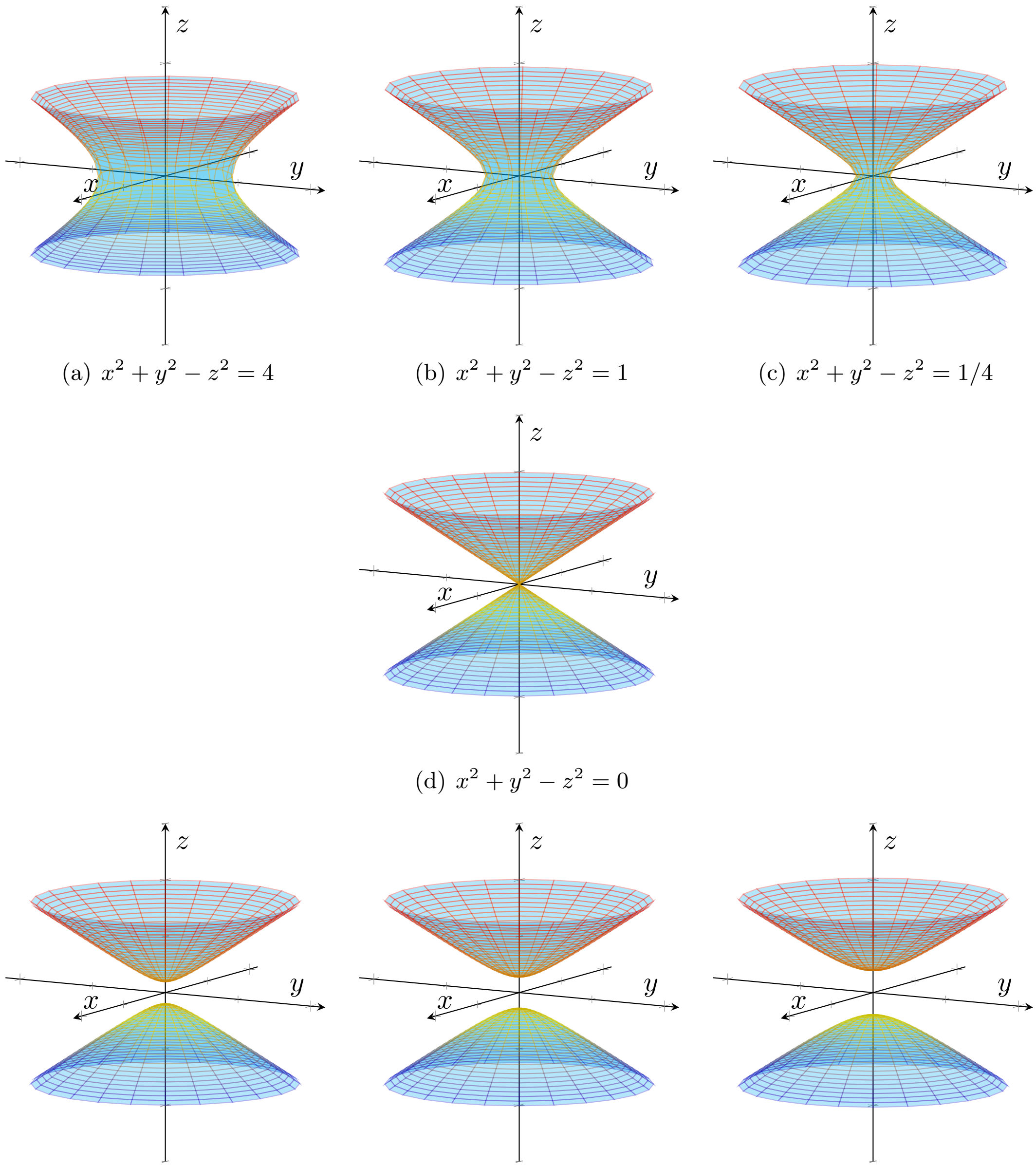
Signatur von A	r	Geometrie	Bezeichnung
$(3,0)$	>0		Ellipsoid
	$=0$		Punkt
	<0		leer
$(2,1)$	>0		einschaliges Hyperboloid
	$=0$		Doppelkegel
	<0		zweischaliges Hyperboloid

ABBILDUNG 18. Quadriken im $\mathbb{R}^3$.



