

Wie werden die Vorlesungen/Übungen organisiert?

- ▶ Einzige Änderung: die Hausaufgaben können zuzweit abgegeben werden.
- ▶ Sonsts wie im Wintersemester:
 - ▶ Sie müssen an den Übungen regelmäßig und aktiv teilnehmen
 - ▶ Mind. 60% der Punkte von der Hausaufgaben sammeln
 1. Am fast jeden Montag wird ein Übungsblatt mit in der regel vier Aufgaben in Netz gestellt (CAJ, bitte unter <http://caj.informatik.uni-jena.de/> sich anmelden).
 2. Heute kommen die erste Hausaufgaben.
 3. Sie müssen die Aufgaben lösen und vor der darauffolgenden Montagsvorlesung abgeben
 4. Dieses Mal ist die Regel „vor der Vorlesung“ streng.
 - ▶ in der Woche 16.06–22.06 Schreiben wir eine Probe-Klausur (Freiwillig, bis zum 20 % der Hausaufgabenpunkten)
 - ▶ Sie bekommen 2 Bonusblättern.
- ▶ Sie müssen die Modul-Anmeldungen ausdrücken, ausfüllen, und im Prüfungsamt (Physik-Studenten: mir) abgeben

Es gibt 2 oder 3 Übungsgruppen:

1. Mi 8–10. (Dr. Uta Freiberg)
 2. Do 8–10. (Dr. Konrad Schöbel)
- ▶ Nach Bedarf gibt es noch eine Übungsgruppe.

Verteilung in die Übungsgruppen findet morgen statt (weil die Anzahl von Übungsgruppen von Anzahl von Teilnehmern abhängt).

Was werden wir zuerst lernen?

- ▶ Grundlagen der Algebra und mathematischen Logik (Körpererweiterungen, Kardinalitäten, Paradoxen).
- ▶ Jordan-Normalformen (Verallgemeinerung von Diagonalisierung).

Körper – Wiederholung

Körper (Def. 13 und 16, LAAG I) ist die Menge mit zwei innerverknüpfungen $(\mathbb{K}, +, \cdot)$ mit der folgenden Eigenschaften:

- (R1) $(\mathbb{K}, +)$ ist eine abel'sche Gruppe, deren neutrales Element wir 0 bezeichnen;
- (R2) die Multiplikation „ \cdot “ ist assoziativ und kommutativ.
- (R3) es gilt das **Distributivgesetz**, d. h. für alle $a, b, c \in \mathbb{K}$ ist $a \cdot (b + c) = a \cdot b + a \cdot c$.
- (K4) $(\mathbb{K} \setminus \{0\}, \cdot)$ eine (abelsche) Gruppe ist.

Unterkörper ist eine nichtleere Teilmenge \mathbb{K}' der Körper, die eine Untergruppe bzgl. „ $+$ “ ist, und sodass $\mathbb{K}' \setminus \{0\}$ eine Untergruppe bzgl. „ \cdot “ ist.

Bemerkung $0 \in \mathbb{K}'$, weil \mathbb{K}' eine Untergruppe bzgl. $+$ ist, und deswegen das neutrale Element enthalten muß.

Bemerkung. Aus Def. 4/ Satz 5 LAAG I folgt, dass die Verknüpfungen „ $+$ “ und „ \cdot “ auf \mathbb{K}' wohldefiniert sind, und dass die Teilmengen \mathbb{K}' und $\mathbb{K}' \setminus \{0\}$ Gruppen bzgl. „ $+$ “ bzw. „ \cdot “ sind. **Deswegen ist ein Unterkörper auch ein Körper.**

Körpererweiterung

Zwei Körper $(\mathbb{K}_1, +_1, \cdot_1)$ und $(\mathbb{K}_2, +_2, \cdot_2)$ sind **isomorph**, falls es eine Bijektion $\phi : \mathbb{K}_1 \rightarrow \mathbb{K}_2$ gibt, die die Verknüpfungen erhält:
 $\phi(x \cdot_1 y) = \phi(x) \cdot_2 \phi(y)$, $\phi(x +_1 y) = \phi(x) +_2 \phi(y)$.

Def. 1 Körper \mathbb{H} heißt eine **Körpererweiterung** von einem Körper \mathbb{K} , falls es einen Unterkörper $\mathbb{H}' \subseteq \mathbb{H}$ gibt, der zu \mathbb{K} isomorph ist.
(In der Regel wird \mathbb{K} bereits ein Unterkörper von \mathbb{H} sein; in dem Fall wird Isomorphismus die Identitätsabbildung.)

Bsp. \mathbb{C} ist eine Körpererweiterung von \mathbb{R} .

Bsp. \mathbb{R} ist eine Körpererweiterung von \mathbb{Q} .

Wicht. Bsp. Quadratische Körpererweiterung von \mathbb{Q} .

Sei $s \in \mathbb{Q}_{\geq 0}$ sodass $\sqrt{s} \notin \mathbb{Q}$ (Z.B. $s = 2$).

Setze $\mathbb{Q}(\sqrt{s}) := \{x + y\sqrt{s} \mid x, y \in \mathbb{Q}\} \subseteq \mathbb{R}$.

Lemma 1 $\mathbb{Q}(\sqrt{s})$ ist eine Körpererweiterung von \mathbb{Q} .

Beweis. Offensichtlich ist $\mathbb{Q} \subseteq \mathbb{Q}(\sqrt{s})$. Nach Def. 1 müssen wir zeigen, dass $\mathbb{Q}(\sqrt{s})$ ein Körper ist.

Es genügt z.z. (s. Bemerkung oben), dass $\mathbb{Q}(\sqrt{s})$ ein Unterkörper von \mathbb{R} . Wir müssen zeigen, dass $\mathbb{Q}(\sqrt{s})$ und $\mathbb{Q}(\sqrt{s}) \setminus \{0\}$ Untergruppen von \mathbb{Q} ist, d.h., $\mathbb{Q}(\sqrt{s})$ abgeschlossen bzgl. „+“, „·“, und invertieren ist.

Addition: $x_1 + y_1\sqrt{s} + x_2 + y_2\sqrt{s} = \underbrace{x_1 + x_2}_{x \in \mathbb{Q}} + \underbrace{(y_1 + y_2)}_{y \in \mathbb{Q}} \sqrt{s}$.

Invertieren bzgl. „+“: $x + y\sqrt{s} + (-x - y\sqrt{s}) = 0$.

Multiplikation: $(x_1 + y_1\sqrt{s}) \cdot (x_2 + y_2\sqrt{s}) = \underbrace{x_1x_2 + y_1y_2s}_{x \in \mathbb{Q}} + \underbrace{(x_1y_2 + x_2y_1)}_{y \in \mathbb{Q}} \sqrt{s}$

Invertieren bzgl. „·“:

$$(x + y\sqrt{s}) \cdot \left(\underbrace{\frac{x}{x^2 - y^2s}}_{\in \mathbb{Q}} - \underbrace{\frac{y}{x^2 - y^2s}}_{\in \mathbb{Q}} \sqrt{s} \right) = \frac{(x+y\sqrt{s})(x-y\sqrt{s})}{x^2 - y^2s} = 1. \quad \square$$

Bemerkung: Für $x + y\sqrt{s} \neq 0$ ist der Nenner $x^2 - y^2s \neq 0$, weil sonst $\frac{x^2}{y^2} = s$ ist, also $\sqrt{s} = \frac{|x|}{|y|} \in \mathbb{Q}$, was Voraussetzungen widerspricht.

In Beweis von Lemma 1 kann man \mathbb{Q} mit jedem Unterkörper $\mathbb{K} \subseteq \mathbb{R}$ (oder sogar $\mathbb{K} \subseteq \mathbb{C}$) ersetzen (die Zahl s erfüllt dann die Bedingung $\sqrt{s} \notin \mathbb{K}$). Beweis wird buchstäblich wiederholt. Also gilt:

Lemma 1' Sei \mathbb{K} eine Unterkörper von \mathbb{R} (bzw. \mathbb{C}), und $s \in \mathbb{K}$ mit $\sqrt{s} \notin \mathbb{K}$. Dann ist die Menge $\mathbb{K}(\sqrt{s}) := \{x + y\sqrt{s} \mid x, y \in \mathbb{K}\} \subseteq \mathbb{R}$ (bzw. $\subseteq \mathbb{C}$) auch eine Unterkörper von \mathbb{R} (bzw. \mathbb{C}).

Dieser Unterkörper heißt eine **quadratische Erweiterung** von \mathbb{K} .

Bsp. $\mathbb{C} = \mathbb{R}(\sqrt{-1})$. Tatsächlich, jede Zahl $z \in \mathbb{C}$ kann man in der Form $\underbrace{x}_{\in \mathbb{R}} + \underbrace{y}_{\in \mathbb{R}} \cdot i$ darstellen.

Endliche Körpererweiterungen

Lemma 2 Sei \mathbb{K}' eine Unterkörper der Körper $(\mathbb{K}, +, \cdot)$. Dann ist $(\mathbb{K}, +, \cdot)$ ein \mathbb{K}' -Vektorraum.

(Diese Aussage war in Blatt 8 LAAG I).

Beweis. $(\mathbb{K}', +, \cdot)$ ist ein Körper; $(\mathbb{K}, +)$ ist eine abelsche Gruppe. Wir müssen die Axiomen (V1 – V4) nachweisen: für $\lambda, \mu \in \mathbb{K}'$ und $v, u \in \mathbb{K}$ soll

(V1) $\lambda(\mu v) = (\lambda\mu)v$ ($=: \lambda\mu v$) \iff Assoziativität von \mathbb{K} bzgl. „ \cdot “.

(V2) $(\lambda + \mu)v = (\lambda v) + (\mu v)$ \iff (R3) (das Distributivgesetz)

(V3) $\lambda(v + w) = (\lambda v) + (\lambda w)$ \iff (R3) (das Distributivgesetz)

(V4) $1v = v$ weil Einselement der Untergruppe auch Einselement der Gruppe ist. □

Eine Körpererweiterung $\mathbb{K}' \subseteq K$ heißt **endlich**, falls die Dimension von $(\mathbb{K}, +, \cdot)$ ein **endlichdimensionaler** \mathbb{K}' -Vektorraum ist.

Bsp. $\mathbb{Q}(\sqrt{s})$ ist eine endliche Erweiterung von \mathbb{Q} (weil jedes Element die Form $x + y \cdot \sqrt{s}$ hat, also die Menge $\{1, \sqrt{s}\}$ ist erzeugend.)

Bsp. \mathbb{C} ist eine endliche Erweiterung von \mathbb{R} , weil die Menge $\{1, i\}$ erzeugend ist. z 1

Satz 1. \mathbb{R} ist keine endliche Erweiterung von \mathbb{Q} .

Exkurs: Mächtigkeit (Ordnung) einer Menge

Def. 2 Seien A, B Mengen. Wir sagen, dass A höchstens gleichmächtig als B ist, falls es eine injektive Abbildung $f : A \rightarrow B$ gibt. Schreibweise: $|A| \leq |B|$. Wir sagen, dass A und B gleichmächtig sind, falls $|A| \geq |B|$ und $|B| \geq |A|$ ist. Schreibweise: $|A| = |B|$.

Bsp. Die Mengen A, B seien endlich, $|A| := a$, $|B| := b$. Dann gilt:
 $|A| \geq |B| \iff a \geq b$, folglich $|A| = |B| \iff a = b$.

Bemerkung. Gibt es eine Bijektion $f : A \rightarrow B$, so gilt: $|A| = |B|$.

Beweis. f ist bijektiv, deswegen injektiv, deswegen $|A| \leq |B|$. Da f bijektiv ist, gibt es eine inverse Abbildung $g : B \rightarrow A$, die auch bijektiv (folglich injektiv) ist; also $|A| \geq |B|$. □

Bsp. $I_1 := \{x \in \mathbb{R} \mid 0 < x < 1\}$ ist gleichmächtig mit $I_2 := \{x \in \mathbb{R} \mid 0 < x < 2\}$.



Tatsächlich, die Abbildung $f : x \mapsto 2 \cdot x$ ist eine Bijektion zwischen I_1 und I_2 .

Bsp. $|\mathbb{Z}| = |\mathbb{N}|$. **Beweis.** Die Abbildung $f : \mathbb{N} \rightarrow \mathbb{Z}$,
$$f(n) = \begin{cases} (n-1)/2 & \text{falls } n \text{ ungerade ist} \\ -n/2 & \text{falls } n \text{ gerade ist} \end{cases}$$
 ist eine Bijektion

Ein kuriozes Beispiel

Hilberts Hotel (Hilbert:  1862 –1943) besteht aus

unendlich viel Zimmer (mit natürlichen Zahlen nummeriert), die leider alle belegt sind. Es kommen aber noch unendlich viele Gäste (auch mit natürlichen Zahlen durchnummeriert). Kann man Sie unterbringen, ohne die Zimmern doppelt belegt sind?

Antwort: Ja!

Den Gast aus Zimmer k geht in das Zimmer $2k$

Der Neugekommene mit Nummer k geht in das Zimmer $2k - 1$.

Keine Doppelbelegung: die alte Guesste sind jetzt in Zimmern mit geraden Nummern, die neue in in Zimmern mit ungeraden Nummern. Alle sind untergebracht.

Wicht. Bsp. $|\mathbb{N}| = |\mathbb{N} \times \mathbb{N}|$.

Beweis: Die injektive Abbildung $\mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N}$ ist einfach zu finden:
 $n \mapsto (1, n)$.

Wir konstruieren eine injektive Abbildung $f : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$.

Sei p_n die n -te Primzahl. Z.B., $p_1 = 2$, $p_2 = 3$, $p_3 = 5$, $p_4 = 7$.

Wir definieren $f : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$, $f(n, m) = p_n^m$.

Die Abbildung ist injektiv: ist $f(n, m) = f(n', m')$, so ist $p_{n'}^{m'} = p_n^m$

folglich $p_{n'}^{m'} \equiv p_n^m \pmod{p_n}$

folglich $p_{n'}^{m'} \equiv 0 \pmod{p_n}$,

was unmöglich ist, weil \mathbb{Z}_{p_n} ein Körper ist und $p_{n'} \not\equiv 0 \pmod{p_n}$ ist

.



Daraus folgt insbesondere, dass $|\mathbb{N}| = |\mathbb{Q}|$.

In der Tat, $|\mathbb{Q}| \geq |\mathbb{N}|$, weil die Abbildung $I : \mathbb{N} \rightarrow \mathbb{Q}$, $I(n) = n$ eine Injektion ist.

Um zu zeigen, dass $|\mathbb{Q}| \leq |\mathbb{N}|$, benutzen wir die Abbildung

$f \circ g : \mathbb{Q} \rightarrow \mathbb{N}$, wobei

f die injektive Abbildung $\mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N}$ aus dem Wicht. Bsp. ist,

und $g : \mathbb{Q} \rightarrow \mathbb{N} \times \mathbb{N}$ ist die folgende injektive Abbildung:

$$g(p/q) = \begin{cases} (|p|, |q|) & \text{falls } p/q \geq 0 \\ (2 \cdot |p|, 2 \cdot |q|) & \text{falls } p/q < 0 \end{cases} \in \mathbb{N} \times \mathbb{N}.$$

(Wir nehmen an, dass p und q Teilfremd sind.)

Da die Verkettung von injektiven Abbildungen injektiv ist, ist

$f \circ g : \mathbb{Q} \rightarrow \mathbb{N}$ auch injektiv, folglich $|\mathbb{Q}| \leq |\mathbb{N}|$. Dann $|\mathbb{Q}| = |\mathbb{N}|$. \square

Satz 2. Zwei Mengen A, B sind genau dann gleichmächtig, wenn es eine Bijektion $h: A \rightarrow B$ gibt.

Beweis. \Leftarrow ist trivial (und ist bereits oben bewiesen; wir wiederholen den Beweis): h ist injektiv; deswegen $|A| \leq |B|$. Die inverse Abbildung h^{-1} ist auch eine Bijektion, folglich injektiv, folglich $|A| \geq |B|$.

Es gelte jetzt $|A| = |B|$, und zwar seien $f: A \rightarrow B$ und $g: B \rightarrow A$ injektiv. Wir definieren jetzt eine Abbildung $h: A \rightarrow B$ wie folgt: Sei $A_1 = A \setminus \text{Bild}_g(B)$ und dann rekursiv $A_{n+1} = \text{Bild}_{g \circ f}(A_n)$. Sei $C = \bigcup_{n \geq 1} A_n$.

* Ist $a \in C$, so setze $h(a) = f(a)$.

* Ist $a \notin C$, so folgt insbesondere $a \in \text{Bild}_g(B)$, wir können somit $h(a) = g^{-1}(a)$ setzen.

Diese Abbildung h ist injektiv: Es gelte $h(a_1) = h(a_2)$ für zwei Elemente $a_1, a_2 \in A$.

* Ist $a_1 \in C$ und $a_2 \notin C$, so ergibt sich mit $a_2 = g(h(a_2)) = g(h(a_1)) = g(f(a_1)) \in A_{n+1}$ ein Widerspruch.

* Der Fall $a_2 \in C$ und $a_1 \notin C$ ist ebenso ausgeschlossen.

* Ist $a_1 \in C$ und $a_2 \in C$, so folgt $f(a_1) = h(a_1) = h(a_2) = f(a_2)$, also $a_1 = a_2$.

* Ist weder $a_1 \notin C$ und $a_2 \notin C$, so folgt $a_1 = g(h(a_1)) = g(h(a_2)) = a_2$.

Die Abbildung ist aber auch surjektiv: Sei $b \in B$ beliebig.

* Ist $g(b) \in A_n$ für ein $n \geq 1$, so folgt nach Definition von A_1 , dass sogar $n > 1$ gilt. Folglich ist $g(b) = g(f(a))$ für ein $a \in A_{n-1}$ und $h(a) = f(a) = b$.

* Ansonsten gilt $h(g(b)) = g^{-1}(g(b)) = b$.

Folglich ist $h: A \rightarrow B$ eine Bijektion,

Folgerung. $|\underbrace{\mathbb{N} \times \cdots \times \mathbb{N}}_{k \text{ Stuck}}| = \mathbb{N}.$

Beweis für $k = 3$. Wir haben in Wicht. Bsp. gezeigt, dass $|\underbrace{\mathbb{N} \times \mathbb{N}}| = \mathbb{N}.$

Nach Satz 2 gibt es dann eine Bijektion $\phi : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}.$ Wir konstruieren eine Bijektion $\Phi : \mathbb{N} \times \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}.$ Wir setzen

$$\Phi(n_1, n_2, n_3) := \phi(\phi(n_1, n_2), n_3) \in \mathbb{N}.$$

Φ ist injektiv: gilt $\Phi(n_1, n_2, n_3) = \Phi(m_1, m_2, m_3),$ so ist

$\phi(\phi(n_1, n_2), n_3) = \phi(\phi(m_1, m_2), m_3).$ Da ϕ bijektiv ist, folgt daraus, dass $\phi(n_1, n_2) = \phi(m_1, m_2)$ und $n_3 = m_3.$ Da ϕ bijektiv ist, folgt aus $\phi(n_1, n_2) = \phi(m_1, m_2)$ dass $n_1 = m_1$ und $n_2 = m_2.$ Also, Φ ist injektiv.

Φ ist surjektiv: Es sei $n \in \mathbb{N}.$ Da ϕ surjektiv ist, gibt es n_1, n_2 mit

$\phi(n_1, n_2) = n.$ Da ϕ surjektiv ist, gibt es m_1, m_2 mit $\phi(m_1, m_2) = n_1.$

Dann ist $\Phi(m_1, m_2, n_2) = \phi(\phi(m_1, m_2), n_2) = \phi(n_1, n_2) = n. \quad \square$

Bemerkung: Beweis für beliebiges k erfolgt nach Induktion (im Beweis für $k = 3$ haben wir praktisch Induktionsschritt gemacht).

Def. 3 Wir sagen, dass A **mächtige** als B ist, falls $|A| \geq |B|$, aber $|A| \neq |B|$. **Schreibweise:** $|A| > |B|$.

Sei A eine Menge. Die Menge aller Teilmengen von A wird 2^A bezeichnet und **Potenzmenge** heißen.

Bsp. Für $A = \emptyset$ ist $2^A = \{\emptyset\} \neq \emptyset$.

Bsp. Sei $A = \{1, 2\}$. Dann ist $2^A = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$.

Sei $A = \{0, 1, 2\}$. Dann ist

$$2^A = \{\emptyset, \{0\}, \{1\}, \{2\}, \{0, 1\}, \{1, 2\}, \{0, 2\}, \{0, 1, 2\}\}.$$

Bemerkung. Sei A endlich. Dann gilt: $|2^A| = 2^{|A|}$.

Beweis. Angenommen $A = \{0, \dots, n\}$ (d.h. $|A| = n + 1$).

Wir ordnen jeder Teilmenge A' die folgende binäre Zahl

$$\alpha_n \alpha_{n-1} \dots \alpha_0 \text{ zu: } \alpha_i = \begin{cases} 1 & \text{falls } i \in A' \\ 0 & \text{falls } i \notin A' \end{cases}.$$

Z.B. $\{1, 2\} \mapsto 110_2$.

$\{0, 2\} \mapsto 101_2$.

Diese Abbildung (von 2^A nach binäre Zahlen aus höchstens $n + 1$ Ziffern) ist injektiv und surjektiv. Dann ist

$$|2^A| = \#\{\text{Binäre Zahlen von } 0 \text{ bis } \underbrace{1\dots 1}_{n+1}\} = 2^{n+1}.$$

Satz 3 Für eine beliebige Menge A gilt: $|A| < |2^A|$.

Beweis. Es sind die folgenden beiden Aussagen zu zeigen:

(i) Es gibt eine Injektion $f : A \rightarrow 2^A$ (daraus folgt, dass $|A| \leq |2^A|$)

(ii) Es gibt keine Bijektion zwischen A und 2^A (gibt es eine Injektion von 2^A nach A , so gibt es nach Satz 2 eine Bijektion zwischen A und 2^A .)

(i) offensichtlich: die Abbildung $f : a \mapsto \{a\}$ leistet das Verlangte.

(ii) Angenommen, irgendeine Abbildung $f : A \rightarrow 2^A$ wäre bijektiv. Dies wird nun zum Widerspruch geführt.

Die Teilmenge $M \subseteq A$ wird definiert als $M := \{a \in A \mid a \notin f(a)\}$.

Da f bijektiv und deswegen surjektiv ist, und da $M \in 2^A$ ist, hat M ein Element $a \in A$ mit $f(a) = M$. Nun gilt:

$$a \in M \iff a \notin f(a) \iff a \notin M.$$

(Die erste **Äquivalenz** beinhaltet die Definition von M , die zweite **Äquivalenz** benutzt nur die Urbildeigenschaft.)

Damit ist der gewünschte Widerspruch vorhanden. □

Satz 4 $|2^{\mathbb{N}}| = |\mathbb{R}|$.

Beweis. Z.z.: (i) Es gibt eine injektive Abbildung $f : \mathbb{R} \rightarrow 2^{\mathbb{Q}}$.

Da $|\mathbb{N}| = |\mathbb{Q}|$, ist $|2^{\mathbb{N}}| = |2^{\mathbb{Q}}|$. Deswegen genügt es eine injektive Abbildung $f : \mathbb{R} \rightarrow 2^{\mathbb{Q}}$ zu konstruieren.

(ii) Es gibt eine injektive Abbildung $g : 2^{\mathbb{N}} \rightarrow \mathbb{R}$.

(i): Konstruktion von f : Man kann jede Zahl $a \in \mathbb{R}$ in der dezimalen Form schreiben: $a = \alpha_k \alpha_{k-1} \dots \alpha_0, \alpha_{-1} \alpha_{-2} \dots = \sum_{i=k}^{-\infty} \alpha_i \cdot 10^i$, wobei $\alpha_i \in \{0, \dots, 9\}$.

Z.B. ist

$$\pi = 3.14159\dots = 3 + 1/10 + 4/100 + 1/1000 + 5/10000 + 9/100000 + \dots$$

Falls nur endlich viele α_i von Null verschieden sind, ist die Zahl rational.

Wir ordnen der Zahl $a = \alpha_k \alpha_{k-1} \dots \alpha_0, \alpha_{-1} \alpha_{-2} \dots$, die folgende Teilmenge von \mathbb{Q} zu: $\{\alpha_k \cdot 10^k, \alpha_{k-1} \cdot 10^{k-1}, \dots, \alpha_0, \alpha_{-1} \cdot 10^{-1}, \dots\}$.

Die Abbildung ist offensichtlich injektiv. Also, $|2^{\mathbb{Q}}| \geq |\mathbb{R}|$.

(ii) Konstruktion von g : Man ordne der Teilmenge $A = \{\dots a_i \dots\} \subseteq \mathbb{N}$ die Zahl $\sum_i 10^{-a_i}$ zu.

Z.B., falls $A = \{1, 3\}$, dann ist $g(A) = 0,101$. Die Abbildung ist

offensichtlich injektiv. Dann $|2^{\mathbb{Q}}| \leq |\mathbb{R}|$, folglich $|2^{\mathbb{Q}}| = |\mathbb{R}|$. □

Beweis vom Satz 1

Satz 1. \mathbb{R} ist keine endliche Erweiterung von \mathbb{Q} .

Beweis. Laut Definition, ist jede endliche Körpererweiterung \mathbb{K} von \mathbb{Q} ein endlichdimensionaler Vektorraum über \mathbb{Q} . Dann ist er zu \mathbb{Q}^k isomorph. Dann gibt es eine Bijektion zwischen \mathbb{Q}^k und \mathbb{K} .

Wir haben jedoch gezeigt, dass es keine Bijektion zwischen \mathbb{Q}^k und \mathbb{R} gibt. In der Tat,

$$\left. \begin{array}{l} |\mathbb{Q}| \stackrel{\text{gestern bewiesen}}{=} |\mathbb{N}| \implies |2^{\mathbb{Q}}| \stackrel{\text{Satz 2}}{=} |2^{\mathbb{N}}| \stackrel{\text{Satz 4}}{=} |\mathbb{R}| \\ |\mathbb{Q}| \stackrel{\text{gestern bewiesen}}{=} |\mathbb{N}| \stackrel{\text{Folgerung}}{=} |\mathbb{N}^k| \stackrel{\text{Satz 2}}{=} |\mathbb{Q}^k| \\ |2^{\mathbb{Q}}| \stackrel{\text{Satz 3}}{>} |\mathbb{Q}| \end{array} \right\} \implies |\mathbb{R}| > |\mathbb{Q}^k|$$

Nach Definition von “>” gibt es keine Bijektion zwischen \mathbb{R} und \mathbb{Q}^k .

Komplexifizierung von \mathbb{R} -Vektorräumen

Komplexifizierung ist eine Operation, die einem \mathbb{R} -Vektorraum einen \mathbb{C} -Vektorraum zuordnet, der sehr ähnliche Eigenschaften hat.

Es sei $(V, +, \cdot)$ ein \mathbb{R} -Vektorraum.

Wir konstruieren ein \mathbb{C} -Vektorraum $V_{\mathbb{C}}$.

Die Menge: $V_{\mathbb{C}} := \{u + i \cdot v \mid u, v \in V\}$ der formalen Ausdrücken der Form $u + i \cdot v$ (also ist $V \times V$ als die Menge).

Verknüpfungen:

$$+ : V_{\mathbb{C}} \times V_{\mathbb{C}} \rightarrow V_{\mathbb{C}}, (u_1 + i \cdot v_1) + (u_2 + i \cdot v_2) = (u_1 + u_2) + (v_1 + v_2) \cdot i.$$

$$\bullet : \mathbb{C} \times V_{\mathbb{C}} \rightarrow V_{\mathbb{C}}, (\alpha + \beta \cdot i) \cdot (u + i \cdot v) = \underbrace{(\alpha u - \beta v)}_{\in V} + \underbrace{(\alpha v + \beta u)}_{\in V} \cdot i.$$

Bsp. $(\alpha + \beta \cdot i)(u + \vec{0} \cdot i) = \alpha u + \beta u \cdot i.$

Bsp. $\alpha(u + v \cdot i) = \alpha u + \alpha v \cdot i.$

Aussage. $(V_{\mathbb{C}}, +, \cdot)$ ist ein \mathbb{C} -Vektorraum.

Beweis. Direkt nach Definition: man muss die Axiome (V1 – V4) überprüfen. □

Basis und Dimension der komplexifizierten Vektorraum

Sei (b_1, \dots, b_n) eine Basis von V . Dann ist $(b_1 + \vec{0} \cdot i, \dots, b_n + \vec{0} \cdot i)$ eine Basis in $V_{\mathbb{C}}$. Daraus folgt, dass $\dim(V) = \dim(V_{\mathbb{C}})$.

Tatsächlich, die Vektoren $(b_1 + \vec{0} \cdot i, \dots, b_n + \vec{0} \cdot i)$ sind linear unabhängig:
ist

$$\vec{0} + \vec{0} \cdot i = (\alpha_1 + \beta_1 \cdot i)(b_1 + \vec{0} \cdot i) + \dots + (\alpha_n + \beta_n \cdot i)(b_n + \vec{0} \cdot i), \text{ so ist}$$
$$\vec{0} + \vec{0} \cdot i = (\alpha_1 b_1 + \beta_1 b_1 \cdot i) + \dots + (\alpha_n b_n + \beta_n b_n \cdot i).$$

Die letzte Gleichung ist äquivalent zu den folgenden zwei Gleichungen:

$$\vec{0} = \alpha_1 b_1 + \dots + \alpha_n b_n \quad (*)$$

$$\vec{0} = \beta_1 b_1 + \dots + \beta_n b_n \quad (**)$$

Wir sehen, dass die beide Gleichungen die Gleichungen in V sind; da (b_1, \dots, b_n) eine Basis in V ist, sind $\alpha_i = \beta_i = 0$ folglich $\alpha + \beta \cdot i = 0$ folglich die Menge $\{b_1 + \vec{0} \cdot i, \dots, b_n + \vec{0} \cdot i\}$ ist linear unabhängig.

Die Menge ist auch erzeugend: wir betrachten einen Vektor

$u + v \cdot i \in V_{\mathbb{C}}$. Da u bzw. v Vektoren aus V sind, kann man sie als Linearkombinationen von Basisvektoren darstellen:

$$u = \alpha_1 b_1 + \dots + \alpha_n b_n, \quad v = \beta_1 b_1 + \dots + \beta_n b_n. \text{ Dann ist}$$

$$u + v \cdot i = (\alpha_1 + \beta_1 \cdot i)(b_1 + \vec{0} \cdot i) + \dots + (\alpha_n + \beta_n \cdot i)(b_n + \vec{0} \cdot i). \text{ Also, die}$$

Menge ist linear unabhängig und erzeugend folglich ist $(b_1 + \vec{0} \cdot i, \dots, b_n + \vec{0} \cdot i)$ eine Basis.

Komplexifizierung von linearen Abbildungen

Sei $f : V \rightarrow W$ eine lineare Abbildung von \mathbb{R} -Vektorraum V nach \mathbb{R} -Vektorraum W .

Frage: Kann man die Abbildung f linear auf $V_{\mathbb{C}}$ verlängern? Gibt es eine Abbildung $f_{\mathbb{C}}$ sodass $f_{\mathbb{C}}(u + \vec{0} \cdot i) = f(u) + \vec{0} \cdot i$?

Antwort: Ja!

Die folgende Abbildung $f_{\mathbb{C}} : V_{\mathbb{C}} \rightarrow W_{\mathbb{C}}$: $f_{\mathbb{C}}(u + v \cdot i) = f(u) + f(v) \cdot i$ hat offensichtlich diese Eigenschaft (weil $f(\vec{0}) = \vec{0}$), und ist linear:

$$\begin{aligned} f_{\mathbb{C}}(u_1 + v_1 \cdot i + u_2 + v_2 \cdot i) &= f(u_1 + u_2) + f(v_1 + v_2) \cdot i = \\ f(u_1) + f(u_2) + (f(v_1) + f(v_2)) \cdot i &= f_{\mathbb{C}}(u_1 + v_1 \cdot i) + f_{\mathbb{C}}(u_2 + v_2 \cdot i). \end{aligned}$$

Ähnlich bzgl. Multiplizieren mit Skalaren.

Sei A die Matrix der Abbildung f bzgl. Basen (b_1, \dots, b_n) in V und (c_1, \dots, c_m) in W .

Frage: Welche Matrix hat die Abbildung F_C bzgl. der Basen $(b_1 + \vec{0} \cdot i, \dots, b_n + \vec{0} \cdot i)$ in V_C und $(c_1 + \vec{0} \cdot i, \dots, c_m + \vec{0} \cdot i)$ in W_C ?

Antwort: Dieselbe Matrix A .

Beweis. Die j -te Spalte der Matrix f_C ist der Koordinatenvektor von

$f_C(b_j)$ in der Basis c_i , also die Zahlen $\begin{pmatrix} \lambda_1 + \mu_1 \cdot i \\ \vdots \\ \lambda_m + \mu_m \cdot i \end{pmatrix}$ sodass

$$f_C(b_j + \vec{0} \cdot i) = (\lambda_1 + \mu_1 \cdot i)(c_1 + \vec{0} \cdot i) + \dots + (\lambda_m + \mu_m \cdot i)(c_m + \vec{0} \cdot i) \quad (*)$$

Die Gleichung $(*)$ ist äquivalent zu zwei Gleichungen:

$$\vec{0} = \mu_1 c_1 + \dots + \mu_m c_m \implies \text{alle } \mu_i = 0.$$

$$f(b_j) \stackrel{\text{Weil } \mu_i = 0}{=} \lambda_1 c_1 + \dots + \lambda_m c_m.$$

Aber diese Gleichung hat nur eine Lösung, und zwar die Einträge der j -te Spalte der Matrix von f sind die Lösung. Also, die Matrizen von f_C und von f sind gleich.

Konjugieren

$V_{\mathbb{C}}$ sein die komplexifizierung des \mathbb{R} -Vektorraums V . Dann heißt die Abbildung $u + v \cdot i \mapsto u - v \cdot i$ **konjugieren**, und wird wie übliche Konjugation mit „quer“ oben bezeichnet

$\overline{u + v \cdot i} := u - v \cdot i$. Ferner gilt:

$$\overline{u + v \cdot i} = u + v \cdot i \iff v = \vec{0}.$$

Rechenregeln für Konjugieren:

- ▶ $\overline{f_{\mathbb{C}}(u + v \cdot i)} = f_{\mathbb{C}}(u - v \cdot i)$.
- ▶ $\overline{(\alpha + \beta \cdot i)(u + v \cdot i)} = (\alpha - \beta \cdot i)(u - v \cdot i)$.

(Beweis: nachrechnen)

Anwendung von Komplexifizierung

Satz 5 Sei $f : V \rightarrow V$ ein Endomorphismus von n -dimensionalen \mathbb{R} -Vektorraum V . Angenommen, die Komplexifizierung $f_{\mathbb{C}}$ von f ist diagonalisierbar (als Endomorphismus von \mathbb{C} -Vektorraum $V_{\mathbb{C}}$.) Dann gibt es ein Basis B in V sodass die Matrix von f die folgende Form hat

$$\left(\begin{array}{c} \boxed{\begin{array}{ccc} \lambda_1 & & \\ & \ddots & \\ & & \lambda_k \end{array}} & & \\ & \boxed{\begin{array}{cc} \alpha_1 & \beta_1 \\ -\beta_1 & \alpha_1 \end{array}} & \\ & & \ddots \\ & & \boxed{\begin{array}{cc} \alpha_m & \beta_m \\ -\beta_m & \alpha_m \end{array}} \end{array} \right), \text{ wobei } \beta_j \neq 0 \text{ ist.}$$

(k oder m können auch gleich 0 sein. Selbstverständlich gilt $2 \cdot m + k = \dim(V)$.)

Bemerkung. Beweis enthält auch einen Algorithmus, wie man die Basis finden kann.

Bemerkung. In LAAG I hatten wir zwei Diagonalisierbarkeitkriterien, Sätze 55, 58.

Exkurs: Hauptsatz der Algebra

Warum ist \mathbb{C} oft besser als \mathbb{R} ? Weil in \mathbb{C} der Hauptsatz der Algebra gilt.

Hauptsatz der Algebra (Beweis in Analysis – Vorlesungen oder in Funktionentheorie) *Jedes $P \in \mathbb{C}[x]$ mit $\text{Grad}(P) \geq 1$ hat mind. eine Nullstelle*

Bsp. In $\mathbb{R}[x]$ ist die Aussage falsch: z.B. $x^2 + 1$ hat keine Nullstell in \mathbb{R} . (wenn wir das Polynom als Polynom über Komplexe koeffizienten auffassen, hat das Polynom die Nullstellen $x_1 = i$, $x_2 = -i$.)

Folgerung A Jedes $P \in \mathbb{C}[x]$, $P \neq 0$, kann man in lineare Faktoren zehrlegen (d.h. in der Form $P = a(x - x_1)(x - x_2)\dots(x - x_n)$ schreiben, wobei $a, x_i \in \mathbb{C}$ sind). Diese Zerlegung ist eindeutig bis zum umstellen von Faktoren.

Beweis Existenz: Sei P ein Polynom des Grades $n \geq 1$. Nach Hauptsatz der Algebra hat er eine Nullstelle x_1 . Nach Lemma 6 ist dann

$P = (x - x_1)g$, wobei $\text{Grad}(g) = \text{Grad}(P) - 1$. Ist $\text{Grad}(g) = 0$, so ist $g = a \in \mathbb{C}$, und wir sind fertig. Sonst hat g eine Nullstelle x_2 , und

deswegen (Lemma 6) ist $g = (x - x_2)h$, wobei

$\text{Grad}(h) = \text{Grad}(g) - 1 = \text{Grad}(P) - 2$, also $P = (x - x_1)(x - x_2)h$,

U.S.W. Nach n Schritte bekommen wir $P = a(x - x_1)\dots(x - x_n)$.

Eindeutigkeit Induktion nach n . **I.A.** ist trivial: hat P Grad 0, so ist

$P = a = b$, also $a = b$. **I.V.:** Angenommen, jedes Polynom des Grades

$n - 1$ kann man Eindeutig in der Form $P = a(x - x_1)(x - x_2)\dots(x - x_n)$

darstellen. **I.S.** Z.z.: die Eindeutigkeit gilt auch für Polynome des Grades

n . Sei $P = a(x - x_1)\dots(x - x_n) = b(x - y_1)\dots(x - y_n)$, wobei $a \neq 0 \neq b$.

Da x_1 eine Nullstelle des Polynoms $a(x - x_1)\dots(x - x_n) = b(x - y_1)\dots(x - y_n)$ ist, ist

$b(x_1 - y_1)(x_1 - y_2)\dots(x_1 - y_n) = 0$, und deswegen ist ein von y_i gleich x_1 . OBdA ist

$y_1 = x_1$. Dividieren des Polynoms durch $(x - x_1)$ gibt

$a(x - x_2)\dots(x - x_n) = b(x - y_2)\dots(x - y_n)$. Nach **I.V.** ist $a = b$ und die

Faktoren $(x - x_i)$ sind die Faktoren $(x - y_i)$ ($i \geq 2$), möglicherweise in

andere Reihenfolge.

Ist $P \in \mathbb{R}[x]$, so ist $P \in \mathbb{C}[x]$, weil $\mathbb{R} \subseteq \mathbb{C}$.

Folgerung B Jedes $P \in \mathbb{R}[x]$, $\text{Grad}(P) > 0$, kann man in Produkt von lineare und quadratischen Faktoren $g_i \in \mathbb{R}[x]$ zerlegen: $P := g_1 g_2 \dots g_m$, wobei $\text{Grad}(g_i) \in \{1, 2\}$.

Hilfsaussage. Es sei $P = \sum_{k=0}^n a_k x^k \in \mathbb{R}[x] \subseteq \mathbb{C}[x]$. Dann gilt für jedes $z \in \mathbb{C}$: $P(\bar{z}) = \overline{P(z)}$ (wobei \bar{z} komplexe Konjugation ist)

Wiederholung: Für $z = a + ib$ ist $\bar{z} = a - ib$.

Wir erinnern zunächst an folgende Eigenschaften komplexer Zahlen:

$$\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}, \quad \overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}.$$

(Konjugieren ist ein Autoisomorphismus des Körpers \mathbb{C})

Deswegen ist $\overline{z^k} = \overline{z}^k$ für alle $k \in \mathbb{N}$. Damit erhalten wir wegen $a_k \in \mathbb{R}$

$$\begin{aligned} \overline{P_n(z)} &= \overline{a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0} = \\ &= \overline{a_n z^n} + \overline{a_{n-1} z^{n-1}} + \dots + \overline{a_1 z} + \overline{a_0} = a_n \overline{z}^n + a_{n-1} \overline{z}^{n-1} + \dots + a_1 \overline{z} + a_0 = P_n(\overline{z}). \end{aligned}$$

Hilfsaussage ist bewiesen. Daraus folgt insbesondere, daß für ein $P \in \mathbb{R}[x] \subseteq \mathbb{C}[x]$ zusammen mit $P_n(c) = 0$ stets auch $P_n(\bar{c}) = 0$ gilt.

Damit ist eine Nullstelle von P_n entweder reell oder die zu ihr komplex konjugierte Zahl ist ebenfalls eine Nullstelle.

Ziel: Satz 5 zu beweisen

Satz 5 Sei $f : V \rightarrow V$ ein Endomorphismus von n -dimensionalen \mathbb{R} -Vektorraum V . Angenommen, die komplexifizierung $f_{\mathbb{C}}$ von f ist diagonalisierbar (als Endomorphismus von \mathbb{C} -Vektorraum $V_{\mathbb{C}}$.) Dann gibt es ein Basis B in V sodass die Matrix von f die folgende Form hat

$$\left(\begin{array}{c|c|c} \boxed{\begin{matrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_k \end{matrix}} & & \\ \hline & \boxed{\begin{matrix} \alpha_1 & \beta_1 \\ -\beta_1 & \alpha_1 \end{matrix}} & \\ \hline & & \ddots \\ \hline & & \boxed{\begin{matrix} \alpha_m & \beta_m \\ -\beta_m & \alpha_m \end{matrix}} \end{array} \right), \text{ wobei } \beta_j \neq 0 \text{ ist.}$$

(k oder m können auch gleich 0 sein. Selbstverständlich gilt $2 \cdot m + k = \dim(V)$.)

Bemerkung. In LAAG I hatten wir zwei Diagonalisierbarkeitkriterien, Sätze 55, 58. Mit Hilfe von diesen Kriterien kann man entscheiden, ob eine Matrix $A \in \text{Mat}(n, n, \mathbb{R}) \subseteq \text{Mat}(n, n, \mathbb{C})$ diagonalisierbar ist.

Wir benutzen: Hauptsatz der Algebra

Hauptsatz der Algebra (Beweis in Analysis – Vorlesungen oder in Funktionentheorie) Jedes $P \in \mathbb{C}[x]$ mit $\text{Grad}(P) \geq 1$ hat mind. eine Nullstelle

Folgerung A Jedes $P \in \mathbb{C}[x]$, $P \neq 0$, kann man in lineare Faktoren zehrlegen (d.h. in der Form $P = a(x - x_1)(x - x_2)\dots(x - x_n)$ schreiben, wobei $a, x_i \in \mathbb{C}$ sind). Diese Zerlegung ist eindeutig bis zum umstellen von Faktoren.

Folgerung B Jedes $P \in \mathbb{R}[x]$, $\text{Grad}(P) > 0$, kann man in Produkt von lineare und quadratischen Faktoren $g_i \in \mathbb{R}[x]$ zerlegen: $P := g_1 g_2 \dots g_m$, wobei $\text{Grad}(g_i) \in \{1, 2\}$.

Beweis von Folgerung B

Hilfsaussage – letztes Mal bewiesen Es sei

$P = \sum_{k=0}^n a_k x^k \in \mathbb{R}[x] \subseteq \mathbb{C}[x]$. Dann gilt für jedes $z \in \mathbb{C}$: $P(\bar{z}) = \overline{P(z)}$
(wobei \bar{z} komplexe Konjugation ist)

Daraus folgt insbesondere, daß für ein $P \in \mathbb{R}[x] \subseteq \mathbb{C}[x]$ zusammen mit $P_n(c) = 0$ stets auch $P_n(\bar{c}) = 0$ gilt. **Damit ist eine Nullstelle von P_n entweder reell oder die zu ihr komplex konjugierte Zahl ist ebenfalls eine Nullstelle.**

Folgerung B Jedes $P \in \mathbb{R}[x]$, $\text{Grad}(P) > 0$, kann man in Produkt von lineare und quadratischen Faktoren $g_i \in \mathbb{R}[x]$ zerlegen: $P := g_1 g_2 \dots g_m$, wobei $\text{Grad}(g_i) \in \{1, 2\}$.

Beweis – Vorsetzung: Es sei nun $c = \alpha + i\beta$ mit $\alpha, \beta \in \mathbb{R}$ eine Nullstelle von P_n . Dann ist

$$(x - c)(x - \bar{c}) = (x - \alpha - i\beta)(x - \alpha + i\beta) = (x - \alpha)^2 + \beta^2 = x^2 + Ax + B \quad (*)$$

mit $A = -2\alpha \in \mathbb{R}$, $B = \alpha^2 + \beta^2 \in \mathbb{R}$. Wir dividieren P durch $x - c$ und durch $x - \bar{c}$, und erhalten wegen $(*)$ die Darstellung

$$P_n = (x - c)(x - \bar{c})Q_{n-2} = (x^2 + Ax + B)Q_{n-2}.$$

Da sowohl P_n als auch $x^2 + Ax + B$ ausschließlich reelle Koeffizienten besitzen, hat auch Q_{n-2} nur reelle Koeffizienten. Also läßt sich die Prozedur wiederholen. Nach endlich viel Schritten bekommen wir

$P := g_1 g_2 \dots g_{m_1} Q_{n-2m_1}$, wobei g_1, \dots, g_{m_1} quadratische Polynome $\in \mathbb{R}[x]$ sind, und Q hat nur reelle Nullstelle. Nach Folgerung A kann man Q_{n-2m_1} in Form $a(x - x_1) \dots (x - x_{n-2m_1})$ schreiben, wobei $a, x_i \in \mathbb{R}$. □

Bemerkung. OBdA haben die quadratische Polynome aus Folgerung B keine reellen Nullstellen.

Beweis. Falls ein quadratisches Polynom $g \in \mathbb{R}[x]$ aus der Zerlegung $f = g_1 \dots g_m$ eine reelle Nullstelle $c \in \mathbb{R}$ hat, dann ist nach Lemma 27 Vorl. Fricke LAAG I $g = (x - c)f$. Da $\text{Grad}(g) = \text{Grad}(f) + 1 = 2$, ist f jedenfalls ein lineares Polynom. □

Folgerung C Die Vielfachheit einer nichtreellen Nullstelle $\alpha + \beta \cdot i$ eines Polynoms $P \in \mathbb{R}[x]$ ist gleich die Vielfachheit der konjugierten Nullstelle $\alpha - \beta \cdot i$.

Beweis. Ist $\alpha + \beta \cdot i$ eine Nullstelle eines quadratischen Polynoms $g \in \mathbb{R}[x]$, so ist $\overline{\alpha + \beta \cdot i} = \alpha - \beta \cdot i$ die zweite Nullstelle von g . Also, jedes quadratische g aus der Zerlegung von P gibt uns zwei konjugierten Nullstellen. □

Beweis von Satz 5

Zuerst als Beispiel führen wir Beweis in Dim 2 durch.

Wir betrachten das charakteristische Polynom $\chi_{f_{\mathbb{C}}}$ vom Endomorphismus $f_{\mathbb{C}}$. Da die Matrix von $f_{\mathbb{C}}$ gleich die Matrix von f ist, ist $\chi_{f_{\mathbb{C}}} = \chi_f$, und deswegen sind die Koeffizienten des charakteristisches Polynoms reell. Deswegen sind alle Nulstellen von $\chi_{f_{\mathbb{C}}}$ reell (in dem Fall ist f bereits über \mathbb{R} diagonalisierbar nach Satz 58 LAAG I), oder hat einen Eigenwert $\alpha + \beta \cdot i$ mit $\beta \neq 0$.

Sei $u + v \cdot i$ ein Eigenvektor zu $\alpha + \beta \cdot i$. Wir zeigen: $u - v \cdot i$ ist auch ein Eigenvektor zu $\alpha - \beta \cdot i$.

In der Tat,

$$f_{\mathbb{C}}(u - v \cdot i) \stackrel{\text{Rechenregeln}}{=} \overline{f_{\mathbb{C}}(u + v \cdot i)} = \overline{(\alpha + \beta \cdot i)(u + v \cdot i)} \stackrel{\text{Rechenregeln}}{=} (\alpha - \beta \cdot i)(u - v \cdot i).$$

Da die Vektoren $u - v \cdot i$ und $u + v \cdot i$ Eigenvektoren mit verschiedenen Eigenwerten sind, sind sie linear unabhängig und bilden deswegen eine Basis.

Dann bilden die Vektoren $u + \vec{0} \cdot i$ und $v + \vec{0} \cdot i$ auch eine Basis in $V_{\mathbb{C}}$ (weil Anzahl der Vektoren gleich die Dimension des Raums ist, und die Menge ist erzeugend). Dann ist (u, v) eine Basis in V .

Offensichtlich gilt: $u + \vec{0} \cdot i = \frac{1}{2}(u + v \cdot i + u - v \cdot i)$,

$$v + \vec{0} \cdot i = -\frac{i}{2}(u + v \cdot i - (u - v \cdot i)),$$

Wir rechnen jetzt die Matrix der Abbildung $f_{\mathbb{C}}$ in der Basis.

$$f_{\mathbb{C}}(u + \vec{0} \cdot i) = f_{\mathbb{C}}\left(\frac{1}{2}(u + v \cdot i + u - v \cdot i)\right) \stackrel{\text{Linearität}}{=} \frac{1}{2}f_{\mathbb{C}}(u + v \cdot i) + \frac{1}{2}f_{\mathbb{C}}(u - v \cdot i)$$

$$\stackrel{\text{Weil Eigenvektoren}}{=} \frac{1}{2}(\alpha + \beta \cdot i)(u + v \cdot i) + \frac{1}{2}(\alpha - \beta \cdot i)(u - v \cdot i) = \alpha u - \beta v + \vec{0} \cdot i.$$

$$\stackrel{\text{Analog}}{=} f_{\mathbb{C}}(u + \vec{0} \cdot i) = \alpha v + \beta u + \vec{0} \cdot i.$$

$$\stackrel{\text{Analog}}{=} f_{\mathbb{C}}(u + \vec{0} \cdot i) = \alpha v + \beta u + \vec{0} \cdot i.$$

Dann ist die Matrix von $f_{\mathbb{C}}$ gleich $\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$ wie im Satz 5.

Beweis für eine beliebige Dimension n

Wir suchen eine Basis sodass die Matrix von f wie im Satz 5 ist.

Wiederholung: Satz 54 LAAG I: Eine Matrix ist diagonalisierbar \iff es gibt eine Basis, die nur aus Eigenvektoren besteht.

Angenommen ist $f_{\mathbb{C}}$ diagonalisierbar. Seien $\lambda_1, \dots, \lambda_s$ die reelle Eigenwerten von $f_{\mathbb{C}}$.