(a) $x^2 + y^2 - z^2 = 4$ (b) $x^2 + y^2 - z^2 = 1$ (c) $x^2 + y^2 - z^2 = 1/4$

(d) $x^2 + y^2 - z^2 = 0$

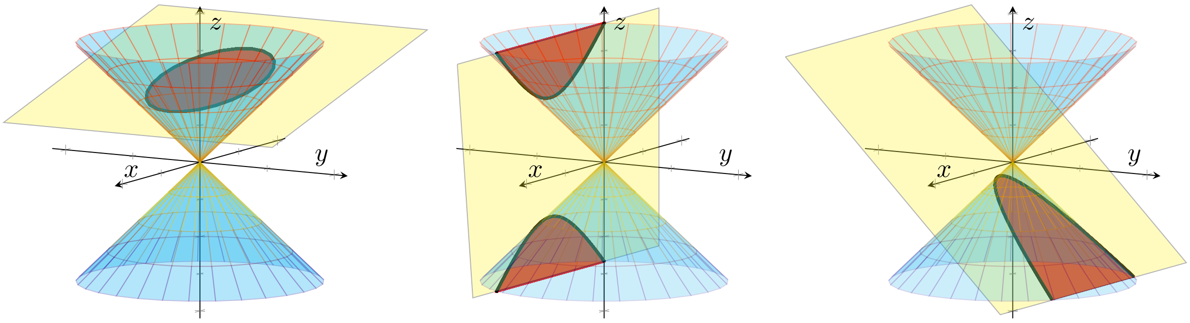
(e) $x^2 + y^2 - z^2 = -1/4$ (f) $x^2 + y^2 - z^2 = -1/2$ (g) $x^2 + y^2 - z^2 = -1$

Abbildung 19. Indefinite Quadriken im $\mathbb{R}^3$.

Der Schnitt von Kegel und Ebene ergibt sich als Lösungsmenge des Systems aus beiden Gleichungen von Kegel und Ebene. Substituieren wir z entsprechend der umgestellten Ebenengleichung, so erhalten wir

$$x^2 + y^2 = (d - ax - by)^2,$$

und das ist die Gleichung einer affinen Quadrik. Die Koordinaten x, y für die Schnittebene sind im Allgemeinen keine rechtwinkligen Koordinaten. Man muß zuerst einen linearen Basiswechsel durchführen zu einer ONB, so daß ein Teil davon den Translationsraum der affinen Ebene aufspannt. Aber für die Begründung des Namens „Kegelschnitte“ reicht die qualitative Beschreibung oben aus.



(a) Ellipse: $\mathscr{K} \cap \{x + 2z = 10\}$. (b) Hyperbel: $\mathscr{K} \cap \{y = -2\}$. (c) Parabel: $\mathscr{K} \cap \{2y + z = -2\}$.

Abbildung 20. Schnitt des Kegels $\mathscr{K} = \{4x^2 + 4y^2 - z^2 = 0\}$ mit einer Ebene.

Proposition 12.27 (Quadratisches Ergänzen)**.** *Sei $2 \in K^\times$, und sei q die affine quadratische Form $q(x) = \langle x, Ax\rangle + \langle b, x\rangle + c$ auf dem K^n mit $A \in \mathrm{GL}_n(K)$. Dann hat q die Form*

$$q(x) = \langle x - a, A(x - a)\rangle - r$$

für ein $a \in K^n$ und $r \in K$.

Beweis. Wir rechnen

$$\langle x - a, A(x - a)\rangle - r = \langle x, Ax\rangle - 2\langle Aa, x\rangle + \langle a, Aa\rangle - r,$$

so daß Koeffizientenvergleich die folgenden Gleichungen für a und r liefert:

$$-2Aa = b \quad \text{und} \quad c = \langle a, Aa\rangle - r.$$

Dies lösen wir zu

$$a = -\frac{1}{2}A^{-1}b \quad \text{und} \quad r = \frac{1}{4}\langle A^{-1}b, b\rangle - c. \qquad \square$$

Dies bedeutet, daß bis auf eine Translation eine affine Quadrik die Lösungsmenge der Form

$$\{x \in K^n \ ; \ \langle x, x\rangle_A = r\}$$

ist, also eine Quadrik.

Übungsaufgaben zu §12

Übungsaufgabe 12.1. Bringen Sie die quadratische Form in den Variablen X, Y, Z

$$q(X, Y, Z) = 2XY + 2YZ + 2ZX$$

durch eine orthogonale lineare Transformation der Variablen in Diagonalgestalt.

Übungsaufgabe 12.2. Entscheiden Sie, ob die folgenden quadratischen Formen $q(X, Y)$ zu Ellipsen oder zu Hyperbeln als Quadriken

$$\left\{ \begin{pmatrix} x \\ y \end{pmatrix} \ ; \ q(x, y) = 1 \right\}$$

führen:

(a) $X^2 + 4\,XY + 9\,Y^2$,
(b) $3\,X^2 + 8\,XY + 2\,Y^2$,
(c) $4\,X^2 + 12\,XY + 9\,Y^2$.

Jakob Stix, Institut für Mathematik, Goethe–Universität Frankfurt, Robert-Mayer-Str. 6–8, 60325 Frankfurt am Main, Germany

E-mail address: stix@math.uni-frankfurt.de