Wir zeigen: $geo_{f_{\mathbb{C}}}(\lambda_i) = geo_f(\lambda_i)$. Sei A die Matrix von f (und deswegen auch von $f_{\mathbb{C}}$).

$$geo_f(\lambda_i) \stackrel{\text{Definition}}{=} \dim(\text{Kern}_{f - \lambda_i \cdot Id}) \stackrel{\text{1ste Dimensionsformel}}{=} n - \dim(\text{Bild}_{f - \lambda_i \cdot Id}) = n - rk(A - \lambda_i \cdot I) \stackrel{\text{Weil die Matrix von } f_{\mathbb{C}} \text{ auch } A \text{ ist}}{=} geo_{f_{\mathbb{C}}}(\lambda_i).$$

Dann gibt es **reelle** linear unabhängige Eigenvektoren $b_1^i, \dots, b_{alg_f(\lambda_i)}^i$ zu λ_i . (Offensichtlich ist $alg_f(\lambda_i) = alg_{f_{\mathbb{C}}}(\lambda_i)$).

Jetzt betrachten wir die komplexen Eigenwerte von $f_{\mathbb{C}}$. Da das Polynom $\mathfrak{N}_{f_{\mathbb{C}}} = \mathfrak{N}_f \in \mathbb{R}[x]$, gilt:

Für jeden Eigenwert $\alpha + \beta \cdot i$ ist $\overline{\alpha + \beta \cdot i} = \alpha - \beta \cdot i$ auch ein Eigenwert (Hilfsaussage zu Folgerung B).

Angenommen, $\alpha_1 + \beta_1 \cdot i, \alpha_1 - \beta_1 \cdot i, \dots, \alpha_m + \beta_m \cdot i, \alpha_m - \beta_m \cdot i$ sind die komplexe Eigenwerten. Hier stets $\beta_j \neq 0$.

Seien $u_1 + v_1 \cdot i, \dots, u_{\text{alg}_{f_{\mathbb{C}}}(\alpha + \beta i)} + v_{\text{alg}_{f_{\mathbb{C}}}(\alpha + \beta i)} \cdot i$ linear unabhängige Eigenvektoren zum Eigenwert $\alpha + \beta \cdot i$. Dann gilt:

(i) die Vektoren

$$\overline{u_1 + v_1 \cdot i} = u_1 - v_1 \cdot i, \dots,$$

$$\overline{u_{\text{alg}_{f_{\mathbb{C}}}(\alpha + \beta i)} + v_{\text{alg}_{f_{\mathbb{C}}}(\alpha + \beta i)} \cdot i} = u_{\text{alg}_{f_{\mathbb{C}}}(\alpha + \beta i)} + v_{\text{alg}_{f_{\mathbb{C}}}(\alpha + \beta i)} \cdot i \text{ sind}$$

$$\text{Eigenvektoren zu } \alpha - \beta \cdot i = \overline{\alpha + \beta \cdot i}.$$

(ii) Diese Eigenvektoren sind ebenfalls linear unabhängig (und deswegen bilden eine Basis in $\text{Eig}_{\alpha - \beta i}$).

(i): wie in Dim 2: $f_{\mathbb{C}}(u - v \cdot i) \stackrel{\text{Rechenregeln}}{=} \overline{f_{\mathbb{C}}(u + v \cdot i)} =$
 $\overline{(\alpha + \beta \cdot i)(u + v \cdot i)} \stackrel{\text{Rechenregeln}}{=} (\alpha - \beta \cdot i)(u - v \cdot i).$

(ii): Angenommen,

$(\gamma_1 + \delta_1 \cdot i)(u_1 - v_1 \cdot i) + \dots + (\gamma_k + \delta_k \cdot i)(u_k - v_k \cdot i) = \vec{0} + \vec{0} \cdot i.$ Wir konjugieren die beiden Seiten:

$$\overline{(\gamma_1 + \delta_1 \cdot i)(u_1 - v_1 \cdot i) + \dots + (\gamma_k + \delta_k \cdot i)(u_k - v_k \cdot i)} \stackrel{\text{Rechenregeln}}{=} (\gamma_1 - \delta_1 \cdot i)(u_1 + v_1 \cdot i) + \dots + (\gamma_k - \delta_k \cdot i)(u_k + v_k \cdot i) = \vec{0} + \vec{0} \cdot i = \vec{0} + \vec{0} \cdot i.$$

Da die Vektoren $u_j + v_j \cdot i$ linear unabhängig sind, sind die Koeffiziente $\gamma_j - \delta_j \cdot i = 0$ folglich $\gamma_j + \delta_j \cdot i = 0.$

Jetzt betrachten wir das Tupel

$$\underbrace{(b_1^1 + \vec{0} \cdot i, \dots, b_{\text{alg}_f(\lambda_j)}^1 + \vec{0} \cdot i, \dots, b_1^s + \vec{0} \cdot i, \dots, b_{\text{alg}_f(\lambda_j)}^s + \vec{0} \cdot i)}_{\text{Basis in } \text{Eig}_{\lambda_1}(f)} \quad \underbrace{u_1^1 + v_1^1 \cdot i, \dots, u_{\text{alg}_f(\alpha_1 + \beta_1 \cdot i)}^1 + v_{\text{alg}_f(\alpha_1 + \beta_1 \cdot i)}^1 \cdot i}_{\text{Basis in } \text{Eig}_{\lambda_s}(f)} \quad \underbrace{u_1^1 + v_1^1 \cdot i, \dots, u_{\text{alg}_f(\alpha_1 + \beta_1 \cdot i)}^1 + v_{\text{alg}_f(\alpha_1 + \beta_1 \cdot i)}^1 \cdot i}_{\text{Basis in } \text{Eig}_f(\alpha_1 + \beta_1 \cdot i)},$$

$$\underbrace{(u_1^1 - v_1^1 \cdot i, \dots, u_{\text{alg}_f(\alpha_1 + \beta_1 \cdot i)}^1 - v_{\text{alg}_f(\alpha_1 + \beta_1 \cdot i)}^1 \cdot i, \dots)}_{\text{Basis in } \text{Eig}_{\alpha_1 - \beta_1 \cdot i}(f)}.$$

Die Vektoren sind linear unabhängig, weil die Vektoren aus verschiedenen Eigenräumen nach Satz 53 LAAG I linear unabhängig sind, und weil sie in jedem Eigenraum eine Basis bilden und deswegen auch linear unabhängig sind. Dann bilden sie eine Basis in $V_{\mathbb{C}}$.

Jetzt können wir die Basis konstruieren, sodass die Matrix von A wie im Satz 5 ist: wir betrachten das Tupel

$$\left(\underbrace{b_1^1, \dots, b_{\text{alg}_f(\lambda_1)}^1}_{\text{Basis in } \text{Eig}_{\lambda_1}(f)}, \dots, \underbrace{b_1^s, \dots, b_{\text{alg}_f(\lambda_s)}^s}_{\text{Basis in } \text{Eig}_{\lambda_s}(f)}, \dots, \right. \\ \left. \underbrace{u_1^1, v_1^1, \dots, u_{\text{alg}_{\alpha_1+\beta_1 \cdot i}(f)}^1, v_{\text{alg}_{\alpha_1+\beta_1 \cdot i}(f)}^1}_{\text{reelle und imaginäre Anteil von Basisvektoren in } \text{Eig}_{\alpha_1+\beta_1 \cdot i}}, \dots, \right.$$

reelle und imaginäre Anteil von Basisvektoren in $\text{Eig}_{\alpha_1+\beta_1 \cdot i}$

$$\left. u_1^m, v_1^m, \dots, u_{\text{alg}_{\alpha_m+\beta_m \cdot i}(f)}^m, v_{\text{alg}_{\alpha_m+\beta_m \cdot i}(f)}^m \right).$$

Anzahl Elementen im Tupel ist gleich

$$\text{alg}_{\lambda_1} + \dots + \text{alg}_{\lambda_s} + 2\text{alg}_{\alpha_1+\beta_1 \cdot i} + \dots + 2\text{alg}_{\alpha_m+\beta_m \cdot i} \stackrel{\text{Folgerung A}}{=} n.$$

Die (komplexifizierung) von Vektoren sind erzeugen in $V_{\mathbb{C}}$: man kann mit deren Hilfe die Basiselemente von komplexen Basis erzeugen:

$$u_i + \vec{0} \cdot i + i \cdot (v_i + \vec{0} \cdot i) = u_i + v_i \cdot i,$$

und deswegen auch alle Vektoren von $V_{\mathbb{C}}$.

Dann ist das Tupel eine Basis in V .

Lass uns die Matrix von f in der Basis ausrechnen:

$f(b_j) = \lambda_j b_j$, da die Vektoren b_j aus der Basis Eigenvektoren sind, folglich ist der Koordinatenvektor von $f(b_j) = e_j \lambda_j$, folglich ist die j -te Spalten der Matrix wie im Satz.

$$\begin{aligned} f(u_i) &= \frac{1}{2} f_{\mathbb{C}}(u_i + v_i \cdot i + u_i - v_i \cdot i) = \\ &= \frac{1}{2} (f_{\mathbb{C}}(u_i + v_i \cdot i) + f_{\mathbb{C}}(u_i - v_i \cdot i)) = \\ &= \frac{1}{2} ((\alpha_i + \beta_i \cdot i)(u_i + v_i \cdot i) + (\alpha_i - \beta_i \cdot i)(u_i - v_i \cdot i)) \stackrel{\text{Ausrechnen}}{=} \end{aligned}$$

$\alpha_i u_i - \beta_i v_i$. Dann ist die entsprechende Spalte der Matrix wie im Satz. Analog:

$$\begin{aligned} f(v_i) &= -\frac{i}{2} f_{\mathbb{C}}(u_i + v_i \cdot i - (u_i - v_i \cdot i)) = \\ &= -\frac{i}{2} (f_{\mathbb{C}}(u_i + v_i \cdot i) - f_{\mathbb{C}}(u_i - v_i \cdot i)) = \\ &= -\frac{i}{2} ((\alpha_i + \beta_i \cdot i)(u_i + v_i \cdot i) - (\alpha_i - \beta_i \cdot i)(u_i - v_i \cdot i)) \stackrel{\text{Ausrechnen}}{=} \end{aligned}$$

$\alpha_i v_i + \beta_i u_i$. Dann ist die entsprechende Spalte der Matrix wie im Satz. □

Anwendung von Körpererweiterung: Konstruktionen mit Zirkel und Lineal:

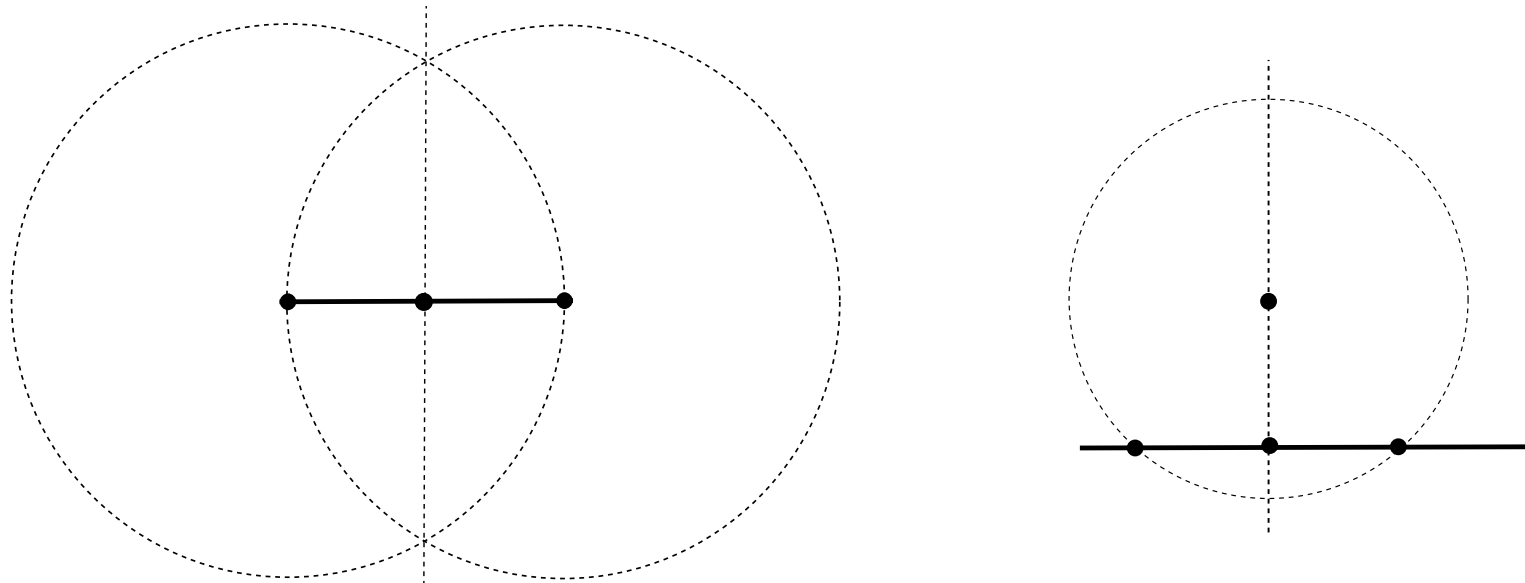
Frage: (Euklid) Welche geometrischen Objekten sind allein mit Zirkel und Lineal konstruierbar?

Wir definieren den Begriff „**konstruierbar**“ durch die folgenden Festlegungen: (Was darf man machen?)

(a) Die Gerade durch zwei gegebene verschiedene Punkte ist konstruierbar. (b) Der Kreis um einen gegebenen Punkt dessen Radius gleich Abstand zwischen zwei gegebenen Punkten ist konstruierbar. (c) Der Schnittpunkt von zwei sich schneidenden Geraden, (d) die Schnittpunkte eines gegebenen Kreises und einer den Kreis schneidenden gegebenen Geraden, (e) Die Schnittpunkte von zwei sich schneidenden gegebenen Kreisen sind konstruierbar.

Geometrische Gebilde (wie z.B. Punkte, Geraden, Strecken, Kreise, Dreiecke, Polygone,) die jeweils durch eine endliche Punktmenge festgelegt werden können, wollen wir vorübergehend als „**Objekte**“ bezeichnen. Wir sagen dann, das Objekt a sei bei Vorgabe der Objekte a_1, \dots, a_k **konstruierbar**, wenn es Objekte $a_{k+1}, \dots, a_n = a$ gibt, so dass a_j bei Vorgabe der Objekte a_1, \dots, a_{j-1} konstruierbar ist für $j = k + 1, \dots, n$.

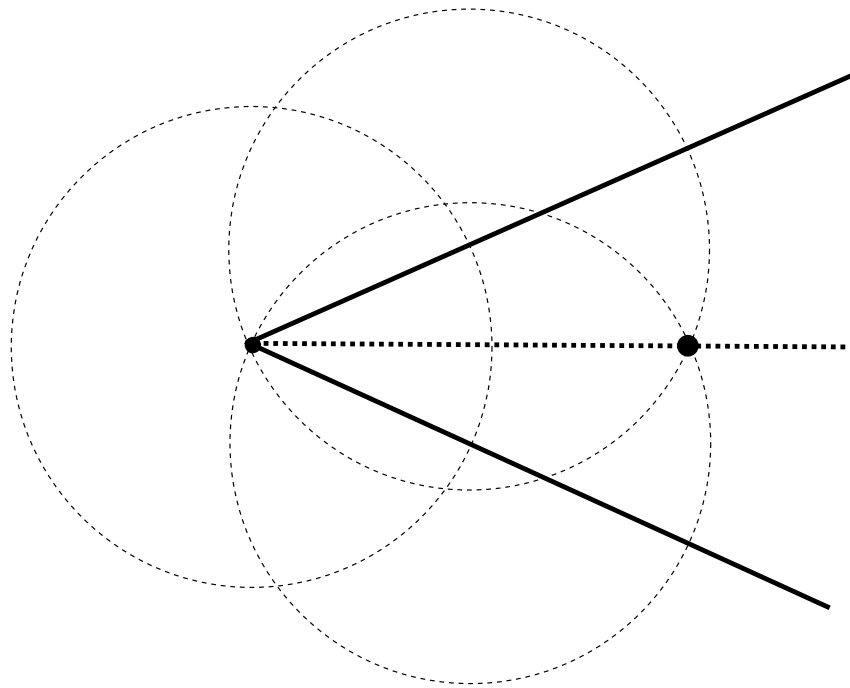
Grundkonstruktionen aus der Schule: Mittelpunkt einer gegebenen Strecke ist konstruierbar.



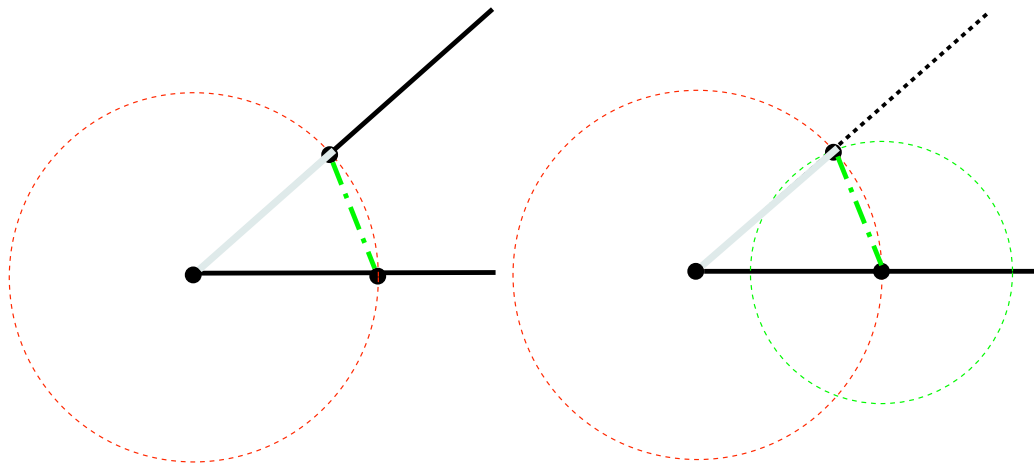
Wir haben mehr gemacht: wir haben die Gerade konstruiert, die zur gegebenen Strecke orthogonal ist. Deswegen ist

- ▶ die Senkrechte in einem gegebenen Punkt einer gegebenen Geraden konstruierbar.
- ▶ Projektion eines Punktes auf einer Gerade ist konstruierbar.

Winkelhalbierende eines gegebenen Winkels ist konstruierbar.

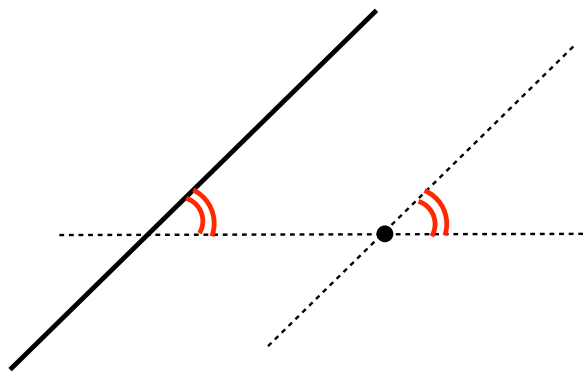


An einem gegebenen Strahl ist vom Anfangspunkt aus der Winkel abzutragen, der die gleiche Größe hat wie ein gegebener Winkel.



Ähnlich, eine Gerade durch einen gegebenen Punkt, die zu einer gegebenen Geraden parallel ist, ist konstruierbar.

An einem gegebenen Strahl ist vom Anfangspunkt aus der Winkel abzutragen, der die gleiche Größe hat wie ein gegebener Winkel.



Ähnlich, eine Gerade durch einen gegebenen Punkt, die zu einer gegebenen Geraden parallel ist, ist konstruierbar.

Konstruierbare Zahlen

Def. 4 Die Zahl $a \in \mathbb{R}$ heißt *konstruierbar*, wenn bei gegebener Strecke der Länge 1 eine Strecke der Länge $|a|$ konstruierbar ist.

Satz 6. Sind die Zahlen $a, b \in \mathbb{R}$ konstruierbar, so auch die Zahlen $a + b, a - b, ab, a/b$ (falls $b \neq 0$), \sqrt{a} (falls $a > 0$).

Beweis. Seien Strecken der Längen 1, a, b gegeben. Die Konstruktion von Strecken der Längen $a + b$ und $a - b$ ist trivial. Die Konstruktion von Strecken der Längen a/b und ab läßt sich an den folgenden ähnlichen Dreiecken ablesen:



Solches ähnliches Dreieck ist konstruierbar, weil eine Gerade durch einen gegebenen Punkt, die zu einer gegebenen Geraden parallel ist, konstruierbar ist.

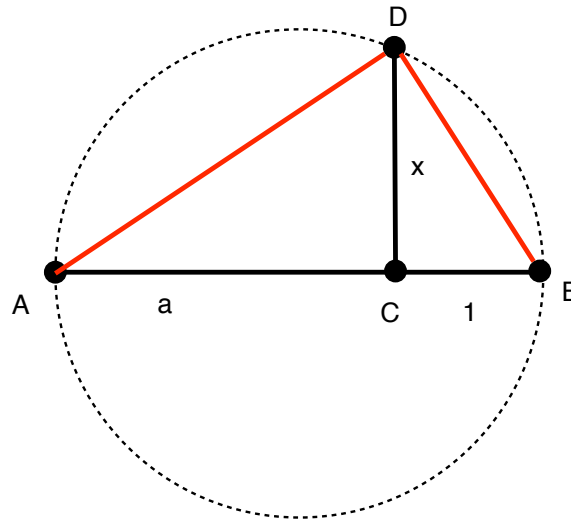
Konstruktion von \sqrt{a}

Konstruiere die Strecke AB der Länge $a + 1$.

Konstruiere den Kreis von Radius $(a+1)/2$ um dem Mittelpunkt der Strecke.

Konstruiere die Gerade durch C , die orthogonal zu AB ist. Sei D ein Schnittpunkt der Geraden mit dem Kreis

Die Länge von CD ist \sqrt{a} . Tatsächlich, der Winkel ADB ist $\frac{\pi}{2}$. (Wird auf der nächsten Folie erklärt)



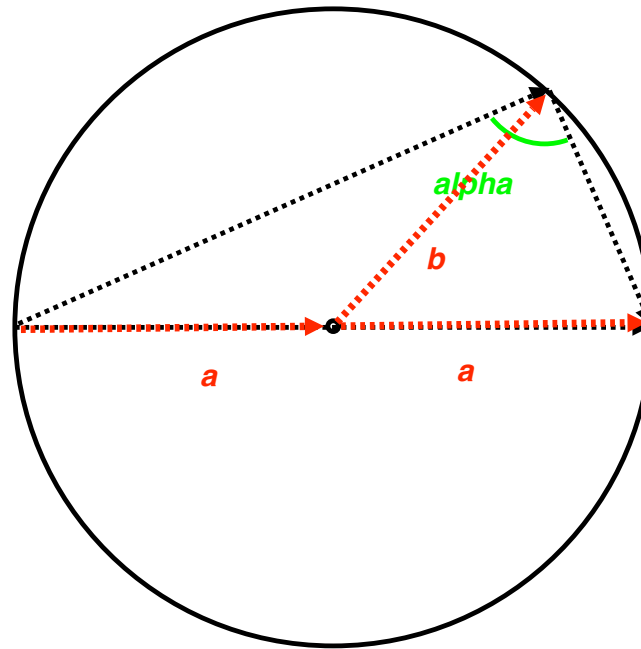
Nach Pythagoras ist

$$\begin{aligned} |AD|^2 &+ |BD|^2 &= |AB|^2 \\ x^2 + a^2 &+ x^2 + 1 &= (a + 1)^2. \end{aligned}$$

Dann $x^2 = a$, also $x = \sqrt{a}$



Z.Z. Im Dreieck auf dem Bild ist der Winkel *alpha* gleich $\pi/2$.



Beweis. Betrachte die Vektoren \vec{a}, \vec{b} wie auf dem Bild. Wir zeigen, dass

$$\langle \underbrace{\vec{a} + \vec{b}}_{\vec{u}}, \underbrace{\vec{a} - \vec{b}}_{\vec{v}} \rangle = 0. \quad \text{Wegen Bilinearität und Symmetrie, ist}$$

$$\langle \vec{a} + \vec{b}, \vec{a} - \vec{b} \rangle = \langle \vec{a}, \vec{a} \rangle + \underbrace{\langle \vec{b}, \vec{a} \rangle - \langle \vec{a}, \vec{b} \rangle}_{=0, \text{ Symmetrie}} - \langle \vec{b}, \vec{b} \rangle = \langle \vec{a}, \vec{a} \rangle - \langle \vec{b}, \vec{b} \rangle =$$

$$|a|^2 - |b|^2 = 0. \quad \text{Dann ist } \arccos\left(\frac{\langle \vec{u}, \vec{v} \rangle}{|\vec{u}| |\vec{v}|}\right) = \arccos(0) = \frac{\pi}{2}. \quad \square$$

Def. 5 Wir sagen, dass ein Unterkörper $\mathbb{K} \subseteq \mathbb{R}$ eine **iterierte quadratische Erweiterung** von \mathbb{Q} ist, falls es eine endliche Folge von quadratischen Erweiterungen

$\mathbb{K}_0, \mathbb{K}_1 = \mathbb{K}_0(\sqrt{s_1}), \mathbb{K}_2 = \mathbb{K}_1(\sqrt{s_2}), \dots, \mathbb{K}_k = \mathbb{K}_{k-1}(\sqrt{s_k}) \subseteq \mathbb{R}$ gibt, sodass $\mathbb{K}_0 = \mathbb{Q}$ und $\mathbb{K}_k = \mathbb{K}$.

Folgerung *Liegt $a \in \mathbb{R}$ in einer iterierten quadratischen Erweiterung von \mathbb{Q} , so ist a konstruierbar.*

Beweis: Hausaufgabe.

Satz 7. Eine Zahl $a \in \mathbb{R}$ ist genau dann konstruierbar, wenn a enthalten ist in einer iterierten quadratischen Erweiterung von \mathbb{Q} .

Bemerkung Ist die Zahl a konstruierbar, so ist bei gegebener Strecke AB eine Strecke der Länge $|a| \cdot |AB|$ auch konstruierbar.

Beweis: „ \Leftarrow “ ist Folgerung aus Satz 6 und Ihre Hausaufgabe. Wir beweisen in „ \Rightarrow “. Wir werden die Konstruktionsschritte (a)–(e) in Standard-Koordinaten $\begin{pmatrix} x \\ y \end{pmatrix}$ auf $(\mathbb{R}^2, \langle \cdot, \cdot \rangle)$ nachvollziehen.

Es genügt zu zeigen: Sind p_1, \dots, p_n Punkte, deren Koordinaten in einem Körper $K \subseteq \mathbb{R}$ liegen, und ist der Punkt p konstruierbar aus p_1, \dots, p_n , so liegen die Koordinaten von p in einer iterierten quadratischen Erweiterung von K .

Tatsächlich, oBdA sind $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ die Eckpunkte der gegebenen Strecke der Länge 1. Deren Koordinaten liegen also in \mathbb{Q} . Falls die Aussage oben richtig ist, liegen die Koordinaten jedes konstruierbaren Punkts in einer iterierten quadratischen Erweiterung von \mathbb{Q} . Dann ist die Länge jeder konstruierbaren Strecke gleich

$$\sqrt{\underbrace{(x_1 - x_2)^2}_{\text{in einer iter. quadr. Erweiterung von } \mathbb{Q}} + \underbrace{(y_1 - y_2)^2}_{\text{in einer iter. quadr. Erweiterung von } \mathbb{Q}}} \in \begin{matrix} \text{iter. quadr.} \\ \text{Erweiterung} \\ \text{von } \mathbb{Q} \end{matrix}.$$

Es genügt nachzuprüfen, dass:

- (i) Schnittpunkt der Geraden $\mathcal{G}_1 := \left\{ \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + t \begin{pmatrix} x_2 - x_1 \\ y_2 - y_1 \end{pmatrix}, \text{ wobei } t \in \mathbb{R} \right\}$ und $\mathcal{G}_2 := \left\{ \begin{pmatrix} x_3 \\ y_3 \end{pmatrix} + s \begin{pmatrix} x_4 - x_3 \\ y_4 - y_3 \end{pmatrix}, \text{ wobei } s \in \mathbb{R} \right\}$, wobei $x_i, y_i \in \mathbb{K}$, in einer iterierten quadratischen Erweiterung von \mathbb{K} liegen.
- (ii) Schnittpunkte der Geraden $\mathcal{G}_1 := \left\{ \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + t \begin{pmatrix} x_2 - x_1 \\ y_2 - y_1 \end{pmatrix}, \text{ wobei } t \in \mathbb{R} \right\}$ und des Kreises um $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$, dessen Radius gleich Abstand zwischen $\begin{pmatrix} x_3 \\ y_3 \end{pmatrix}$ und $\begin{pmatrix} x_4 \\ y_4 \end{pmatrix}$, wobei $x_i, y_i \in \mathbb{K}$, in einer iterierten quadratischen Erweiterung von \mathbb{K} liegen.
- (iii) Schnittpunkte des Kreises um x_0 , dessen Radius gleich Abstand zwischen $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ und $\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$ ist, mit der Kreis um x'_0 , dessen Radius gleich Abstand zwischen $\begin{pmatrix} x'_1 \\ y'_1 \end{pmatrix}$ und $\begin{pmatrix} x'_2 \\ y'_2 \end{pmatrix}$ ist, in einer iterierten quadratischen Erweiterung von \mathbb{K} liegen.

(i)

Falls die Gerade $\mathcal{G}_1 := \left\{ \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + t \begin{pmatrix} x_2 - x_1 \\ y_2 - y_1 \end{pmatrix}, \text{ wobei } t \in \mathbb{R} \right\}$ und $\mathcal{G}_2 := \left\{ \begin{pmatrix} x_3 \\ y_3 \end{pmatrix} + s \begin{pmatrix} x_4 - x_3 \\ y_4 - y_3 \end{pmatrix}, \text{ wobei } s \in \mathbb{R} \right\}$, nicht parallel sind, ist der Schnittpunkt die Lösungsmenge des Systems (auf s, t)

$$\begin{cases} x_1 + t(x_2 - x_1) = x_3 + s(x_4 - x_3) \\ y_1 + t(y_2 - y_1) = y_3 + s(y_4 - y_3) \end{cases},$$

dessen Matrixform

$$\begin{pmatrix} x_2 - x_1 & -(x_4 - x_3) \\ y_2 - y_1 & -(y_4 - y_3) \end{pmatrix} \begin{pmatrix} t \\ s \end{pmatrix} = \begin{pmatrix} x_3 - x_1 \\ y_3 - y_1 \end{pmatrix} \quad \text{ist}$$

Da die Gerade nichtparallel sind, ist die Koeffizientenmatrix des Systems nichtausgeartet, also ist die Lösung

$$\begin{pmatrix} t \\ s \end{pmatrix} = \begin{pmatrix} x_2 - x_1 & -(x_4 - x_3) \\ y_2 - y_1 & -(y_4 - y_3) \end{pmatrix}^{-1} \begin{pmatrix} x_3 - x_1 \\ y_3 - y_1 \end{pmatrix} = \frac{1}{\det \begin{pmatrix} x_2 - x_1 & -(x_4 - x_3) \\ y_2 - y_1 & -(y_4 - y_3) \end{pmatrix}} \begin{pmatrix} -(y_4 - y_3) & (x_4 - x_3) \\ -(y_2 - y_1) & x_2 - x_1 \end{pmatrix} \begin{pmatrix} x_3 - x_1 \\ y_3 - y_1 \end{pmatrix}$$

Wir sehen, dass die Koordinaten des Schnittpunkts in \mathbb{K} liegen.

(ii)

Betrachte den Kreis um $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ mit Radius $r = \sqrt{(x_4 - x_3)^2 + (y_4 - y_3)^2}$.

Da $(x_4 - x_3)^2 + (y_4 - y_3)^2 \in \mathbb{K}$, liegt r in einer quadratischen Erweiterung von \mathbb{K} (in \mathbb{K} oder in $\mathbb{K}(\sqrt{(x_4 - x_3)^2 + (y_4 - y_3)^2})$).

Der Schnittpunkt der Geraden $\mathcal{G} := \left\{ \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + t \begin{pmatrix} x_2 - x_1 \\ y_2 - y_1 \end{pmatrix}, \text{ wobei } t \in \mathbb{R} \right\}$ mit dem Kreis ist der Punkt der Form $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + t \begin{pmatrix} x_2 - x_1 \\ y_2 - y_1 \end{pmatrix}$, der auf dem Kreis liegt, i.e.

$$(x_0 - x_1 - t(x_2 - x_1))^2 + (y_0 - y_1 - t(y_2 - y_1))^2 = r^2.$$

Dies ist eine quadratische Gleichung $at^2 + bt + c = 0$ auf t , deren Koeffizienten a, b, c Elemente von \mathbb{K} oder $\mathbb{K}(r)$ sind.

Deren Lösungen sind $t_{\pm} = -\frac{b}{2} \pm \sqrt{\left(\frac{b}{2}\right)^2 - ac}$. Sie liegen in einem iterierten quadratischen Erweiterung von \mathbb{K} .

Die Schnittpunkte der Geraden \mathcal{G}_1 und des Kreises sind die Punkte $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + t_{\pm} \begin{pmatrix} x_2 - x_1 \\ y_2 - y_1 \end{pmatrix}$. Deren Koordinaten liegen in einem iterierten quadratischen Erweiterung von \mathbb{K} .

(iii)

Den Kreis um $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ (bzw. um $\begin{pmatrix} x'_0 \\ y'_0 \end{pmatrix}$) dessen Radius gleich Abstands r zwischen $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ und $\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$ (bzw. Abstands r' zwischen $\begin{pmatrix} x'_1 \\ y'_1 \end{pmatrix}$ und $\begin{pmatrix} x'_2 \\ y'_2 \end{pmatrix}$) ist, ist die Lösungsmenge der Systems

$$\begin{cases} (x - x_0)^2 + (y - y_0)^2 - r^2 = 0, \\ (x - x'_0)^2 + (y - y'_0)^2 - r'^2 = 0. \end{cases}$$

Subtraktion ergibt

$$2x(x_0 - x'_0) + 2y(y_0 - y'_0) + (r^2 - x_0^2 - y_0^2) - (r'^2 - x_0'^2 - y_0'^2) = 0.$$

Da wir o.B.d.A. $(x_0, y_0) \neq (x'_0, y'_0)$ annehmen können, können wir y durch x (oder x durch y) ausdrücken, dies in eine der Kreisgleichungen einsetzen und dann die entstehende quadratische Gleichung lösen. In jedem Fall sind, um die Koordinaten der konstruierten Punkte aus den Koordinaten der gegebenen Punkte zu berechnen, nur rationale Operationen und das Ziehen einer Quadratwurzel erforderlich. Daraus liegen Sie in einer iterierten quadratischen Erweiterung von \mathbb{K} . □

Nichtkonstruierbarkeit.

Wir werden

1. Ummöglichkeit von 3 klassischen Konstruktionsproblemen besprechen:
 - ▶ die *Dreiteilung des Winkels* beweisen,
 - ▶ die *Verdoppelung des Würfels (Delisches Problem)* beweisen
 - ▶ und die *Quadratur des Kreises* (nur besprechen),
2. Unmöglichkeit von konstruieren von 7–Eck und 9–Eck beweisen.

Vorbereitungssätze

Frage Wie beweist man, dass eine Zahl nicht in einem iterierten quadratischen Erweiterung von \mathbb{Q} liegt?

Def. 6 Eine kubische Gleichung $x^3 + lx^2 + mx + n = 0$ heißt *irreduzibel*, wenn die Koeffizienten l, m, n rational sind, aber keine Lösung der Gleichung rational ist.

Satz 8 Ist die Zahl x Lösung einer irreduziblen kubischen Gleichung

$$x^3 + lx^2 + mx + n = 0, \quad (1)$$

so liegt x nicht in einer iterierten quadratischen Erweiterung von \mathbb{Q} .

Def. vor dem Beweis Liegt y in einem Körper, der durch k -malige quadratische Erweiterung aus \mathbb{Q} entsteht, so sagen wir, y sei *auf dem Niveau k* .

Bsp. $1/2$ ist auf dem Niveau 0, $1 + \sqrt{3}$ ist auf dem Niveau 1.

Widerspruchsbeweis. Angenommen, eine Lösung der Gleichung (1) läge in einer iterierten quadratischen Erweiterung von \mathbb{Q} . Sei x_1 die Lösung von (1), die auf dem kleinstem Niveau k ist. Da $x_1 \notin \mathbb{Q}$, ist $k \geq 1$. Es gilt also $x_1 = a + b\sqrt{s}$ mit geeigneten Zahlen a, b, s auf dem Niveau $k - 1$.

Wir setzen dies in (1) ein und erhalten

$$(a + b\sqrt{s})^3 + l(a + b\sqrt{s})^2 + m(a + b\sqrt{s}) + n = 0$$

$$\underbrace{(a^3 + 3ab^2s + a^2l + b^2sl + ma + n)}_A + \underbrace{(3a^2b + b^3s + 2abl + bm)}_B \sqrt{s} = 0.$$

A auf dem Niveau $k - 1$ B auf dem Niveau $k - 1$

Ist $A \neq 0 \neq B$, so ist \sqrt{s} auf dem Niveau $k - 1$ (weil $\sqrt{s} = -A/B$ ist, und nach deswegen nach Satz 6 auf dem Niveau $k - 1$ liegen muss), also ist x_1 nach Satz 6 auf dem Niveau $k - 1$, was Voraussetzungen widerspricht. Dann ist $A = B = 0$, und deswegen $x_2 := a - b\sqrt{s}$ auch eine Nullstelle der Gleichung (1), denn es ist

$$x_2^3 + lx_2^2 + mx_2 + n = (a^3 + 3ab^2s + a^2l + b^2sl + ma + n) - (3a^2b + b^3s + 2abl + bm)\sqrt{s} = 0.$$

Nach Satzgruppe von Viëta ist die Summe der drei Nullstellen von (1) gleich $-l$, also ist $-l - 2a$ ebenfalls eine Nullstelle. Sie ist auf dem Niveau $k - 1$. Das ist ein Widerspruch. □

Folgerung Um zu beweisen, dass eine Zahl nichtkonstruierbar ist, können wir zeigen, dass die Zahl eine Nullstelle einer irreduziblen kubischen Gleichung ist.

Satz 9 Sei $x^3 + lx^2 + mx + n = 0$ eine kubische Gleichung sodass $l, m, n \in \mathbb{Z}$. Dann gilt:

Diese Gleichung ist g.d. irreduzibel, wenn sie keine ganzzahlige Lösung hat.

Beweis. „ \implies “ ist offensichtlich: ist die Gleichung irreduzibel, so sind nach Def. 6 die Lösungen irrational.

Widerspruchsbeweis in „ \longleftarrow “ Angenommen, die Gleichung ist nicht irreduzibel, obwohl keine Lösungen ganzzahlig ist. Dann gibt es eine rationale Lösung $x = r/s$, wobei $r, s \in \mathbb{Z}$. OBdA ist $\text{ggT}(r, s) = 1$. Einsetzen in die Gleichung ergibt $r^3 = -s(lr^2 + smr + ns^2)$. Ist $|s| > 1$, so hat s einen Primfaktor p . Dieser muss auch Primfaktor von r^3 sein, und damit von r , ein Widerspruch. Also ist $s = 1$ und daher x ganzzahlige Lösung, im Widerspruch zur Voraussetzung. □

Dreiteilung des Winkels

Problem: *Kann man einen beliebige gegebene Winkel mit Zirkel und Lineal dreiteilen?*

Antwort: Nein.

Wäre die Winkeldreiteilung mit Zirkel und Lineal lösbar, so wäre insbesondere der Winkel $\pi/9$ konstruierbar, damit wäre $\cos(\pi/9)$ konstruierbar. Wegen der Identität

$$\cos 3\theta = \cos 2\theta \cos \theta - \sin 2\theta \sin \theta =$$

$$(2\cos^2\theta - 1)\cos\theta - 2\sin^2\theta \cos\theta = 4\cos^3\theta - 3\cos\theta$$

und wegen $\cos \frac{\pi}{3} = 1/2$ genügt die Zahl $c = 2\cos \frac{\pi}{9}$ der Gleichung

$$c^3 - 3c - 1 = 0.$$

Diese Gleichung hat keine ganzzahligen Lösungen (denn jede Lösung x erfüllt $x(x^2 - 3) = 1$, aber $x = \pm 1$ ist keine Lösung). Nach Satz 9 ist dann die Gleichung irreduzibel, und deswegen nach Satz 8 ist die Zahl $\cos(\pi/9)$ nicht konstruierbar. Die Unlösbarkeit der Winkeldreiteilung ist damit gezeigt. Zugleich ist mit diesem Beweis gezeigt, daß das reguläre 9-Eck nicht konstruierbar ist.

Verdoppelung des Würfels (Delisches Problem)

(Konstruktion eines Würfels mit dem doppelten Volumen eines vorgegebenen Würfels)

Der Sage nach wurde die Stadt Delos einmal von einer Seuche heimgesucht. Die Bewohner befragten ein Orakel und erhielten den Rat, einen ihrer Altäre zu verdoppeln. Plato interpretierte den Orakelspruch so, dass der würfelförmige Altar durch einen Würfel mit doppeltem Volumen ersetzt werden sollte. Er erklärte, Gott wolle die Griechen beschämen, weil sie das Studium der Mathematik vernachlässigt hätten. Daher ist die Verdopplung des Würfels auch als „Delisches Problem“ bekannt.

Wäre es lösbar, so könnte man aus einer Strecke der Länge 1 eine Strecke der Länge $\sqrt[3]{2}$ konstruieren. Nach Satz 6 liegt dann $\sqrt[3]{2}$ in einer iterierten quadratischen Erweiterung von \mathbb{Q} . Da $\sqrt[3]{2}$ eine Lösung der kubischen Gleichung

$$x^3 - 2 = 0$$

und diese nach Satz 9 irreduzibel ist, ist das ein Widerspruch zu Satz 8.

Quadratur des Kreises (ohne Beweis)

Aufgabe: *Mit Lineal und Zirkel aus einem gegebenen Kreis ein Quadrat mit demselben Flächeninhalt zu konstruieren.*

Satz (Lindemann 1882) *Das ist unmöglich*

Beweisidee: Wir müssen zeigen, daß die Zahl π in keiner iterierten quadratischen Erweiterung von \mathbb{Q} liegt. Dies folgt daraus,

- ▶ daß π **transzendent** ist (=nullstelle von keinem Polynom mit rationalen Koeffizienten),
- ▶ jede konstruierbare Zahl **algebraisch** (=nicht transzendent) ist.

Mit diesem Nachweis wurde das Problem der Quadratur des Kreises endgültig erledigt. Wir werden diese Aussage nicht beweisen. Sie dürfen sie trotzdem selbstverständlich u.a. in Hausaufgaben benutzen.

Beweis der Transzendenz von π /Quadratur des Kreises (von Rudolf Fritsch) <http://www.minet.uni-jena.de/matveev/Lehre/LA07/transzendens-von-pi.pdf>

Konstruierbarkeit von regulären n -Ecken

(**Reguläres** = alle Seiten und alle Winkel sind gleich)

Frage Welche regulären n -Ecken kann man mit Zirkel und Lineal konstruieren?

(Falls wir ein reguläres n -Eck konstruieren können, dann können wir ein reguläres n -Eck mit einer vorgegebenen Seite konstruieren.)

Bsp. Reguläres Dreieck, reguläres Viereck sind konstruierbar; reguläres 5-Eck ist ebenfalls konstruierbar – Beweis kommt.

Satz 10 *Reguläres n -Eck ist g.d. konstruierbar, wenn die Zahl $\cos(\frac{2\pi}{n})$ konstruierbar ist.*

Beweis. „ \Leftarrow “ Angenommen ist die Zahl $\cos(\frac{2\pi}{n})$ konstruierbar. Dann können wir der Winkel $\frac{2\pi}{n}$ konstruieren. Dann können wir einen Kreis in n gleichen Sektoren teilen, und so die Ecken eines regulären n -Eck konstruieren.

Beweis „ \Rightarrow “

Angenommen reguläres n -Eck ist konstruierbar. Dann können wir das n -Eck in einen Kreis einbeschreiben. (Für je zwei Seiten nehme die Geraden, die die Seiten orthogonal im Mittelpunkt schneiden. Deren Durchschnitt ist der Mittelpunkt des Kreises)

Dann können wir den Winkel $\frac{2\pi}{n}$ konstruieren, und deswegen die Zahl $\cos(\frac{2\pi}{n})$. □

Geometrische Konstruktionen und komplexe Zahlen.

Beobachtung $z := e^{\frac{2\pi}{n}i} \in \mathbb{C}$ ist eine Nullstelle der Gleichung
 $z^{n-1} + z^{n-2} + \dots + 1 = 0. \quad (*)$

Tatsächlich, $1, z, z^2, \dots, z^{n-1}$ ist die geometrische Progression, deren

Summe ist $\frac{z^n - 1}{z - 1} = \frac{(e^{\frac{2\pi}{n}i})^n - 1}{e^{\frac{2\pi}{n}i} - 1} = \frac{(e^{2\pi \cdot i}) - 1}{e^{\frac{2\pi}{n}i} - 1} = \frac{1 - 1}{e^{\frac{2\pi}{n}i} - 1} = 0.$

Bsp. Reguläres 5-Eck ist konstruierbar

Tatsächlich, für $n = 5$ ist die Gleichung $(*)$

$z^4 + z^3 + z^2 + z + 1 = 0.$ Wir können diese Gleichung lösen. Wir dividieren durch z^2 und bekommen

$z^2 + \frac{1}{z^2} + 2 + z + \frac{1}{z} - 1 = 0$ Da $(z + \frac{1}{z})^2 = z^2 + \frac{1}{z^2} + 2,$ ist die Gleichung äquivalent zu

$w^2 + w - 1 = 0,$ wobei $w = z + \frac{1}{z}.$ Dann ist $w = \frac{\pm\sqrt{5}-1}{2} \in \mathbb{Q}(\sqrt{5}).$

Dann ist $z + \frac{1}{z} = \frac{\pm\sqrt{5}-1}{2}.$ Das sind quadratische Gleichungen, deren Koeffizienten in $\mathbb{Q}(\sqrt{5})$ liegen. Dann haben die Nullstellen der Form

$z = a + ib,$ wobei $a \in \mathbb{Q}(\sqrt{5})$ ist. Da nach Beobachtung oben

$z = e^{\frac{2\pi}{5}i} = \cos(\frac{2\pi}{5}) + i \sin(\frac{2\pi}{5})$ eine Nullstelle ist, ist $\cos(\frac{2\pi}{5}) \in \mathbb{Q}(\sqrt{5}),$ und deswegen nach Satz 10 ist reguläres 5-Eck konstruierbar.

Reguläres 7-Eck ist nicht konstruierbar

Nach Satz 10 müssen wir zeigen, dass $\cos(2\pi/7)$ in keinem iterierten quadratischen Erweiterung von \mathbb{Q} liegt. Nach Beobachtung, ist $e^{\frac{2\pi}{7}i} = \cos(2\pi/7) + i \sin(2\pi/7)$ eine Lösung der Gleichung $z^6 + z^5 + \dots + 1 = 0$ (1).

Für die Zahl $y := z + \frac{1}{z}$ gilt daher

$$y^3 + y^2 - 2y - 1 = 0, \quad (2)$$

wie sich sofort durch Einsetzen in (1) und Umformung ergibt. Dann ist

$$y_1 := z_1 + \frac{1}{z_1} = e^{i2\pi/7} + e^{-i2\pi/7} = 2 \cos \frac{2\pi}{7}$$

eine Lösung von (2). Wäre nun das reguläre 7-Eck konstruierbar, so wäre die Zahl $\cos \frac{2\pi}{7}$ konstruierbar. Nach Satz 6 liegt y_1 dann in einer iterierten quadratischen Erweiterung von \mathbb{Q} . Nach Satz 8 folgt daraus, daß die Gleichung (2) nicht irreduzibel ist, sie hat also nach Satz 9 eine ganzzahlige Lösung. Die Lösungen von (2) sind neben $2 \cos \frac{2\pi}{7}$ noch die Zahlen $2 \cos \frac{4\pi}{7}$ und $2 \cos \frac{6\pi}{7}$ (die man genauso findet). Keine davon ist ganzzahlig. Das ist ein Widerspruch.

Jordan'sche Normalform

Ziel: Finde auch für nicht diagonalisierbare Endomorphismen eine möglichst einfache Matrixdarstellung.

Bsp. $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ ist nicht diagonalisierbar. Tatsächlich,

$$\chi_{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}} = \det \begin{pmatrix} -t & 1 \\ 0 & -t \end{pmatrix} = t^2.$$

Wir haben nur eine Nullstelle $\lambda = 0$. Kann man eine Basis aus

Eigenvektoren finden? Nein, weil $\text{Kern} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ hat nach Dimensionsformel

die Dimension $\underbrace{2}_{\dim(V)} - \underbrace{1}_{\text{rk}(A)} = 1$. Also, es gibt keine 2 linear unabhängige

Eigenvektoren, und die Matrix ist nach Satz 54 LAAG I nicht diagonalisierbar.

Wiederholung: Seien $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ paarweise verschieden, $k \in \mathbb{N}$, sei $f = (x - \lambda_1)^{\gamma_1} \dots (x - \lambda_k)^{\gamma_k} \in \mathbb{K}[x]$. Dann gilt:

(i) Falls ein Polynom $g \in \mathbb{K}[x]$ das Polynom f teilt, so ist es von der Form $g = \alpha(x - \lambda_1)^{\mu_1} \dots (x - \lambda_k)^{\mu_k}$ mit $\alpha \in \mathbb{K} \setminus \{0\}$ und $\mu_j \in \mathbb{N} \cup \{0\}$, $\mu_j \leq \gamma_j$.

(ii) Es gibt Polynome $h_1, h_2 \in \mathbb{K}[x]$, sodass

$$1 = h_1 \cdot \underbrace{(x - \lambda_1)^{\gamma_1}}_{f_1} + h_2 \underbrace{(x - \lambda_2)^{\gamma_2} \dots (x - \lambda_k)^{\gamma_k}}_{f_2}.$$

(i) \iff Lemma 33 Vorl. 20 LAAG I

(ii) \iff Satz 14 / Beweis von Satz 14 Vorl. Fricke LAAG I

Satz 11 Sei $f = (x - \lambda_1)^{\gamma_1} \dots (x - \lambda_k)^{\gamma_k} \in \mathbb{K}[x]$, wobei λ_i paarweise verschieden sind. Sei V ein endlichdimensionaler \mathbb{K} -Vektorraum. Sei $\phi : V \rightarrow V$ ein Endomorphismus. Dann ist

$$\text{Kern}_{f(\phi)} = \text{Kern}_{(\phi - \lambda_1 \cdot \text{Id})^{\gamma_1}} \oplus \dots \oplus \text{Kern}_{(\phi - \lambda_k \cdot \text{Id})^{\gamma_k}}.$$

Wiederholung: $(\phi - \lambda \cdot Id)(v) := \phi(v) - \lambda v.$

$$(\phi - \lambda \cdot Id)^2(v) = (\phi(v) - \lambda \cdot Id) \circ (\phi(v) - \lambda \cdot Id)(v) = \phi^2(v) - 2\lambda\phi(v) + \lambda^2 v \text{ u.s.w.}$$

Falls $A \in Mat(dim(V), dim(V), \mathbb{K})$ die Matrix von ϕ ist, so ist

$$A - \lambda Id = \begin{pmatrix} a_{11} - \lambda & \cdots & a_{1 \dim(V)} \\ \cdots & \ddots & \cdots \\ a_{\dim(V) 1} & \cdots & a_{\dim(V) \dim(V)} - \lambda \end{pmatrix} \text{ die Matrix von } \phi - \lambda \cdot Id .$$

Für $P = a_n x^n + \dots + a_0 \in \mathbb{K}[x]$ ist

$$P(\phi) = a_n \cdot \underbrace{\phi \circ \dots \circ \phi}_{n \text{ Stück}} + \dots + a_0 \cdot Id \text{ ein Endomorphismus von } V.$$

Falls $A \in Mat(dim(V), dim(V), \mathbb{K})$ die Matrix von ϕ ist, so ist die Matrix von $P(\phi)$ die Matrix $a_n A^n + \dots + a_0 \cdot Id.$

Bemerkung $P(\phi) \circ Q(\phi) = (P \cdot Q)(\phi) = (Q \cdot P)(\phi).$ (Obwohl

$$P(\phi) \cdot Q(\psi) \stackrel{\text{in der Regel}}{\neq} Q(\psi) \cdot P(\phi).)$$

Wiederholung – Def. 49 Vorl. 19 LAAG I Seien V_1, \dots, V_k Untervektorräume von V . Die Menge $W := \{v_1 + \dots + v_k : \text{wobei } v_i \in V_i\}$ heißt **die Summe** von V_i und wird $V_1 + \dots + V_k$ bezeichnet. Das ist ein Untervektorraum (Hausaufgabe 1 Blatt 11 LAAG I) Wir sagen, dass die Summe **direkt** ist (Bezeichnung: $W = V_1 \oplus \dots \oplus V_k$), falls jedes $w \in W$ schreibt sich **eindeutig** in der Form $w = v_1 + \dots + v_k$ mit $v_i \in V_i$

Satz 11 Sei $f = (x - \lambda_1)^{\gamma_1} \dots (x - \lambda_k)^{\gamma_k} \in \mathbb{K}[x]$, wobei λ_i paarweise verschieden sind. Sei V ein endlichdimensionaler \mathbb{K} -Vektorraum. Sei $\phi : V \rightarrow V$ ein Endomorphismus. Dann ist
$$\text{Kern}_{f(\phi)} = \text{Kern}((\phi - \lambda_1 \cdot \text{Id})^{\gamma_1}) \oplus \dots \oplus \text{Kern}((\phi - \lambda_k \cdot \text{Id})^{\gamma_k}).$$

Beweis. Induktion nach k . **IA:** Für $k = 1$ ist die Aussage offensichtlich: links und rechts stehen gleiche Ausdrücke.

IV: die Aussage gelte für alle $k - 1$.

IS: $k - 1 \rightarrow k$: Setze wieder $f_1 := (x - \lambda_1)^{\gamma_1}$,
 $f_2 := (x - \lambda_2)^{\gamma_2} \dots (x - \lambda_k)^{\gamma_k}$.

Wie oben wiederholt $\implies \exists h_1, h_2 \in \mathbb{K}[x]$ sodass $h_1 \cdot f_1 + h_2 \cdot f_2 = 1$.

Wir zeigen: $\text{Kern}_{f(\phi)} = \text{Kern}_{f_1(\phi)} \oplus \text{Kern}_{f_2(\phi)}$. Dann folgt die Aussage nach **InduktionsVoraussetzung** für $\text{Kern}_{f_2(\phi)}$.

Schema.

Wir zeigen $\overset{(a)}{\text{Kern}_{f_1(\phi)} \subseteq \text{Kern}_f(\phi)}$, $\overset{(b)}{\text{Kern}_{f_1(\phi)} + \text{Kern}_{f_2(\phi)} \supseteq \text{Kern}_f(\phi)}$.

Daraus folgt, dass $\text{Kern}_{f_1(\phi)} + \text{Kern}_{f_2(\phi)} = \text{Kern}_f(\phi)$.

Dann ((c)) zeigen wir, dass $\text{Kern}_{f_1(\phi)} \cap \text{Kern}_{f_2(\phi)} = \{\vec{0}\}$.

Daraus wird folgen, dass die Summe direkt ist.

(a) **Behauptung:** $\text{Kern}_{f_1(\phi)} \subseteq \text{Kern}_f(\phi)$.

Beweis von (a): ist $v \in \text{Kern}_{f_1(\phi)}$, so ist

$$\begin{aligned} f(\phi)(v) &= (f_1(\phi) \circ f_2(\phi))(v) \stackrel{\text{Bemerkung oben}}{=} \\ &= (f_2(\phi) \circ f_1(\phi))(v) \stackrel{\text{Weil } \text{Kern}_{f_1(\phi)} \subseteq \text{Kern}_f(\phi)}{=} f_2(\phi)(\vec{0}) = \vec{0}. \implies \\ &\text{Kern}_{f_1(\phi)} \subseteq \text{Kern}_f(\phi). \end{aligned}$$

Analog: $\text{Kern}_{f_2(\phi)} \subseteq \text{Kern}_f(\phi)$.

Es folgt: $\text{Kern}_{f_1(\phi)} + \text{Kern}_{f_2(\phi)} \subseteq \text{Kern}_f(\phi)$.

(b) Behauptung: $\text{Kern}_{f_1(\phi)} + \text{Kern}_{f_2(\phi)} \supseteq \text{Kern}_f(\phi)$.

Beweis von (b): Sei $v \in \text{Kern}_f(\phi)$.

Wie wir oben wiederholt haben, gibt es $h_1, h_2 \in \mathbb{K}[x]$, sodass

$$1 = h_1 \cdot f_1 + h_2 \cdot f_2. \text{ Da } 1(\phi) \stackrel{\text{Def}}{=} Id, \text{ folgt}$$
$$v = Id(v) = \underbrace{h_1(\phi) \circ f_1(v)}_{:=v_2} + \underbrace{h_2(\phi) \circ f_2(v)}_{:=v_1}.$$

Wir haben:

$$f_2(\phi)(v_2) = f_2(\phi) \circ h_1(\phi) \circ f_1(v) = h_1(\phi) \circ f(\phi)(v) = h_1(\phi)(\vec{0}) = \vec{0},$$

also $v_2 \in \text{Kern}_{f_2(\phi)}$.

Analog: $v_1 \in \text{Kern}_{f_1(\phi)}$.

Dann ist jedes v eine Summe von $v_1 \in \text{Kern}_{f_1(\phi)}$ und $v_2 \in \text{Kern}_{f_2(\phi)}$, folglich ist $\text{Kern}_{f_1(\phi)} + \text{Kern}_{f_2(\phi)} \supseteq \text{Kern}_f(\phi)$.

(c) Behauptung: $\text{Kern}_{f_1(\phi)} \cap \text{Kern}_{f_2(\phi)} = \{\vec{0}\}$.

Beweis von (c). Sei $v \in \text{Kern}_{f_1(\phi)} \cap \text{Kern}_{f_2(\phi)}$. Dann ist

$$v = \text{Id}(v) \stackrel{\text{wie in (b)}}{=} \underbrace{h_1(\phi) \circ f_1(v)}_{:=\vec{0}} + \underbrace{h_2(\phi) \circ f_2(v)}_{:=\vec{0}} = \vec{0}.$$

As (c) folgt, dass die Summe $\text{Kern}_{f_1(\phi)} + \text{Kern}_{f_2(\phi)}$ direkt ist. Nach Definition müssen wir zeigen, dass jedes v eindeutig als Summe

$$\underbrace{v_1}_{\in \text{Kern}_{f_1(\phi)}} + \underbrace{v_2}_{\in \text{Kern}_{f_2(\phi)}} \text{ darstellbar ist. Sei}$$
$$v = \underbrace{v_1}_{\in \text{Kern}_{f_1(\phi)}} + \underbrace{v_2}_{\in \text{Kern}_{f_2(\phi)}} = \underbrace{v'_1}_{\in \text{Kern}_{f_1(\phi)}} + \underbrace{v'_2}_{\in \text{Kern}_{f_2(\phi)}}.$$

Dann ist $v_1 - v'_1 \in \text{Kern}_{f_1(\phi)} = v'_2 - v_2 \in \text{Kern}_{f_2(\phi)}$.

Wegen $\text{Kern}_{f_1(\phi)} \cap \text{Kern}_{f_2(\phi)} = \{\vec{0}\}$ ist dann $v_1 - v'_1 = \vec{0}$ und

$$v_2 - v'_2 = \vec{0},$$

□

Gleichzeitig haben wir Induktionsschritt gemacht und deswegen Satz 11 bewiesen.

Folgerung = Satz 58 Vorl. 21 LAAG I:

$Min_A = (\lambda_1 - x) \dots (\lambda_k - x)$, wobei $\lambda_i \iff A$ ist diagonalisierbar.
paarweise verschieden sind

Die Richtung „ \Leftarrow “ ist einfach: man muss zeigen, dass Minimalpolynom einer Diagonalmatrix die Form $(\lambda_1 - x) \dots (\lambda_k - x)$ mit paarweise verschieden λ_i hat, siehe Vorl. 21 LAAG I, falls es nicht offensichtlich ist.

Beweis „ \implies “: Sei ϕ der Endomorphismus mit der Matrix A . Da $\text{Min}_A(A) = \mathbf{0}$, ist $\text{Kern}_{\text{Min}_A(f)} = V$. Nach Satz 11 ist

$$V = \underbrace{\text{Kern}_{f - \lambda_1 \cdot \text{Id}}}_{= \text{Eig}_{\lambda_1}} \oplus \cdots \oplus \underbrace{\text{Kern}_{f - \lambda_k \cdot \text{Id}}}_{= \text{Eig}_{\lambda_k}}.$$

Seien $(b_1, \dots, b_{\text{Geo}_{\lambda_1}})$, $(b_{\text{Geo}_{\lambda_1}+1}, \dots, b_{\text{Geo}_{\lambda_1} + \text{Geo}_{\lambda_2}})$, \dots ,
 $(b_{\text{Geo}_{\lambda_1} + \dots + \text{Geo}_{\lambda_{k-1}} + 1}, \dots, b_{\text{Geo}_{\lambda_1} + \dots + \text{Geo}_{\lambda_k}})$ Basen jeweils in
 $\text{Eig}_{\lambda_1}, \text{Eig}_{\lambda_2}, \dots, \text{Eig}_{\lambda_k}$. Dann ist die Menge $\{b_1, \dots, b_{\text{Geo}_{\lambda_1} + \dots + \text{Geo}_{\lambda_k}}\}$

(i) erzeugend, weil man nach Satz 11 jedes v als Summe
 $\underbrace{v_1}_{\in \text{Eig}_{\lambda_1}} + \dots + \underbrace{v_k}_{\in \text{Eig}_{\lambda_k}}$ darstellen kann, und jedes v_i eine Linearkombination
von Elementen der i -ten Basis sind.

(ii) **EINDEUTIG** erzeugend ist, weil man jedes v eindeutig als Summe
 $\underbrace{v_1}_{\in \text{Eig}_{\lambda_1}} + \dots + \underbrace{v_k}_{\in \text{Eig}_{\lambda_k}}$ schreiben kann, und jedes v_i kann man eindeutig als

Summe von Elementen von Basis von Eig_{λ_i} schreiben.

Dann ist das Tupel $(b_1, \dots, b_{\text{Geo}_{\lambda_1} + \dots + \text{Geo}_{\lambda_k}})$ eine Basis. Da alle Elemente Eigenvektoren sind, ist die Matrix von ϕ in der Basis nach Satz 54 LAAG 1 (Vorl. 19) diagonal, □

Verallgemeinere Eigenräume

Sei V ein endlich-dimensionaler \mathbb{K} -Vektorraum und $\phi : V \rightarrow V$ ein Endomorphismus. Das Minimalpolynom $Min_\phi = (x - \lambda_1)^{\gamma_1} \cdots (x - \lambda_k)^{\gamma_k}$ zerfalle in Linearfaktoren mit paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_k$. Dann gilt nach Satz 11:

$$V = \underbrace{Kern((\phi - \lambda_1 \cdot Id)^{\gamma_1})}_{:=W_{\lambda_1}} \oplus \cdots \oplus \underbrace{Kern((\phi - \lambda_k \cdot Id)^{\gamma_k})}_{:=W_{\lambda_k}}.$$

Def. 7 Die Räume $W_{\lambda_1}, \dots, W_{\lambda_k}$ heißen **verallgemeinere Eigenräume**.

Satz 11 in Worten: Das Minimalpolynom von $\phi : V \rightarrow V$ zerfalle in Linearfaktoren. Dann ist V direkte Summe von verallgemeinerten Eigenräumen.

Def. 8 Ein Untervektorraum $W \subseteq V$ heißt ϕ -invariant, falls $Bild_\phi(W) \subseteq W$.

Triv. Bsp. $\{\vec{0}\}$ ist ein ϕ -invarianter Untervektorraum; V selbst ist ein ϕ -invarianter Untervektorraum.

Interes. Bsp. Verallgemeinere Eigenräume W_i (von ϕ) sind ϕ -invariante Untervektorräume.

Beweis. Für jedes $v \in W_\lambda$ ist

$$(\phi - \lambda \cdot Id)^\gamma \circ \phi(v) = \phi \circ (\phi - \lambda \cdot Id)(v) = \phi(\vec{0}) = \vec{0},$$

Allgemeine Überlegung

Das Minimalpolynom $Min_\phi = (x - \lambda_1)^{\gamma_1} \cdots (x - \lambda_k)^{\gamma_k}$ zerfalle in Linearfaktoren. Für jedes $i = 1, \dots, k$ sei B_i die Basis in W_{λ_i} . Wir setzen die Basen B_i zu einer Basis $B = (B_1, \dots, B_k)$ von V zusammen (Beweis, dass es eine Basis ist haben wir in Beweis von Folgerung aus Satz 11 gemacht).

Da W_{λ_i} ϕ -invariant ist, für jedes $b \in B_i$ ist das Vektor $\phi(b)$ eine Linearkombination des Vektoren aus der Basis B_i , ist die Matrix A von ϕ bzgl. der Basis blockdiagonal:

$$A := \begin{pmatrix} \boxed{A_1} & & & \\ & \boxed{A_2} & & \\ & & \ddots & \\ & & & \boxed{A_k} \end{pmatrix},$$

wobei A_i ist die Matrix von $\phi|_{W_{\lambda_i}}$ in der Basis B_i .

Wir haben also eine Basis konstruiert, so dass die Matrix von ϕ blockdiagonal ist. Wir werden die Basis noch verbessern.

Dazu werden wir $\phi|_{W_i} - \lambda_i \cdot Id$ untersuchen.

Nilpotente Endomorphismen

Def. 9 Ein Endomorphismus $\varphi : W \rightarrow W$ heißt **nilpotent**, falls es ein $n \in \mathbb{N}$ gibt mit $\varphi^n \equiv \mathbf{0}$.

Für $w \in W$ heißt $k \in \mathbb{N}$ die φ -Periode von w , falls $\varphi^{k-1}(w) \neq \mathbf{0}$, aber $\varphi^k(w) = \mathbf{0}$.

Bemerkung. Es ist stets $n \geq k$ (falls φ wie in Def. 9 ist).

Lemma 3. Sei $\varphi : W \rightarrow W$ nilpotent, sei $w \in W$ mit φ -Periode k . Dann sind $w, \varphi(w), \varphi^2(w), \dots, \varphi^{k-1}(w)$ linear unabhängig.

Beweis: Sei $a_0 w + a_1 \varphi(w) + \dots + a_{k-1} \varphi^{k-1}(w) = \mathbf{0}$. (*) Wir wenden φ^{k-1} an und bekommen

$$a_0 \varphi^{k-1}(w) + \underbrace{a_1 \varphi^k(w)}_{\vec{0}} + \dots + \underbrace{a_{k-1} \varphi^{2k-2}(w)}_{\vec{0}} = \mathbf{0}. \text{ Da } \varphi^{k-1}(w) \neq \mathbf{0}, \text{ folgt}$$

daraus dass $a_0 = 0$.

Dann (*) ist $a_1 \varphi(w) + \dots + a_{k-1} \varphi^{k-1}(w) = \mathbf{0}$.

Anwendung von φ^{k-2} liefert uns $a_1 = 0$ u.s.w., □

Zusammenfassung der Montag-Vorlesung:

Ziel: Eine Basis finden sodass die Matrix von gegebenen Endomorphismus $\phi : V \rightarrow V$ möglich einfach ist.

Annahme: Das Minimalpolynom Min_ϕ zerfällt in Linearfaktoren:

$$Min_\phi = (x - \lambda_1)^{\gamma_1} \cdots (x - \lambda_k)^{\gamma_k}.$$

Bemerkung. Falls $\mathbb{K} = \mathbb{C}$, ist diese Annahme immer erfüllt (Folgerung A aus dem Hauptsatz der Algebra).

Wir haben gezeigt: Es gibt eine Basis

$$\left(\underbrace{B_1}_{\text{Basis in } W_{\lambda_1}}, \dots, \underbrace{B_k}_{\text{Basis in } W_{\lambda_k}} \right)$$

sodass die Matrix von ϕ blockdiagonal ist:

$$A := \begin{pmatrix} \boxed{A_1} & & & & \\ & \boxed{A_2} & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \boxed{A_k} \end{pmatrix}.$$

Die Dimension des i -ten Blocks ist die Dimension von W_{λ_i} , d.h., Anzahl von Elementen in B_i .

Die Beschränkung $\phi|_{W_{\lambda_i}}$ von ϕ auf W_{λ_i} ist eine Abbildung mit der Eigenschaft

$$(\phi - \lambda_i Id)^{\gamma_i} = \mathbf{0}. \quad (*)$$

Wir werden jetzt eine Basis in W_{λ_i} suchen/finden, sodass die Matrix von $\phi|_{W_{\lambda_i}}$ (mit der Eigenschaft $(*)$) möglich einfach ist.

Also, für uns sind die Endomorphismen der Form $\phi + \lambda \cdot Id$ interessant, sodass $\phi^k = \mathbf{0}$. (Statt ϕ aus der letzten Folie betrachte ich jetzt $\phi := \phi_{alt} - \lambda \cdot Id$). Solche ϕ nennt man **nilpotent** (Def. 9). Für $w \in W$ heißt $k \in \mathbb{N}$ die **ϕ -Periode** von w , falls $\phi^{k-1}(w) \neq \vec{0}$, aber $\phi^k(w) = \vec{0}$.

Wiederholung – Lemma 3. Sei $\phi : W \rightarrow W$ nilpotent, sei $w \in W$ mit ϕ -Periode k . Dann sind $w, \phi(w), \phi^2(w), \dots, \phi^{k-1}(w)$ linear unabhängig.

Bezeichnung. $Z_w := \text{span}(w, \phi(w), \phi^2(w), \dots, \phi^{k-1}(w))$ ist ein invarianter Untervektorraum: in der Tat,

$\phi(a_0 w + a_1 \phi(w) + \dots + a_{k-1} \phi^{k-1}(w)) = a_0 \phi(w) + \dots + a_{k-2} \phi^{k-1}(w)$ ist wieder ein Element von $\text{span}(w, \phi(w), \phi^2(w), \dots, \phi^{k-1}(w))$. Da

die erzeugende Elemente

$$b_1 := \phi^{k-1}(w), b_2 := \phi^{k-2}(w), \dots, b_k := w$$

nach Lemma 3 linear unabhängig sind, bilden sie eine Basis. Da

$\phi(b_i) = b_{i-1}$ ist (für $i \neq 1$) und $\phi(b_1) = \vec{0}$, ist die Matrix von ϕ in

der Basis
$$\begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}$$

Lemma 4 (Zerlegungslemma für nilpotente Endomorphismen) Sei W ein endlich-dimensionaler \mathbb{K} -Vektorraum und $\phi : W \rightarrow W$ ein nilpotenter Endomorphismus. Sei $w \in W$ ein Element mit maximaler ϕ -Periode k . Dann existiert ein ϕ -invarianter Untervektorraum $U \subseteq W$, so dass $W = Z_w \oplus U$.

Beweis. Wähle einen ϕ -invarianten Untervektorraum $U \subseteq W$ mit maximaler Dimension, so dass $Z_w \cap U = \{\vec{0}\}$. Dann ist die Summe $Z_w + U \subseteq W$ eine direkte Summe: ist $z + u = z' + u'$, so ist $z - z' = u' - u \in Z_w \cap U$, d.h., $z = z', u = u'$.

Z.z.: $Z_w \oplus U = W$. Angenommen das wäre nicht der Fall: Dann gäbe es ein $v \in W$ mit $v \notin Z(w) \oplus U$. Wähle j so, dass $\phi^{j-1}(v) \notin Z_w \oplus U$, aber $\phi^j(v) \in Z(w) \oplus U$: Ein solches j existiert, da ϕ nilpotent ist, und deswegen für genügend grosse j gilt $\phi^j(v) = \vec{0}$. Setze $x := \phi^{j-1}(v)$. Dann gilt $x \notin Z_w \oplus U$, aber $\phi(x) \in Z_w \oplus U$.

Schreibe $\phi(x) = \underbrace{a_0 w + \dots + a_{k-1} \phi^{k-1}(w)}_{\in Z_w} + \underbrace{u}_{\in U}$.

$$\begin{aligned} \vec{0} &= \phi^k(x) = \phi^{k-1} \circ \phi(x) \implies \\ &= \phi^{k-1}(a_0 w + \dots + a_{k-1} \phi^{k-1}(w) + u) \quad \text{nur } \phi^{k-1}(w), \phi^{k-1}(u) \stackrel{\text{dürfen}}{\neq} 0 \\ &= \underbrace{a_0 \phi^{k-1}(w)}_{\in Z_w} + \underbrace{\phi^{k-1}(u)}_{\in U} \quad \text{Weil die Summe direkt ist} \implies a_0 = 0, \phi^{k-1}(u) = 0 \end{aligned}$$

Setze $y := x - (a_1 w + \dots + a_{k-1} \phi^{k-2}(w))$.

Nach Konstruktion gilt:

$$\phi(y) = \phi(x) - (a_1 \phi(w) + \dots + a_{k-1} \phi^{k-1}(w)) = u \in U,$$

Wir zeigen: $y \notin Z_w \oplus U$. Sonsts wäre auch

$$x = y + a_1 w + \dots + a_{k-1} \phi^{k-2}(w) \in Z_w \oplus U.$$

Setze $U' := U \oplus \text{span}(y)$. (Die Summe ist direkt, weil $y \notin U$).

$\dim(U') = \dim(U) + 1$; U' ist ϕ -invariant. Wir bekommen Widerspruch mit der Annahme, dass U maximale Dimension hat. □

Lemma 4 auf der Sprache von Matrizen. Sei A eine $(n \times n)$ - Matrix über \mathbb{K} mit $A^\gamma = 0$. Dann existiert eine Matrix $B \in GL(n, \mathbb{K})$ sodass

$$B^{-1}AB = \left(\begin{array}{c|c} \begin{array}{cccc} 0 & 1 & & \\ & 0 & 1 & \\ & & \ddots & \ddots \\ & & & 0 & 1 \\ & & & & 0 \end{array} & \begin{array}{c} \\ \\ \\ \boxed{D} \end{array} \end{array} \right).$$

Bemerkung. Da $A^\gamma = \mathbf{0}$, ist auch $D^\gamma = \mathbf{0}$. Dann kann man D auch (mit einem geeigneten Basiswechsel) in zwei Blöcke zersplittern u.s.w.

Satz 12 (Zerlegungssatz für nilpotente Endomorphismen) Sei W ein endlich-dimensionaler \mathbb{K} -Vektorraum, sei $\phi : W \rightarrow W$ ein nilpotenter Endomorphismus. Dann existieren $w_1, \dots, w_l \in W$, so dass $W = Z_{w_1} \oplus \dots \oplus Z_{w_l}$ und die Dimensionen der Z_{w_j} sind bis auf die Reihenfolge eindeutig durch ϕ festgelegt.

Beweis: Die Existenz (der Zerlegung) folgt induktiv aus dem Zerlegungslemma.

Also, es gibt eine Basis, sodass die Matrix von ϕ gleich

$$\begin{pmatrix} \boxed{C_1} & & & & \\ & \boxed{C_2} & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \boxed{C_l} \end{pmatrix}, \text{ wobei } C_i = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}. \text{ Verschiedene } C_i$$

können selbstverständlich verschiedene Dimensionen haben.

Wir müssen jetzt zeigen, dass der Endomorphismus die Anzahl und Dimensionen von Blöcke bestimmt.

Sie werden es zu Hause ausrechnen (Hausaufgabe 3):Für

$$C = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \text{ gilt: } rk(C) = n - 1 \text{ (in unserem Fall } 5 - 1).$$

$$\text{Für } C^2 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \text{ gilt: } rk(C^2) = n - 2 \text{ (in unserem fall 3).}$$

$$\text{Für } C^3 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \text{ gilt: } rk(C^3) = n - 3 \text{ (in unserem Fall 2).}$$

$$C^5 = \begin{bmatrix} 0 & 0 & \vdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \text{ gilt: } rk(C^5) = 0. \quad C^6 = \mathbf{0}; \\ rk(C^6) = rk(C^5) = 0.$$

Das ist immer der Fall:

$\dim(\text{Kern}_{\phi|Z_{w_j}}) = 1$. (Weil

$\phi(a_0 w_j + \dots + a_{k-1} \phi^{k-1}(w_j)) = a_0 \phi(w_j) + \dots + a_{k-2} \phi^{k-1}(w_j)$ ist genau dann $\vec{0}$, wenn alle $a_0, \dots, a_{k-2} = 0$.)

Bezeichnet n_j die ϕ -Periode von w_j , dann gilt für $m \in \mathbb{N}$:

$$\text{Kern}_{(\phi|Z_{w_j})^m} = \begin{cases} \text{span}(\phi^{n_j-m}(w_j), \dots, \phi^{n_j-1}(w_j)), & \text{falls } n_j > m \\ Z_{w_j}, & \text{falls } m \geq n_j \end{cases}$$

Wir werden diese Beobachtung nutzen um Anzahl und Dimensionen aller Blöcke zu bestimmen; nach dem Beispiel wird es hoffentlich offensichtlich.

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

wir sehen: $rk(A) =$
 $\underbrace{dim(Bild_\phi)} = n -$

von ϕ bestimmt

{Anzahl von Kästchen}.

Dann ist Anzahl von Kästchen ist durch ϕ eindeutig bestimmt.

$$A^2 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

wir sehen: $rk(A^2) =$
 $\underbrace{dim(Bild_{\phi^2})} = n -$

von ϕ bestimmt

{Anzahl von Kästchen} -

{Anzahl von Kästchen von Dimension ≥ 2 }

. Dann ist Anzahl von Kästchen von Dimension ≥ 2 durch ϕ eindeutig bestimmt.

$$A^3 = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

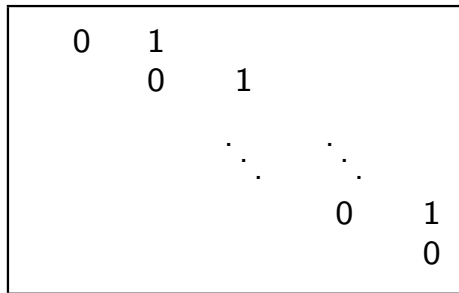
wir sehen: $rk(A^3) = \underbrace{dim(\text{Bild}_{\phi^3})}_{\text{von } \phi \text{ bestimmt}} = n -$

$\{ \text{Anzahl von Kästchen} \} -$
 $\{ \text{Anzahl von Kästchen}$
 $\text{von Dimension } \geq 2 \}$
 $- \{ \text{Anzahl von Kästchen}$
 $\text{von Dimension } \geq 3 \}$

.
 Dann ist Anzahl von
 Kästchen von Dimen-
 sion ≥ 3 durch ! ein-
 deutig bestimmt.

Analog gilt für eine beliebige nilpotente Matrix.

Satz 12 auf der Matrizen-Sprache: Sei A eine $(n \times n)$ - Matrix über \mathbb{K} mit $A^\gamma = \mathbf{0}$. Dann gibt es eine $B \in GL(n, \mathbb{K})$, sodass die Matrix $B^{-1}AB$

gleich $\begin{pmatrix} \boxed{C_1} & & & & \\ & \boxed{C_2} & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \boxed{C_l} \end{pmatrix}$, wobei $C_i =$ 

Bemerkung.

$$B^{-1}(A - \lambda \cdot Id)B \stackrel{\text{Linearität}}{=} B^{-1}AB - \lambda B^{-1}IdB = B^{-1}AB - \lambda \cdot Id.$$

Deswegen ist (falls $(A - \lambda \cdot Id)^\gamma = \mathbf{0}$ wie in Satz 11) $B^{-1}AB =$

$$\begin{pmatrix} \boxed{C_1} & & & & \\ & \boxed{C_2} & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \boxed{C_l} \end{pmatrix} + \lambda \cdot Id.$$

Jordansche Normalform

Def. 10 Die $k \times k$ Matrix $J_i^k := \begin{pmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \ddots & \ddots \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}$ heißt *Jordan-Block*.

Bsp. $J_2^1 = (2)$, $J_1^2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $J_{1+i}^3 = \begin{pmatrix} 1+i & 1 & \\ & 1+i & 1 \\ & & 1+i \end{pmatrix}$

Bemerkung $\text{alg}_{J_\lambda^k}(\lambda) = k$ (da $\chi_{J_\lambda^k} = (\lambda - t)^k$), $\text{geo}_{J_\lambda^k} = 1$ (da

$J_i^k - \lambda \cdot Id_k = \begin{pmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & \ddots & \ddots \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}$ ist; deren Rang ist $k - 1$ und deswegen

$\dim(\text{Kern}_{J_\lambda^k - \lambda \cdot Id_k}) = k - (k - 1) = 1$. Also, Jordan-Block J_i^k ist für $k \geq 2$ nicht diagonalisierbar.)

Die Matrix der Form $\begin{pmatrix} \boxed{J_{\lambda_1}^{k_1}} & & \\ & \ddots & \\ & & \boxed{J_{\lambda_m}^{k_m}} \end{pmatrix}$, wobei $J_{\lambda_j}^{k_j}$ Jordan-Blöcke sind,

heißt **Jordan-Matrix**.

Bsp. Die Matrizen

$\begin{pmatrix} 2 & & \\ & 3 & \\ & & 4 \end{pmatrix}$, $\begin{pmatrix} 2 & 1 & \\ & 2 & 3 \end{pmatrix}$, $\begin{pmatrix} 2 & 1 & \\ & 2 & 2 \end{pmatrix}$, $\begin{pmatrix} 2 & 1 & \\ & 2 & 1 \\ & & 2 \end{pmatrix}$ sind Jordan-Matrizen.

Satz 13 (Jordansche Normalform) Sei \mathbb{K} ein Körper, sei V ein endlichdimensionaler \mathbb{K} -Vektorraum und sei $\phi : V \rightarrow V$ ein Endomorphismus, dessen Minimalpolynom in Linearfaktoren zerfällt. Dann existiert eine Basis von V , so dass die Matrix von ϕ Jordan-Matrix sind. Diese Jordansche Normalform ist bis auf Reihenfolge der Jordanblöcke eindeutig.

Satz 13 ist eine einfache Folgerung von Sätze 11, 12

Satz 11 – Wiederholung Es gibt eine Basis, so dass in der Basis die

Matrix von ϕ die Form $\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{pmatrix}$ hat, wobei

$$(A_i - \lambda_i \cdot Id)^{\gamma_i} = \mathbf{0}.$$

Aus Satz 12, siehe auch Bemerkung oben: Für jedes A_i (der Dimension n_i) mit der Eigenschaft $(A_i - \lambda_i \cdot Id)^{\gamma_i} = 0$ gibt es eine

$$B_i \in GL(n_i, \mathbb{K}) \text{ mit } B_i^{-1} A_i B_i = \begin{pmatrix} c_1^i & & & \\ & c_2^i & & \\ & & \ddots & \\ & & & c_{l_i}^i \end{pmatrix} + \lambda_i Id .$$

Dann ist

$$\begin{pmatrix} B_1 & & & \\ & B_2 & & \\ & & \ddots & \\ & & & B_k \end{pmatrix}^{-1} \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{pmatrix} \begin{pmatrix} B_1 & & & \\ & B_2 & & \\ & & \ddots & \\ & & & B_k \end{pmatrix} \\
 = \begin{pmatrix} B_1^{-1} A_1 B_1 & & & \\ & B_2^{-1} A_2 B_2 & & \\ & & \ddots & \\ & & & B_k^{-1} A_k B_k \end{pmatrix} = \text{Wie wir wollen.}$$

Eindeutigkeit der Jordan-Normalform folgt aus der Beobachtung, dass ϕ die Untervektorräume W_{λ_i} eindeutig bestimmt, und aus der Eindeutigkeitsaussage im Satz 12. □

Anwendung: Beweis von Hamilton-Cayley für Matrizen über \mathbb{C} .

Satz 56 LAAG I Vorl. 20 (Hamilton-Cayley)



1805 --1856



1821 --1895

$$\mathcal{N}_A(A) = \mathbf{0}$$

In Worten: Charakteristisches Polynom einer Matrix annihiliert die Matrix.

Beobachtung 1 Für $J = \begin{pmatrix} \boxed{J_{\lambda_1}^{k_1}} & & \\ & \ddots & \\ & & \boxed{J_{\lambda_m}^{k_m}} \end{pmatrix}$, ist

$$\mathcal{N}_J = (\lambda_1 - t)^{k_1} \dots (\lambda_m - t)^{k_m}.$$

Tatsächlich, da J eine obere Dreiecksmatrix ist (=alle Einträge unterhalb der Hauptdiagonalen 0 sind) ist, ist $J - t \cdot Id$ auch eine obere Dreiecksmatrix, und deswegen $\det(J - t \cdot Id)$ ist Produkt von Diagonalelementen, die $(\lambda_i - t)$ sind.

Beobachtung 2 Die Matrizen A, A' seien ähnlich: $A = B^{-1}A'B$.

Dann gilt: Für jedes $P \in \mathbb{C}[t]$ ist $P(A) = B^{-1}P(A')B$.

Tatsächlich, wir haben dies im wesentlichen in Vorl. 20 LAAG I bewiesen:

$$A^k = \underbrace{A \cdot A \cdot \dots \cdot A}_{k \text{ mal}} = B^{-1}A' \underbrace{BB^{-1}}_{Id} A' B \dots B^{-1}A' B = B^{-1}A'^k B.$$

$$\text{Also, } P(A) = P(B^{-1}A'B) =$$

$$a_k B^{-1}A'^k B + a_{k-1} B^{-1}A'^{k-1} B + \dots + a_0 B^{-1} B$$

$$\stackrel{\text{Linearität}}{=} B^{-1}(a_k A'^k + \dots + a_0 Id) B = B^{-1}P(A')B.$$

Beobachtung 3 Für eine Block-diagonale Matrix

$$M = \begin{pmatrix} \boxed{B_1} & & \\ & \ddots & \\ & & \boxed{B_m} \end{pmatrix}, \text{ wobei } B_i \text{ eine } k_i \times k_i\text{-Matrix ist, ist}$$

$$P(M) = \begin{pmatrix} \boxed{P(B_1)} & & \\ & \ddots & \\ & & \boxed{P(B_m)} \end{pmatrix}.$$

Beobachtung 4 $\mathfrak{N}_{J_\lambda^k}(J_\lambda^k) = \mathbf{0}$.

Tatsächlich,

$$\mathfrak{N}_{J_\lambda^k}(J_\lambda^k) = (\lambda \cdot Id - J_\lambda^k)^k = (-1)^k \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}^k \quad \text{wie in Bsp. vorher} \mathbf{0}$$

Alle vier Beobachtungen zusammen:

Für A , die ähnlich zu $J = \begin{pmatrix} \boxed{J_{\lambda_1}^{k_1}} & & \\ & \ddots & \\ & & \boxed{J_{\lambda_m}^{k_m}} \end{pmatrix}$, gilt

$$1. \ \mathfrak{N}_J = (\lambda_1 - t)^{k_1} \dots (\lambda_m - t)^{k_m} = \mathfrak{N}_{J_{\lambda_1}^{k_1}} \dots \mathfrak{N}_{J_{\lambda_m}^{k_m}}.$$

2. $P(A) \stackrel{\text{für geeignetes } B}{=} B^{-1}P(J)B$, wobei J die Jordan'sche Normalform von A

$$3. \ P(J) = \begin{pmatrix} \boxed{P(J_{\lambda_1}^{k_1})} & & \\ & \ddots & \\ & & \boxed{P(J_{\lambda_m}^{k_m})} \end{pmatrix}, \quad 4. \ \mathfrak{N}_J(J_{\lambda_i}^{k_i}) = P_i(J_{\lambda_i}^{k_i}) \cdot \underbrace{\mathfrak{N}_{J_{\lambda_i}^{k_i}}(J_{\lambda_i}^{k_i})}_0 = \mathbf{0}$$

Dann ist $\mathfrak{N}_A(A) \stackrel{\text{Satz 18 LAAG I}}{=} \mathfrak{N}_J(A) = B^{-1}\mathfrak{N}_J(J)B =$

$$B^{-1} \begin{pmatrix} \boxed{\mathfrak{N}_J(J_{\lambda_1}^{k_1})} & & \\ & \ddots & \\ & & \boxed{\mathfrak{N}_J(J_{\lambda_m}^{k_m})} \end{pmatrix} B = B^{-1} \mathbf{0} B = \mathbf{0},$$

□

Satz 13 (Jordansche Normalform) Sei \mathbb{K} ein Körper, sei V ein endlichdimensionaler \mathbb{K} -Vektorraum und sei $\phi : V \rightarrow V$ ein Endomorphismus, dessen Minimalpolynom in Linearfaktoren zerfällt. Dann existiert eine Basis von V , so dass die Matrix von ϕ Jordan-Matrix sind. Diese Jordansche Normalform ist bis auf Reihenfolge der Jordanblöcke eindeutig.

Wiederholung: *Jordan-Matrizen* sind die Matrices der Form

$$\begin{pmatrix} \boxed{J_{\lambda_1}^{k_1}} & & & \\ & \ddots & & \\ & & \boxed{J_{\lambda_m}^{k_m}} & \\ & & & \ddots \end{pmatrix}, \text{ wobei } J_{\lambda}^k \text{ die folgende } k \times k\text{-Matrix ist}$$
$$\begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}$$

- ▶ **Beobachtung 2 aus der Vorlesung 6** Sei $P \in \mathbb{K}[x]$ ein Polynom, $A \in \text{Mat}(n, n, \mathbb{K})$, $B \in \text{GL}(n, \mathbb{K})$. Dann gilt:
 $P(B^{-1}AB) = B^{-1}P(A)B$.

- ▶ **Beobachtung 3 aus der Vorlesung 6** Für eine Block-diagonale

$$\text{Matrix } M = \begin{pmatrix} \boxed{A_1} & & \\ & \ddots & \\ & & \boxed{A_m} \end{pmatrix}, \text{ wobei } A_i \text{ eine } k_i \times k_i\text{-Matrix ist,}$$
$$\text{ist } P(M) = \begin{pmatrix} \boxed{P(A_1)} & & \\ & \ddots & \\ & & \boxed{P(A_m)} \end{pmatrix}.$$

- Sei J ein $k \times k$ Jordan Block über \mathbb{C} mit Eigenwert λ , d.h., $J = J_\lambda^k$.
Dann gilt:

$$P(J) = \begin{pmatrix} P(\lambda) & \frac{1}{1!}P'(\lambda) & \frac{1}{2!}P^{(2)}(\lambda) & \dots & \frac{1}{n!}P^{(n)}(\lambda) \\ & \ddots & \ddots & & \vdots \\ & & & & \frac{1}{2!}P^{(2)}(\lambda) \\ & & & P(\lambda) & \frac{1}{1!}P'(\lambda) \\ & & & & P(\lambda) \end{pmatrix}, \text{ wobei } P^{(j)} \text{ die}$$

j -te Ableitung des Polynoms P ist.

(Z.B., für Polynom $P = z^3$ ist $P' = 3z^2$, $P^{(2)} = 6z$, $P^{(3)} = 6$,
 $P^{(4)} = P^{(5)} = \dots = 0$;

allgemein: für Polynom $P = \sum_{k=0}^m a_k x^k \in \mathbb{C}[x]$ ist

$$P' = \sum_{k=1}^m k \cdot a_k x^{k-1} \in \mathbb{C}[x] \quad (*)$$

Das ist in wesentlichem Hausaufgabe 3a: in der Tat, da die Ableitung die lineare Abbildung von $\mathbb{C}[x]$ auf $\mathbb{C}[x]$ ist, genügt es, die Formel für Monomen (d.h., für Polynomen der Form x^n) zu prüfen, und Sie haben sie zu Hause ausgerechnet.

Bemerkung. Man kann die Ableitung über beliebige Körper definieren, nicht nur über \mathbb{C} , mit Hilfe der Formel (*). Wir machen es später. In dem Fall die Eigenschaften, die für \mathbb{C} „umsonst“ aus Ana I kommen, separat bewiesen werden müssen.

Analytische Funktionen von Matrizen

Sei $a_0, a_1, \dots, a_k, \dots \in \mathbb{C}$ eine Folge,

$f(z) = a_0 + a_1 z + a_2 z^2 + \dots + a_k z^k + \dots = \sum_{k=0}^{\infty} a_k z^k$ die entsprechende analytische Funktion und A eine Matrix. Wir betrachten die Folge von

Matrizen (partielle Summen): $a_0 \cdot Id, a_0 \cdot Id + a_1 A, \dots,$

$a_0 \cdot Id + a_1 \cdot A + a_2 \cdot A^2 + \dots + a_m A^m = \sum_{k=0}^m a_k \cdot A^k := P_m(A), \dots$

Wir sagen, dass die Folge konvergiert, falls sie komponentenweise konvergiert, d.h., für jede $i, j \in n$ konvergieren die (i, j) -Einträge von Matrizen $a_0 \cdot Id, a_0 \cdot Id + a_1 A, \dots, P_m(A)$.

Def. 10 Falls die Folge konvergiert, setzen wir

$$f(A) = \lim_{m \rightarrow \infty} \sum_{k=0}^m a_k \cdot A^k.$$

(Falls für irgendwelche i, j die Folge divergiert, ist $f(A)$ nicht definiert.)

Bemerkung: Für Polynomen $\sum_{k=0}^m a_k \cdot A^k$ gilt:

$\sum_{k=0}^m a_k \cdot (B^{-1}AB)^k = B^{-1} \left(\sum_{k=0}^m a_k \cdot A^k \right) B$. Wir werden bald sehen:

$f(B^{-1}AB) = B^{-1}f(A)B$.

Dies bedeutet, dass Definition nicht von Wahl von Basis abhängt (d.h., wir in Wirklichkeit analytische Funktionen von Endomorphismen definiert haben).

Natürliche Fragen

Unten $f = \sum_{k=0}^{\infty} a_k z^k$ ist stets eine analytische Funktion, A eine $n \times n$ -Matrix.

1. Für welche A existiert $f(A)$? (d.h., für welche A konvergiert die Reihe $a_0 \cdot Id, a_0 \cdot Id + a_1 A, \dots$, komponentenweise)
2. Wie kann man $f(A)$ ausrechnen?

Bemerkung. Es ist hoffnungslos, bereits für 3×3 -Matrizen einfache analytische Funktionen von Matrizen, z.B. e^A , nach Definition auszurechnen.

3. Wozu braucht man analytische Funktionen von Matrizen auszurechnen?

Bemerkung. In dem Kurs gew. Differentialgleichungen werden analytische Funktionen von linearen Abbildungen auch gemacht, (sogar für Banach-Räumen) teilweise parallel zu linearen Algebra. Das bedeutet, dass sie für Differentialgleichungen, und deswegen auch alle Naturwissenschaften wichtig sind.

Satz 14 Angenommen, alle Eigenwerte von $A \in \text{Mat}(n, n, \mathbb{C})$ liegen in Konvergenzkreise von f . Dann gilt: $f(A)$ ist definiert.

Satz 14 Antwortet auf die erste Frage: wann ist $f(A)$ definiert.

Satz 15 (Spektralsatz) Seien $\lambda_1, \dots, \lambda_m$ verschiedene Eigenwerte von $A \in \text{Mat}(n, n, \mathbb{C})$, k_1, \dots, k_m ihre algebraische Vielfachheiten ($k_1 + \dots + k_m = n$). Angenommen, der Konvergenzradius von f ist größer als $\max|\lambda_i|$.

Dann gilt: Es gibt n Matrizen

$Z_{\lambda_1,1}, \dots, Z_{\lambda_1,k_1}, Z_{\lambda_2,1}, \dots, Z_{\lambda_2,k_2}, \dots, Z_{\lambda_m,1}, \dots, Z_{\lambda_m,k_m} \in \text{Mat}(n, n, \mathbb{C})$, die nur von A abhängen, so dass für jede analytische Funktion f gilt: ist

$f(A)$ definiert, so ist

$$f(A) = \sum_{i=1}^m (f(\lambda_i)Z_{\lambda_i,1} + \dots + f^{(k_i-1)}(\lambda_i)Z_{\lambda_i,k_i}) \quad (*)$$

In Worten. Für jede Matrix A gibt es n Matrizen $Z_{\lambda_i,j}$ sodass jede analytische Funktion $f(A)$ mit genügend großem Konvergenzradius die Linearkombination (*) ist.

Bemerkung. Die Matrizen $Z_{\lambda_i,j}$ hängen nicht von f ab, nur von A . Satz 15 antworten auf die zweite Frage und gibt uns eine effektive Methode, die Matrix $f(A)$ auszurechnen.

Beweis von Sätze 14/15.

Beweisstrategie: wir zeigen zuerst, dass es genügt, die Sätze nur für Jordan-Blöcke zu beweisen, und dann prüfen die Formeln für die Blöcke.

Wir benutzen die folgende Aussage aus Ana I:

Konvergieren die Folgen $\alpha_0, \dots, \alpha_k, \dots$; $\beta_0, \dots, \beta_k, \dots$ gegen α bzw. β , so konvergiert die Folge $C_1 \cdot \alpha_i + C_2 \cdot \beta_i$ gegen $C_1 \cdot \alpha + C_2 \cdot \beta$.

Ist $f(B^{-1}AB)$ definiert, dann ist $f(A)$ auch definiert. In der Tat, für jedes Polynom $P_k = a_0 + \dots + a_k z^k$ ist $P_k(A) = BP_k(B^{-1}AB)B^{-1}$. Also, die Einträge von $P_k(A)$ sind lineare Ausdrücke von Einträgen von $P_k(B^{-1}AB)$ mit konstanten Koeffizienten, die von B kommen.

Falls alle Einträge von $P_k(B^{-1}AB)$ für $k \rightarrow \infty$ konvergieren, konvergieren auch deswegen die Einträge von

$P_k(A) = BP_k(B^{-1}AB)B^{-1}$, und zwar gegen $Bf(B^{-1}AB)B^{-1}$.

Deswegen genügt es, die Sätze 14, 15 für die Jordan-Matrizen zu beweisen.

Da für eine Block-diagonale Matrix $M = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_m \end{pmatrix}$, wobei

A_j eine $k_j \times k_j$ -Matrix ist, ist $P_n(M) = \begin{pmatrix} P_n(A_1) & & \\ & \ddots & \\ & & P_n(A_m) \end{pmatrix}$,

genügend es, die Sätze 14, 15 für die einzelne Jordan-Blöcke zu beweisen.

Beweis für Jordan-Blöcke: ausrechnen

Nach Hausaufgabe 3, ist

$$P_n(J_\lambda^k) = \begin{pmatrix} P_n(\lambda) & \frac{1}{1!} P_n'(\lambda) & \frac{1}{2!} P_n^{(2)}(\lambda) & \cdots & \frac{1}{k!} P_n^{(k)}(\lambda) \\ & \ddots & \ddots & & \vdots \\ & & \ddots & & \vdots \\ & & & P_n(\lambda) & \frac{1}{2!} P_n^{(2)}(\lambda) \\ & & & & \frac{1}{1!} P_n'(\lambda) \\ & & & & P_n(\lambda) \end{pmatrix}$$

Falls λ in der Konvergenzkreise von f liegt, konvergiert $P_n(\lambda)$ gegen $f(\lambda)$, folglich konvergieren die Diagonalelemente gegen $f(\lambda)$. Falls λ in der Konvergenzkreise von f liegt, konvergiert $P_n'(\lambda)$ gegen $f'(\lambda)$, folglich konvergieren die $(i, i + 1)$ Einträge gegen $\frac{1}{1!} f'(\lambda)$, u.s.w. Satz 14 ist bewiesen.

Wir sehen, dass $f(J)$ Linearkombination von Matrizen

$$D_i := \begin{pmatrix} 0 & \cdots & 1 & \cdots & 0 \\ & \ddots & & \ddots & \vdots \\ & & 0 & \ddots & \vdots \\ & & & \ddots & \vdots \\ & & & & 1 \\ & & & & \vdots \\ & & & & 0 \end{pmatrix} \text{ (die nicht von } f \text{ abhängen), mit Koeffizienten}$$

$\frac{1}{i!} f^{(i)}(\lambda)$ ist. Satz 15 ist bewiesen.

Folgerung (Rechenregeln für analytische Funktionen von Matrizen) Es gilt:

1. $Af(A) = f(A)A$ (falls $f(A)$ definiert ist).
2. Sei $h = f + g$. Dann gilt $h(A) = f(A) + g(A)$ (falls $h(A), f(A)$ definiert sind.)
3. Sei $h = fg$. Dann gilt $h(A) = f(A)g(A)$ (falls definiert)
4. Sei $h(z) = f(g(z))$. Dann gilt $h(A) = f(g(A))$ (falls definiert)
5. Sei $f(\lambda) \neq 0$ für jeden Eigenwert λ von A , $h = 1/f$. Dann $h(A) = (f(A))^{-1}$.

Beweis. Benutzen Sie Satz 15 oder die Definition. Z.B. in (1) sei $g(z) = zf(z) = f(z)z$. Dann ist $g(A) = Af(A) = f(A)A$.

Anwendung: die Exponentialabbildung von Matrizen

Die gewöhnliche Exponentialfunktion e^x in der Analysis ist (unabhängig voneinander) durch zwei Eigenschaften charakterisiert:

1. Als Lösung der Differentialgleichung $y' = y$ mit der Anfangsbedingung $y(0) = 1$,

2. Als (überall konvergente) Potenzreihe $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$

Wir betrachten jetzt e^A , wobei A eine Matrix ist, und zeigen, dass Sie die beiden Eigenschaften hat.

Nach Def. 10 ist $e^A := \sum_{k=0}^{\infty} \frac{1}{k!} A^k$

e^J , wobei $J = J_\lambda^k$ ein Jordan-Block ist

Im Beweis von Sätze 14, 15 haben wir gezeigt, dass

$$f(J) = \begin{pmatrix} f(\lambda) & \frac{1}{1!} f'(\lambda) & \frac{1}{2!} f^{(2)}(\lambda) & \cdots & \frac{1}{n!} f^{(n)}(\lambda) \\ & \ddots & \ddots & & \vdots \\ & & \ddots & & \vdots \\ & & & f(\lambda) & \frac{1}{2!} f^{(2)}(\lambda) \\ & & & & \frac{1}{1!} f'(\lambda) \\ & & & & f(\lambda) \end{pmatrix}$$

Da für $f(z) = e^z$ gilt $f^{(j)}(z) = e^z$, ist $f^{(j)}(\lambda) = e^\lambda$, und deswegen

$$e^J = e^\lambda \begin{pmatrix} 1 & \frac{1}{1!} & \frac{1}{2!} & \cdots & \frac{1}{n!} \\ & \ddots & \ddots & & \vdots \\ & & \ddots & & \vdots \\ & & & 1 & \frac{1}{1!} \\ & & & & 1 \end{pmatrix}$$

Für eine Jordan-Matrix $M = \begin{pmatrix} \boxed{J_1} & & \\ & \ddots & \\ & & \boxed{J_m} \end{pmatrix}$, wobei $J_i := J_{\lambda_i}^{k_i}$

ein $k_i \times k_i$ -Jordanblock ist, ist $e^M = \begin{pmatrix} \boxed{e^{J_1}} & & \\ & \ddots & \\ & & \boxed{e^{J_m}} \end{pmatrix}$.

Insbesondere gilt: $\det(e^M) = e^{\text{trace}(M)}$ (die beide Seiten sind $e^{\sum_i k_i \lambda_i}$).

Satz 16 Die Exponentialabbildung hat die folgenden Eigenschaften:
(unten sind A, B stets $n \times n$ -Matrizen über \mathbb{C} .)

(1) e^A konvergiert für jede Matrix A .

(2) Die Zuordnung $A \mapsto e^A$ ist eine C^∞ -Abbildung (sogar reell analytisch) (als Abbildung von \mathbb{C}^{n^2} auf sich selbst.)
 $= \text{Mat}(n, n, \mathbb{C})$

(3) Es gilt $e^{A+B} = e^A \cdot e^B$, falls $AB = BA$ (dies ist das Exponentialgesetz)

(4) e^A ist stets invertierbar, und es gilt $e^A = e^{-A}$ sowie
 $\det(e^A) = e^{\text{trace}(A)}$

(5) Die Abbildung $t \mapsto e^{tA}$ ist bzgl. t differenzierbar mit $e^{0 \cdot A} = Id$ und
 $\frac{d}{dt} e^{t \cdot A} = A e^{tA}$.

Beweis (1): Da Konvergenzradius von e^z -Funktion unendlich ist, ist e^A für alle $A \in \text{Mat}(n, n, \mathbb{C})$ nach Satz 14 definiert.

Beweis (2): Weil die Einträge von e^A konvergente Reihen von Einträgen von A sind.

Beweis (3): Die Gleichung $AB = BA$ impliziert die binomische Formel $(A + B)^2 = A^2 + AB + BA + B^2 = A^2 + 2AB + B^2$, allgemein

$$(A + B)^n = \sum_{k=0}^n \binom{n}{k} A^k B^{n-k}, \text{ wobei } \binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Es folgt

$$\begin{aligned} e^{A+B} &= \sum_{n=0}^{\infty} \frac{1}{n!} (A + B)^n \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^n \binom{n}{k} A^k B^{n-k} \\ &= \sum_{n=0}^{\infty} \sum_{k=0}^n \frac{1}{k!(n-k)!} A^k B^{n-k} \\ &= \sum_{n=0}^{\infty} \sum_{k+l=n} \frac{1}{k!l!} A^k B^l \stackrel{\text{symmetrisch bzgl. } A, B}{=} e^{B+A}. \end{aligned}$$

Beweis (4): Setze $A := -B$. Offensichtlich, $AB = BA (= -B^2)$. Dann folgt nach Teil (3): $Id = e^0 = e^{A-A} = e^A \cdot e^{-A}$, also $e^{-A} = (e^A)^{-1}$.

Die Gleichung $\det(e^A) = e^{\text{trace}(A)}$ bekommt man sofort, wenn A eine Jordan-Matrix steht, wie wir vor dem Satz ausgerechnet haben. In dem Fall, auf der Diagonale von e^A stehen e^{λ_i} , wobei λ_i Eigenwerte sind. $e^{\text{trace}(A)}$ ist ebenfalls Produkt von Exponentialfunktionen von Eigenwerten (mit Vielfachheiten). Da \det und trace nicht von Wahl der Basis abhängen, gilt die Gleichung $\det(e^A) = e^{\text{trace}(A)}$ für alle A .

Beweis (5): Hier rechnet man mit gliedweiser Differentiation folgendes aus:

$$\begin{aligned} \frac{d}{dt}(e^{t \cdot A}) &= \sum_{k=0}^{\infty} \frac{1}{k!} t^k \cdot A^k = \sum_{k=0}^{\infty} \frac{d}{dt} \frac{1}{k!} t^k \cdot A^k \\ &= \sum_{k=1}^{\infty} \frac{d}{dt} \frac{1}{k!} \cdot k \cdot t^{k-1} \cdot A^k \\ &\stackrel{k-1=m}{=} A \sum_{m=0}^{\infty} \frac{d}{dt} \frac{1}{m!} \cdot t^m \cdot A^m. \end{aligned}$$

Anwendung: lösen von linearen Gleichungssysteme mit konstanten Koeffizienten

Für eine gegebene $n \times n$ -Matrix A gesucht wird ein Vektor $u = \begin{pmatrix} u_1(x) \\ \vdots \\ u_n(x) \end{pmatrix}$,

dessen Einträge u_i Funktionen von x sind, s.d.

$$(*) \quad u'(x) = Au$$

$$\text{d.h.,} \quad \begin{pmatrix} u'_1(x) \\ \vdots \\ u'_n(x) \end{pmatrix} = A \begin{pmatrix} u_1(x) \\ \vdots \\ u_n(x) \end{pmatrix}$$

Bsp. Für $n = 1$ ist $(*)$ die bekannte Gleichung $u'(x) = au(x)$, und deren Lösung ist $u(x) = e^{ax}$

wobei C eine Konstante ist.

$(*)$ ist eine äußerst wichtige Gleichung in Physik, und kommt selbstverständlich im Kurs gewöhnliche Differentialgleichungen vor.

Nach Satz 16 löst

$$u(x) := e^{Ax} C \quad (**)$$

diese Gleichung, wobei C ein (konstanter) Vektor ist. Nach Satz von Satz von Picard-Lindelöf (kommt noch im Kurs gewöhnliche Differentialgleichungen) haben alle Lösungen von $(*)$ die Form $(**)$.

Neues Thema: affine Geometrie

Definition 11 **Affiner Raum** über einem \mathbb{K} -Vektorraum V ist die Menge $\mathcal{A} \neq \emptyset$ mit einer Abbildung $+ : \mathcal{A} \times V \rightarrow \mathcal{A}$, für die gilt

(A₁) für alle $a \in \mathcal{A}, v, w \in V$ gilt

$$(a + v) + w = a + \underbrace{(v + w)}$$

Übliche Addition von Vektoren

(A₂) für alle $a_1, a_2 \in \mathcal{A}$ existiert genau ein $v \in V$ s.d. $a_2 = a_1 + v$ (Der Vektor v wird $\overrightarrow{a_1 a_2}$) bezeichnet.

Die **Dimension** des affinen Raums ist die Dimension des V . Die Elemente von \mathcal{A} heißen **Punkten**.

StandardBsp. Sei V ein \mathbb{K} -Vektorraum. Wir setzen $\mathcal{A} = V$. “+” sei die übliche Addition in V . Das ist ein affiner Raum:

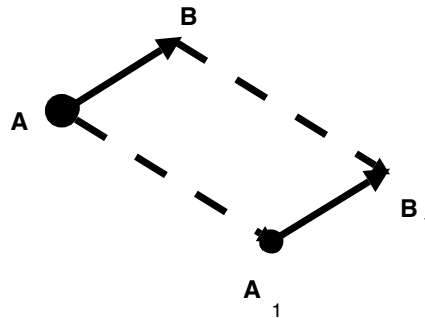
(A₁): $(a + v) + w = a + (v + w)$ entspricht Assoziativität der Addition,

(A₂) entspricht der Existenz der eindeutigen Inversen

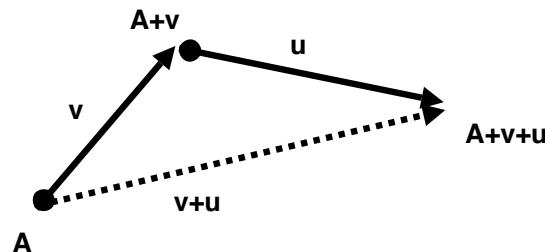
(weil $a_2 = a_1 + v \iff v \stackrel{\text{eindeutig}}{=} a_2 - a_1$.)

MotivationsBsp: \mathcal{E} = Übliche („Schulgeometrische“) Ebene, V – die Menge von geometrischen Vektoren (geordneten Strecken) mit Anfang in A_1 , “+” sei die Addition von Vektoren und Punkten auf der Ebene. Sei $A \in \mathcal{E}$, $\vec{v} = \overrightarrow{A_1 B_1}$ sei ein Vektor auf \mathcal{E} ,. Die **Summe** $A + \vec{v}$ ist ein Punkt $B \in \mathcal{E}$ so dass die geordnete Strecke $\overrightarrow{A, B}$ gleich (= parallel und hat die gleiche Länge und die gleiche Richtung) zu \vec{v} ist.

Addition von Vektoren und Punkten



Eigenschaft (A_1): A sei ein Punkt, \vec{v}, \vec{u} seien Vektoren. Es gilt:
 $(A + \vec{v}) + \vec{u} = A + (\vec{u} + \vec{v})$.



Eigenschaft (A_2): Für $A, B \in \mathcal{E}$ gibt es genau einen Vektor (mit Anfang in A_1), der zur geordneten Strecke $\overrightarrow{A, B}$ gleich ist.

Lass uns noch einmal Eigenschaften (A1), (A2) ansehen.

(A₁) für alle $a \in \mathcal{A}$, $v, w \in V$ gilt $(a + v) + w = a + (v + w)$

(A₂) für alle $a_1, a_2 \in \mathcal{A}$ existiert genau ein $v \in \mathcal{A}$ s.d. $a_2 = a_1 + v$ (Der Vektor v wird $\overrightarrow{a_1 a_2}$) bezeichnet.

Falls wir einen festen Punkt $a \in \mathcal{A}$ gewählt haben, ist \mathcal{A} “fast” ein Vektorraum: jedem a_1 ist eindeutiges $\overrightarrow{a a_1}$ zugeordnet.

Plan für Heute

Theorie von affinen Räumen

- ▶ Affiner Raum
- ▶ Affiner Unterraum
- ▶ Affine Abbildungen
- ▶ Affine Koordinaten
- ▶ Hauptsatz der affiner Geometrie

Theorie von Vektorräumen (LAAG I)

- ▶ Vektorraum (Vorl.9)
- ▶ Untervektorraum (Vorl. 9)
- ▶ Koordinaten (Vorl 11-12)
- ▶ Lineare Abbildungen (Vorl. 12-13)
- ▶ Hauptsatz der linearen Algebra (Vorl. 12)

Lemma 5 Für alle $a, a_1, a_2, a_3 \in \mathcal{A}$ gilt: $a + \vec{0} = a$, $\overrightarrow{aa} = \vec{0}$,
 $\overrightarrow{a_1 a_2} = -\overrightarrow{a_2 a_1}$, $\overrightarrow{a_1 a_2} + \overrightarrow{a_2 a_3} = \overrightarrow{a_1 a_3}$.

Beweis: Hausaufgabe.

Def. 12 Sei \mathcal{A} ein affiner Raum über \mathbb{K} - Vektorraum V . Eine Teilmenge $\mathcal{U} \subseteq \mathcal{A}$ heißt ein **affiner Unterraum**, falls ein Untervektorraum $V_{\mathcal{U}}$ und ein $a_0 \in \mathcal{A}$ existieren so dass

$$\mathcal{U} = \{a_0 + v, \text{ wobei } v \in V_{\mathcal{U}} \}.$$

$\dim(\mathcal{U}) := \dim(V_{\mathcal{U}})$. Die affine Unterräume der Dimension 1 heißen **Gerade**, die affine Unterräume der Dimension 2 heißen **Ebenen**. Falls V n -dimensional ist, die affine Unterräume der Dimension $n - 1$ heißen **Hyperebenen**.

Triviales Bsp. Jede 1-Punkt-Menge $\{a\} \subseteq \mathcal{A}$ ist ein affiner Unterraum (Weil $\{\vec{0}\}$ ein Vektorraum ist, und $a + \vec{0} \stackrel{\text{Lem. 5}}{=} a$) der Dimension 0 .

\mathcal{A} selbst ist auch ein affiner Unterraum, da nach **(A₂)**

$$\{a_0 + v, \text{ wobei } v \in V \} = \mathcal{A}. \quad \dim(\mathcal{A}) = \dim(V)$$

Bemerkung Affiner Unterraum \mathcal{U} ist ein affiner Raum über $V_{\mathcal{U}}$

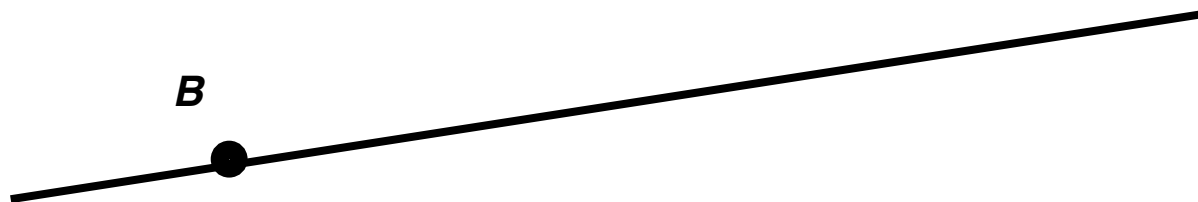
Bsp. Unterräume von \mathcal{E}_2

Unteräume von Ebene

A

Punkt ist ein 0-dim. aff. Unterraum

B



Gerade = $\{ B + t v, \text{ wobei } t \text{ aus } \mathbb{R} \text{ ist} \}$

ist ein 1-dim. aff. Unterraum

=Hyperebene

Die ganze Ebene ist ein 2-dim. aff. Unterraum

Bsp. Unterräume von E_3

Unteräume vom Raum

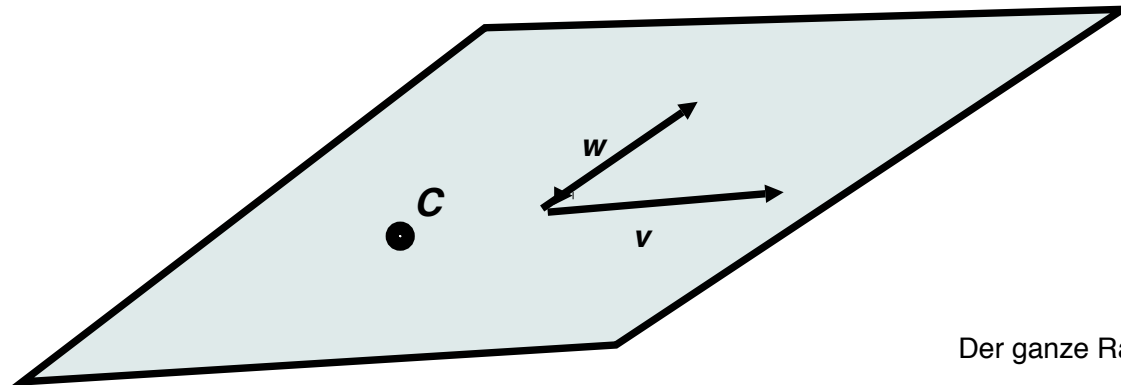
●
A

Punkt ist ein 0-dim. aff. Unterraum

●
B

Gerade = $\{ B + t v, \text{ wobei } t \text{ aus } \mathbb{R} \text{ ist} \}$

ist ein 1-dim. aff. Unterraum



Der ganze Raum

ist ein 3-dim. aff. Unterraum

Die Ebene $\{ C + t v + r w, \text{ wobei } t, r \text{ aus } \mathbb{R} \text{ sind} \}$

und v, w linear unabhängig sind

ist ein 2-dim. aff. Unterraum

ist eine Hyperebene

Lemma 7 Sind $\mathcal{U}_1, \mathcal{U}_2$ affine Unterräume s.d. $\mathcal{U}_1 \cap \mathcal{U}_2 \neq \emptyset$, so ist $\mathcal{U}_1 \cap \mathcal{U}_2$ auch ein Unterraum und $V_{\mathcal{U}_1 \cap \mathcal{U}_2} = V_{\mathcal{U}_1} \cap V_{\mathcal{U}_2}$.

Beweis: Wähle $a_0 \in \mathcal{U}_1 \cap \mathcal{U}_2$. Nach Lemma 6 gilt

$$\mathcal{U}_1 = \{a_0 + v, \text{ wobei } v \in V_{\mathcal{U}_1}\},$$

$$\mathcal{U}_2 = \{a_0 + v, \text{ wobei } v \in V_{\mathcal{U}_2}\}.$$

$$\text{Dann } \mathcal{U}_1 \cap \mathcal{U}_2 = \{a_0 + v \text{ wobei } v \in V_{\mathcal{U}_1} \cap V_{\mathcal{U}_2}\}.$$



Bsp. Betrachte den affine Raum aus StandardBsp, wobei $V = \mathbb{K}^n$. D.h., $\mathcal{A} = V = \mathbb{K}^n$ und die Addition ist die übliche Addition in \mathbb{K}^n . Sei $A \in \text{Mat}(m, n, \mathbb{K})$. Betrachte das lineare Gleichungssystem $Ax = b$. Angenommen, das System ist lösbar. Dann ist die Lösungsmenge \mathcal{U} ein Unterraum und $V_{\mathcal{U}} = \text{Kern}_{f_A}$. Tatsächlich, nach Satz 47(b) LAAG I ist die Lösungsmenge $\{\tilde{x} + v \text{ wobei } v \in \text{Kern}_{f_A}\}$.

Affine Abbildungen

Definition 13 Seien $\mathcal{A}_1, \mathcal{A}_2$ affine Räume über V_1, V_2 . Eine Abbildung $F : \mathcal{A}_1 \rightarrow \mathcal{A}_2$ heißt **affin**, wenn es eine lineare Abbildung $f : V_1 \rightarrow V_2$ existiert, so dass für jede $a, b \in \mathcal{A}_1$ gilt

$$f(\overrightarrow{ab}) = \overrightarrow{F(a)F(b)}. \quad (*)$$

Eine bijektive affine Abbildung heißt **Affinität** oder ein **affiner Isomorphismis**

Bedienung (*) umformulieren: für alle $a_1 \in \mathcal{A}_1, v \in V_1$ gilt

$$F(a_1 + v) = F(a_1) + f(v). \quad (**)$$

Lemma 8

Affine Abbildung F ist Affinität $\iff f$ ist ein Isomorphismus.

Beweis \Leftarrow . Bijektiv = surjektiv und injektiv

f sei injektiv. Betrachte $a \neq b \in \mathcal{A}_1$. Nach Lemma 5 $\overrightarrow{ab} \neq \vec{0}$. Dann ist $\overrightarrow{F(a)F(b)} = f(\overrightarrow{ab}) \neq \vec{0}$. Dann ist $F(a) \neq F(b)$, also F ist injektiv.

f sei surjektiv. Betrachte $a_1 \in \mathcal{A}_1, a_2 := F(a_1) \in \mathcal{A}_2$. Nach (\mathcal{A}_2) gilt:

$\mathcal{A}_1 := \{a_1 + v_1 \text{ wobei } v_1 \in V_1\}$ und $\mathcal{A}_2 := \{a_2 + v_2 \text{ wobei } v_2 \in V_2\}$.

Da alle $v_2 \in V_2$ Bilder von Elementen von V_1 sind, und wegen (**) ist

$\text{Bild}_F = \mathcal{A}_2$. Also, ist f bijektiv, so ist F auch bijektiv.

Definition 13 – noch einmal Seien $\mathcal{A}_1, \mathcal{A}_2$ affine Räume über V_1, V_2 .

Eine Abbildung $F : \mathcal{A}_1 \rightarrow \mathcal{A}_2$ heißt **affin**, wenn es eine lineare Abbildung $f : V_1 \rightarrow V_2$ existiert, so dass für jede $a, b \in \mathcal{A}_1$ gilt

$$f(\overrightarrow{ab}) = \overrightarrow{F(a)F(b)}. \quad (*)$$

Eine bijektive affine Abbildung heißt **Affinität** oder ein **affiner Isomorphismis**

Bedingung (*) ist äquivalent zu:

$$F(a_1 + v) = F(a_1) + f(v). \quad (**)$$

Beweis \implies . Sei F injektiv. Betrachte $u \neq v \in V_1$ und $a_1 \in \mathcal{A}_1$. Nach (A_2) gilt $a_1 + v \neq a_1 + u$. Dann ist $F(a_1 + v) \neq F(a_1 + u)$; also $F(a_1) + f(v) \neq F(a_1) + f(u)$, also $f(v) \neq f(u)$, also f ist injektiv.

Sei F surjektiv. Betrachte ein beliebiges $v_2 = \overrightarrow{a_2 b_2} \in V_2$. Da F surjektiv, ist $a_2 = F(a_1), b_2 = F(b_1)$ für irgendwelche $a_1, b_1 \in \mathcal{A}_1$. Dann ist

$$f(\overrightarrow{a_1 b_1}) \stackrel{(*)}{=} \overrightarrow{F(a_1)F(b_1)}, \text{ also } v_2 \in \text{Bild}_f, \text{ also } f \text{ ist surjektiv.} \quad \square$$

Bsp. (Parallelverschiebung) Sei $\mathcal{A}_1 = \mathcal{A}_2 = \mathcal{A}$ und $V_1 = V_2 = V$. Sei $v_0 \in V$ ein fester Vektor. Dann ist die Abbildung $f : \mathcal{A} \rightarrow \mathcal{A}$,
 $F(a) := a + v_0$ affin. In dem Fall ist $f = Id$. Tatsächlich, $\overrightarrow{F(a)F(b)}$ ist der Vektor $w \in V$ s.d. $F(a) + w = F(b)$. Da $F(a) = a + v_0$, $F(b) = b + v_0$ ist $F(a) + w = a + v_0 + w = a + w + v_0 = F(a + w)$. Also, $f(w) = w$.
 Eine solche Abbildung heißt **Translation (oder Parallelverschiebung)**. Sie ist ein Affinität nach Lemma 8.

HauptBsp. Seien $a_1 \in \mathcal{A}_1$, $a_2 \in \mathcal{A}_2$, $f : V_1 \rightarrow V_2$ sei eine lineare Abbildung. Dann ist

$$F : \mathcal{A}_1 \rightarrow \mathcal{A}_2 \quad , \quad F(a) := a_2 + f(\overrightarrow{a_1 a})$$

eine affine Abbildung.

Tatsächlich, $F(a_1 + v) = F(a_1) + f(v)$.

Also, alle affine Abbildungen sind wie in HauptBsp.

In Worten: Jede affine Abbildung ist eine Verkettung von einer Translation und einer linearen Abbildung

Def. 14 Sei \mathcal{A} affiner Raum über \mathbb{K} - Vektorraum V , $\dim(V) = n \geq 1$. Ein $(n + 1)$ -Tupel (a_0, \dots, a_n) von Punkten a_0, \dots, a_n aus \mathcal{A} heißt **Koordinatensystem** für \mathcal{A} , falls die Vektoren $\overrightarrow{a_0 a_1}, \overrightarrow{a_0 a_2}, \overrightarrow{a_0 a_3}, \dots, \overrightarrow{a_0 a_n}$ linearunabhängig sind (und damit eine Basis von V bilden). Ist $x \in \mathcal{A}$ ein beliebiger Punkt, so gilt

$$\overrightarrow{a_0 x} = \sum_{i=1}^n x_i \overrightarrow{a_0 a_i}$$

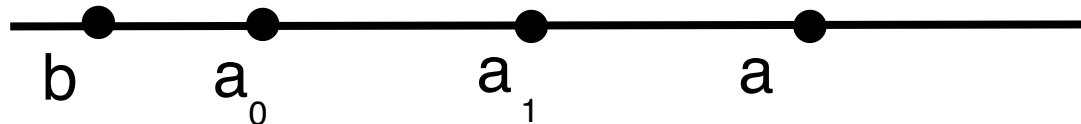
mit durch x eindeutig bestimmten $x_i \in \mathbb{K}$. Diese x_i heißen die **Koordinaten** von x bzg. Koordinatensystem (a_0, \dots, a_n) , und $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ heißt der **Koordinatenvektor** von x .

Koordinaten auf einer Gerade

Gerade = 1-dimensionaler affiner Raum

Koordinatensystem besteht aus $n + 1 = 2$ Punkten (a_0, a_1) s.d. $a_0 \neq a_1$.

Koordinatenvektor eines Punkts a besteht aus einem Zahl x s.d. $a_0 + x\overrightarrow{a_0a_1} = a$



Punkt	Koordinate
a_1	1
a_0	0
a	2
b	$-1/2$

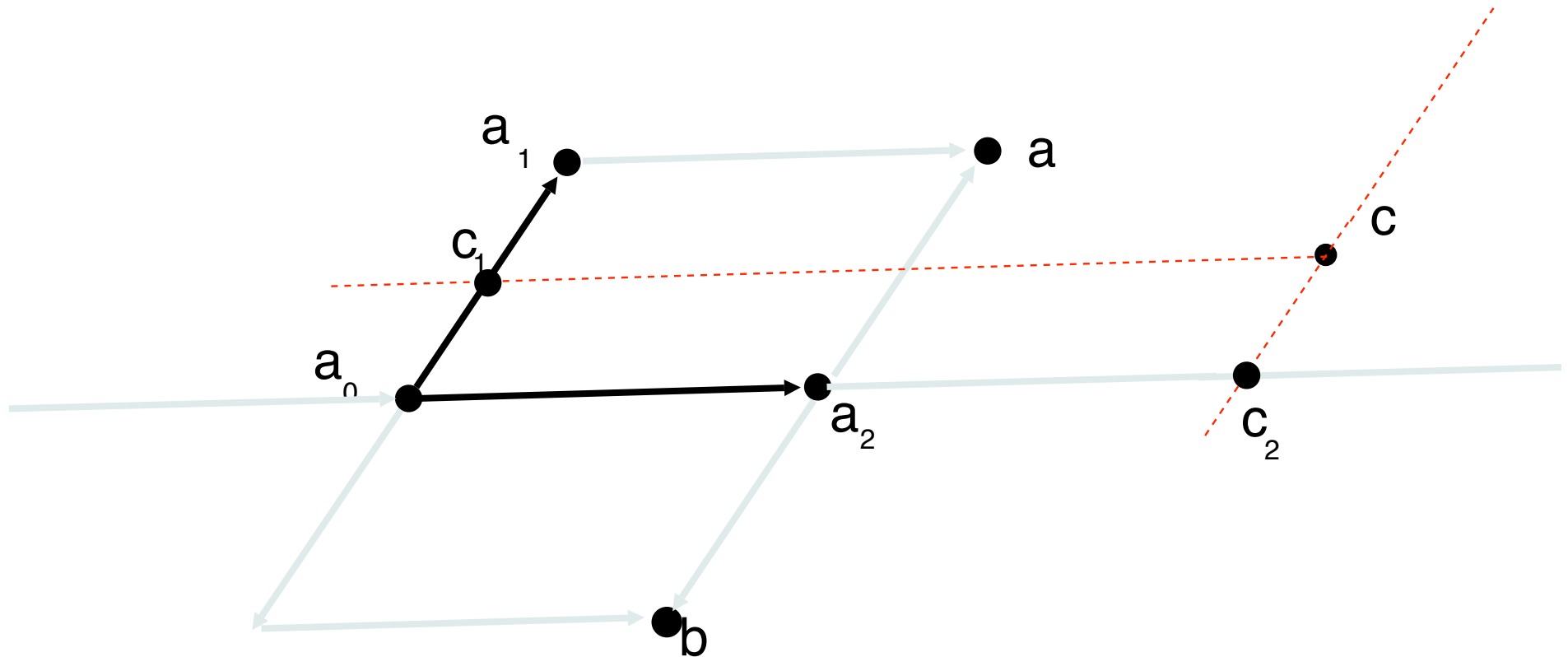
Koordinaten auf einer Ebene

Ebene = 2-dimensionaler affiner Raum

Koordinatensystem besteht aus $n + 1 = 3$ Punkten (a_0, a_1, a_2) s.d. $\{\overrightarrow{a_0a_1}, \overrightarrow{a_0a_2}\}$ linear unabhängig sind. Nach Hausaufgabe 2a, Blatt 7 LAAG I bedeutet dies das $\overrightarrow{a_0a_1}, \overrightarrow{a_0a_2}$ nichtproportional sind.

Koordinatenvektor eines Punkts a ist $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ s.d. $a_0 + x_1 \overrightarrow{a_0a_1} + x_2 \overrightarrow{a_0a_2} = a$

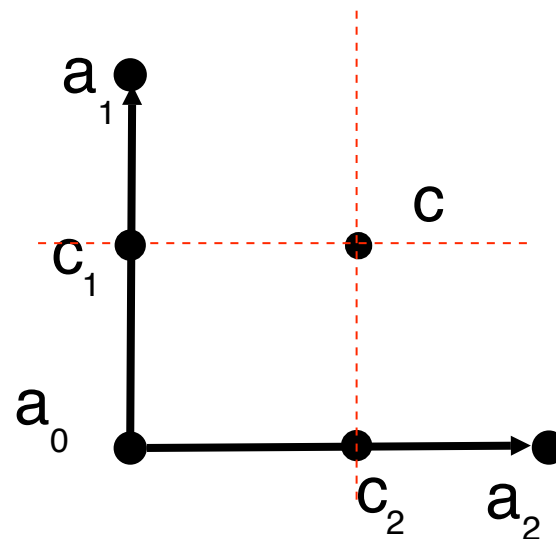
Für a auf dem Bild: $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ Für b : $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ Für c : $\begin{pmatrix} 1/2 \\ 2 \end{pmatrix}$



Frage Wann sind affine Koordinaten auf der (üblichen) Ebene cartesisch?

Antwort Wenn die Strecken (a_0, a_1) Länge 1 haben und paarweise orthogonal sind.

Weil in dem Fall das Parallelogramm $a_0c_1cc_2$ ein Rechteck ist, und deswegen die Punkte c_1, c_2 sind Projektionen des c auf die Geraden a_0a_1, a_0a_2 . Da die Länge von $\lambda\vec{v}$ gleich $|\lambda||\vec{v}|$ ist, sind die Koordinaten von c die Längen von $(a_0, c_1), (a_0, c_2)$ (mit Vorzeichen)



Satz 17 (Hauptsatz der affinen Geometrie) Sei $\mathcal{A}_1, \mathcal{A}_2$ zwei affine Räume der gleichen Dimension n (über \mathbb{K} -Vektorräume V_1 bzw. V_2). Seien (a_0, \dots, a_n) und (b_0, \dots, b_n) die Koordinatensysteme in \mathcal{A}_1 bzw. \mathcal{A}_2 . Dann existiert genau eine affine Abbildung $F : \mathcal{A}_1 \rightarrow \mathcal{A}_2$ s.d. $F(a_i) = b_i$. Diese Abbildung ist Affinität.

Beweis. Nach Definition sind $(\overrightarrow{a_0 a_1}, \overrightarrow{a_0 a_2}, \overrightarrow{a_0 a_3}, \dots, \overrightarrow{a_0 a_n})$ und $(\overrightarrow{b_0 b_1}, \overrightarrow{b_0 b_2}, \overrightarrow{b_0 b_3}, \dots, \overrightarrow{b_0 b_n})$ Basen von V_1 bzw. V_2 . Nach Satz 30 LAAG I (Vorl. 12) gibt es genau eine lineare Abbildung f , so dass $f(\overrightarrow{a_0 a_i}) = \overrightarrow{b_0 b_i}$. Da die Tupel Basen sind, ist f ein Isomorphismus. Betrachte die Abbildung

$$F : \mathcal{A}_1 \rightarrow \mathcal{A}_2, \quad F(a) := b_0 + f(\overrightarrow{a_0 a}) \quad (\text{aus dem HauptBsp.})$$

Die Abbildung ist eine Affinität nach Lemma 8. Z.z.: $F(a_i) = b_i$.
 $F(a_0) = b_0$, da $\overrightarrow{a_0 a_0} \stackrel{\text{Lem.5}}{=} \vec{0}$ und $b_0 + f(\vec{0}) \stackrel{\text{Lem.5}}{=} b_0$

$$F(a_i) = F(a_0 + \overrightarrow{a_0 a_i}) = F(a_0) + f(\overrightarrow{a_0 a_i}) = b_0 + \overrightarrow{b_0 b_i} = b_i.$$

Folgerung *Alle affine Räume (über \mathbb{K} – Vektorraum) der gleichem endlichen Dimension sind affin isomorph.*

Diese Aussage erlaubt, alle Probleme der (endlichdimensionalen) affinen Geometrie in einem \mathbb{K}^n zu betrachten.

Lass uns alle affine Abbildungen von \mathcal{A}_1 nach \mathcal{A}_2 beschreiben.

Satz 18 *Seien $\mathcal{A}_1 = \mathbb{K}^n$, $\mathcal{A}_2 = \mathbb{K}^m$ Standard-Räume (sie sind affine Räume jeweils über \mathbb{K}^n , \mathbb{K}^m). Dann jede affine Abbildung $F : \mathcal{A}_1 \rightarrow \mathcal{A}_2$ kann man als $F(x) = a + Ax$, wobei $a \in \mathbb{K}^m$ und $A \in \text{Mat}(m, n, \mathbb{K})$ sind.*

Beweis. Setze $a = F(\vec{0})$. Da $\vec{0} + x = x$, ist

$F(x) = F(\vec{0} + \vec{x}) \stackrel{(**)}{=} F(\vec{0}) + f(x) = a_1 + Ax$ für eine Matrix $A \in \text{Mat}(m, n, \mathbb{K})$. □

Wiederholung: Die Lösungsmenge eines linearen Gleichungssystems ein affiner Unterraum ist.

Satz 19 Sei \mathcal{U} ein affiner Unterraum von Standard-Raum \mathbb{K}^n . Dann existiert ein $m \leq n$, $m \in \mathbb{N}$, eine Matrix $A \in \text{Mat}(m, n, \mathbb{K})$ und ein $b \in \mathbb{K}^m$ s.d. die Lösungsmenge von $Ax = b$ ist \mathcal{U} .

Beweis. Sei $k := \dim(\mathcal{U})$. Man betrachte eine Basis (v_1, \dots, v_k) von $V_{\mathcal{U}}$. Nach Basisergänzungssatz (Folgerung (c) Vorl. 11 LAAG I) gibt es eine Basis $(v_1, \dots, v_k, w_1, \dots, w_m)$ von V , wobei $m = n - k$. Man betrachte die lineare Abbildung $f : \mathbb{K}^n \rightarrow \mathbb{K}^m$, die die Vektoren v_1, \dots, v_k auf $\vec{0}$, und die Vektoren w_1, \dots, w_m jeweils auf Standard-Basisvektoren $e_1, \dots, e_m \in \mathbb{K}^m$ abbildet (Existiert nach Satz 30 Vorl 12. LAAG I). Sei $A \in \text{Mat}(m, n, \mathbb{K})$ deren Matrix. Nehme ein $a \in \mathcal{U}$ und setze $b = Aa$.

Wir zeigen: \mathcal{U} ist die Lösungsmenge von $Ax = b$. Wir zeigen zuerst: $\text{Kern}_f = V_{\mathcal{U}}$.

Zuerst $V_{\mathcal{U}} \subseteq \text{Kern}_f$: Tatsächlich, jedes $v \in V_{\mathcal{U}}$ ist eine Linearkombination $\sum_{i=1}^k \lambda_i v_i$ und deswegen $f(v) = f(\sum_{i=1}^k \lambda_i v_i) = \sum_{i=1}^k \lambda_i f(v_i) = \sum_{i=1}^k \lambda_i \vec{0} = \vec{0}$. Also jedes $v \in V_{\mathcal{U}}$ liegt in Kern_f .

Wir zeigen: $V_{\mathcal{U}} \supseteq \text{Kern}_f$ (jeder Vektor $w \in \text{Kern}_f$ liegt in $V_{\mathcal{U}}$. Da

$(v_1, \dots, v_k, w_1, \dots, w_m)$ eine Basis in V ist,

$w = \sum_{i=1}^k \lambda_i v_i + \sum_{j=1}^m \mu_j w_j$. Dann ist $f(w) =$

$f(\sum_{i=1}^k \lambda_i v_i) + f(\sum_{j=1}^m \mu_j w_j) = \vec{0} + \sum_{j=1}^m \mu_j f(w_j) = \sum_{j=1}^m \mu_j e_j$.

Also, aus $f(w) = \vec{0}$ folgt $\sum_{j=1}^m \mu_j e_j = 0$; dann alle $\mu_j = 0$; dann $w \in V_{\mathcal{U}}$. Also, $\text{Kern}_f = V_{\mathcal{U}}$.

Wir haben b so konstruiert, dass $Aa = b$ ist, also a ist die Lösung von $Ax = b$. Nach Satz 47(b) LAAG I ist die Lösungsmenge

$\{a + v \text{ wobei } v \in \text{Kern}_f\} = \{a + v \text{ wobei } v \in V_{\mathcal{U}}\} = \mathcal{U}$. □

Affine Eigenschaften ($\mathbb{K} = \mathbb{R}$, falls nicht explicit erwähnt.)

Def. 15 Sei M eine Teilmenge eines affinen Raums \mathcal{A} über V (über \mathbb{K}). Eine Eigenschaft der Menge M heißt **affin**, wenn für jede Affinität $F : \mathcal{A} \rightarrow \mathcal{A}_1$ die Bildmenge $\{F(a) \mid a \in M\}$ auch diese Eigenschaft hat.

Bezeichnung Die Bildmenge einer Menge M unter der Abbildung F werden wir $Bild_F(M) := \{F(a) \mid a \in M\}$ bezeichnen.

Bsp. Eigenschaft „Unterraum zu sein“ ist eine affine Eigenschaft. Tatsächlich, man betrachte einen Unterraum $\mathcal{U} \subseteq \mathcal{A}$. Ist F ein affiner Isomorphismus, so ist $F(a) = F(a_0) + f(\overrightarrow{a_0 a})$; da $Bild_f(V_{\mathcal{U}})$ ein Untervektorraum ist, ist $Bild_F(\mathcal{U})$ ein Unterraum.

Bsp. Eigenschaften „Gerade“, „Ebene“, oder „Hyperebene“ zu sein sind affine Eigenschaften. Tatsächlich, nach Definition 12 ist

Gerade	Unterraum der Dimension 1
Ebene	Unterraum der Dimension 2
Hyperebene	Unterraum der Dimension $n - 1$

Da die Isomorphismen die Dimension des Untervektorraums erhalten, ist Bild einer Geraden, Ebenen oder Hyperebenen jeweils eine Gerade, Ebene oder Hyperebene.

Bsp. Eigenschaft aus 3 Punkten zu bestehen ist eine affine Eigenschaft.

Definition 16 Seien $a_0, a_1 \in \mathcal{A}$. Die **Strecke** mit Endpunkten a_0, a_1 ist die Menge $\{a_0 + \lambda \overrightarrow{a_0 a_1} \mid \text{wobei } 0 \leq \lambda \leq 1\}$. (ist sinnvoll nur wenn $\mathbb{K} = \mathbb{R}$ ist)



Bsp. Eigenschaft einer Strecke zu sein ist auch eine affine Eigenschaft.

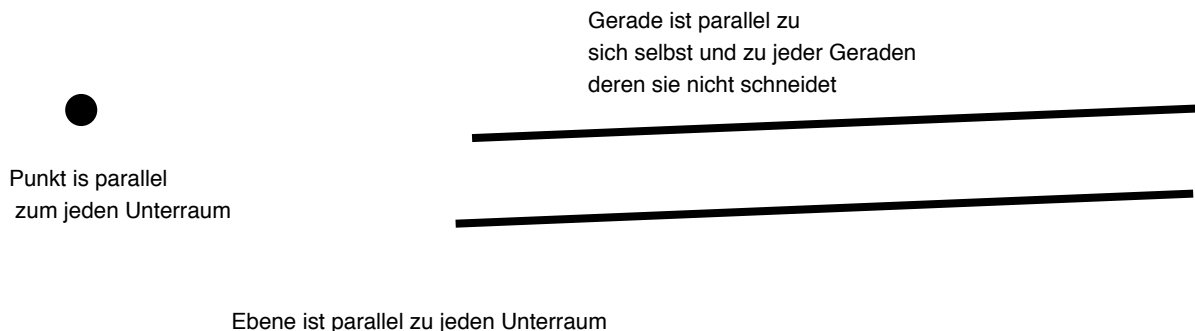
Tatsächlich, für jeden Punkt einer Strecke $\{a_0 + \lambda \overrightarrow{a_0 a_1} \mid \text{wobei } 0 \leq \lambda \leq 1\}$ gilt

$$F(a_0 + \lambda \overrightarrow{a_0 a_1}) \stackrel{(**) \text{ in Def. 13}}{=} F(a_0) + \underbrace{\lambda f(\overrightarrow{a_0 a_1})}_{\substack{\overrightarrow{\lambda F(a_0) F(a_1)}} \text{ nach } (*) \text{ in Def. 13}} .$$

Dann ist die Bildmenge $\{F(a_0) + \lambda \overrightarrow{F(a_0) F(a_1)} \mid \text{wobei } 0 \leq \lambda \leq 1\}$ auch eine Strecke.

Def. 17 Seien $\mathcal{U}_1, \mathcal{U}_2$ Unterräume von \mathcal{A} (man kann es über einem \mathbb{K} -Vektorraum V definieren). Sie heißen *parallel* zu einem anderen, falls $V_{\mathcal{U}_1} \subseteq V_{\mathcal{U}_2}$ oder $V_{\mathcal{U}_2} \subseteq V_{\mathcal{U}_1}$.

Auf der Ebene



Bsp. Jeder Unterraum ist zu sich selbst, zum ganzen Raum und zu einem Punkt Parallel.

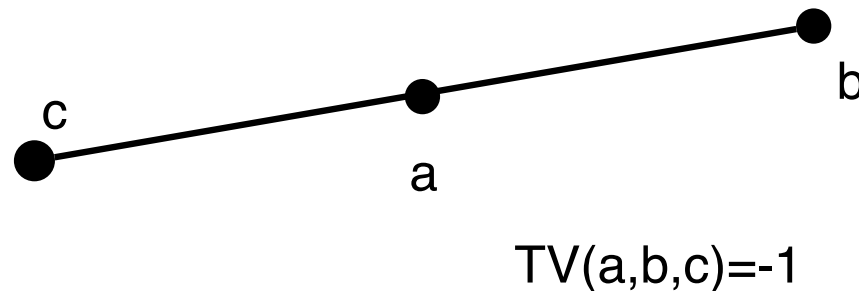
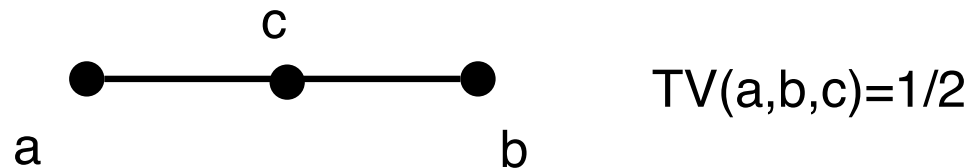
Einfach zu sehen: Seien $\mathcal{U}_1, \mathcal{U}_2$ parallele affine Unterräume. Dann gilt: entweder enthält ein Unterraum den anderen, oder sie haben keine Schnittpunkte.

Bsp. Eigenschaft aus zwei Unterräume zu bestehen, die keine Schnittpunkten haben ist eine affine Eigenschaft.

Definition 18 Sei (a, b, c) ein Tripel von Punkten in \mathcal{A} .

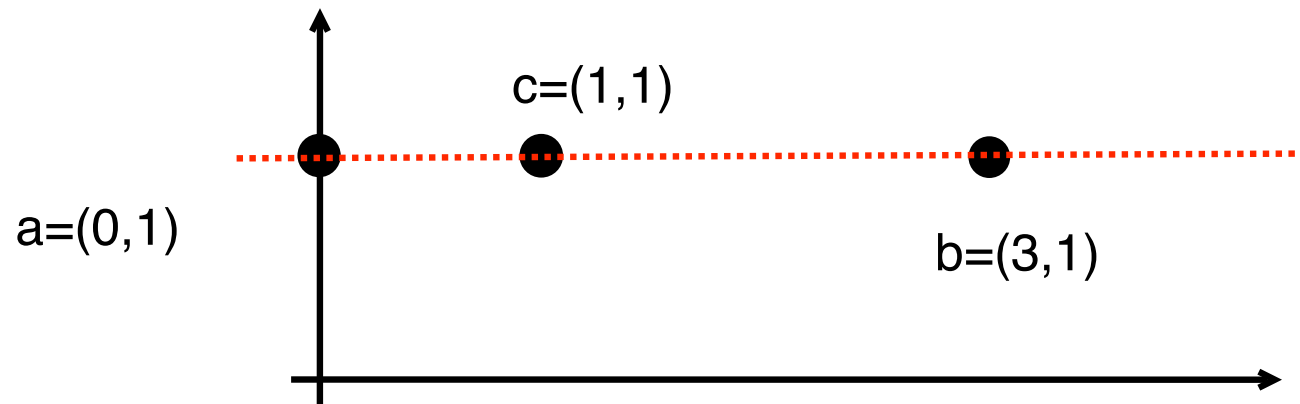
1. Die Punkte a, b, c heißen **kollinear**, falls sie auf einer Geraden liegen.
2. Ist (a, b, c) ein kollineares Punktetripel und $a \neq b$, so heißt der durch die Gleichung $\vec{ac} = \lambda \vec{ab}$ eindeutig bestimmte Skalar λ das **Teilverhältnis** des kollinearen Punktetripels (a, b, c) , bezeichnet durch $TV(a, b, c)$.

Bsp. c heißt der **Mittelpunkt** der Strecke (a, b) , wenn $TV(a, b, c) = \frac{1}{2}$ gilt (hat Sinn falls $\frac{1}{2} = 2^{-1}$ wohldefiniert ist, also falls $1 + 1 \neq 0$ ist.) Das ist äquivalent zu $c = a + \frac{1}{2}\vec{ab}$.



Bsp. In \mathbb{R}^2 sind $a = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $b = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$ und $c = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ kollinear.

Wegen $\vec{ac} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{3}\vec{ab} = \frac{1}{3}\begin{pmatrix} 3 \\ 0 \end{pmatrix}$ gilt $TV(a, b, c) = \frac{1}{3}$.



Geometr. Bedeutung: Sei $\mathcal{A} = \mathcal{E}$. Dann gilt:

$$|TV(a, b, c)| = \frac{\text{Länge}(a, c)}{\text{Länge}(a, b)}.$$

Ferner gilt: liegt c auf der Strecke a, b , so ist $0 \leq TV(a, b, c) \leq 1$.

liegt a auf der Strecke b, c , so ist $TV(a, b, c) \leq 0$.

liegt b auf der Strecke a, c , so ist $TV(a, b, c) \geq 1$.

Lemma 9 Sind (a, b, c) kollinear und F ein affiner Isomorphismus, so sind $(F(a), F(b), F(c))$ auch kollinear; ferner gilt:
 $TV(a, b, c) = TV(F(a), F(b), F(c))$.

In Worten: Affinitäten erhalten Kollinearität und das Teilverhältnis.

Beweis. Liegen a, b, c auf einer Geraden \mathcal{G} , so liegen $F(a), F(b), F(c)$ auf der Bildmenge $Bild_F(\mathcal{G})$. Da $Bild_F(\mathcal{G})$ wieder eine Gerade ist, sind $F(a), F(b), F(c)$ kollinear.

Nach Definition 18 ist

$$\overrightarrow{ac} = TV(a, b, c)\overrightarrow{ab}.$$

Nach Definition 13 ist $\overrightarrow{F(a)F(c)} = f(\overrightarrow{ac})$ und $\overrightarrow{F(a)F(b)} = f(\overrightarrow{ab})$. Also,

$$\overrightarrow{F(a)F(c)} = TV(a, b, c)\overrightarrow{F(a)F(b)},$$

also $TV(F(a), F(b), F(c)) = TV(a, b, c)$

Folgerung:

Punkt c ist der Mittelpunkt der Strecke (a, b)	affine Eigenschaft
Punkt c teilt die Strecke (a, b) in Verhältnis 2 : 1	affine Eigenschaft

$\frac{\text{Flächeninhalt einer Menge}}{\text{Flächeninhalt der zweiten Menge}}$ ist eine affine Eigenschaft

In der Vorl. 16 LAAG I haben wir verstanden, dass eine lineare Abbildung $f_A : \mathcal{E} \rightarrow \mathcal{E}$ (wobei \mathcal{E} die übliche „Schulgeometrische“ Ebene ist)

Flächeninhalt jeder Figure mit Faktor $\det(A)$ multipliziert.

Wir wissen dass jede Affinität von \mathcal{E} nach \mathcal{E} die Form

$F(x) = F(a) + f_A(\overrightarrow{ax})$ hat. Da die Translation (Parallelverschiebung)

offensichtlich den Flächeninhalt erhält, multipliziert die affine Abbildung Flächeninhalt jeder Figure mit Faktor $\det(A)$.

Deswegen erhalten die Affinitäten von \mathcal{E} $\frac{\text{Flächeninhalt einer Menge}}{\text{Flächeninhalt der zweiten Menge}}$.

Selbstverständlich ist das Phänomen mehrdimensional : in $Dim(3)$ und in $Dim(n)$ (und in Lorenz-Geometrie) ist

$Volum(\text{Bild}_F(\text{Menge})) = \det(f)(Volum(\text{Menge}))$ und deswegen

$$\frac{Volum(\text{Bild}_F(\text{Menge}_1))}{Volum(\text{Bild}_F(\text{Menge}_2))} = \frac{\det(f) \cdot Volum(\text{Menge}_1)}{\det(f) \cdot Volum(\text{Menge}_2)} = \frac{Volum(\text{Menge}_1)}{Volum(\text{Menge}_2)}$$

Wiederholung Eine Eigenschaft einer Teilmenge $M \subseteq \mathcal{A}$ heißt **affin**, wenn für jede Affinität $F : \mathcal{A} \rightarrow \mathcal{A}_1$ die Bildmenge $\{F(a) \mid a \in M\}$ auch diese Eigenschaft hat.

Punkt zu sein	affin
Gerade (zu sein)	affin
Strecke	affin
Dreieck	affin
parallele Gerade	affin
Winkel (in \mathcal{E}_2 oder \mathcal{E}_3) zwischen zwei Geraden ist gerade	nicht affin
Länge einer Strecke (in \mathcal{E}_2 oder \mathcal{E}_3) ist gleich 5	nicht affin
Punkt c ist die Mittelpunkt der Strecke (a, b)	affin
Punkt c teilt die Strecke (a, b) in Verhältnis 2 : 1	affin
Flächeninhalt (in \mathcal{E}_2 oder \mathcal{E}_3)	nicht affin
$\frac{\text{Flächeninhalt einer Menge}}{\text{Flächeninhalt der zweiten Menge}} = 5$	affin

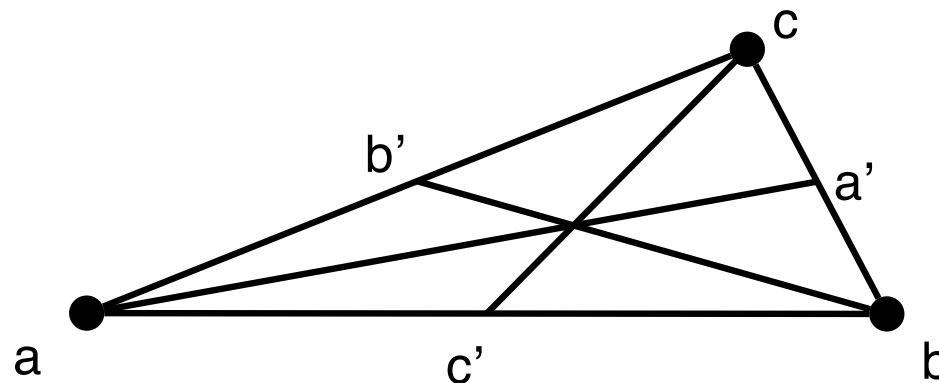
Anwendung in der (Schul)geometrie

Wir sagen, dass eine geometrische Aufgabe **affin** ist, falls nur affine Eigenschaften gegeben sind.

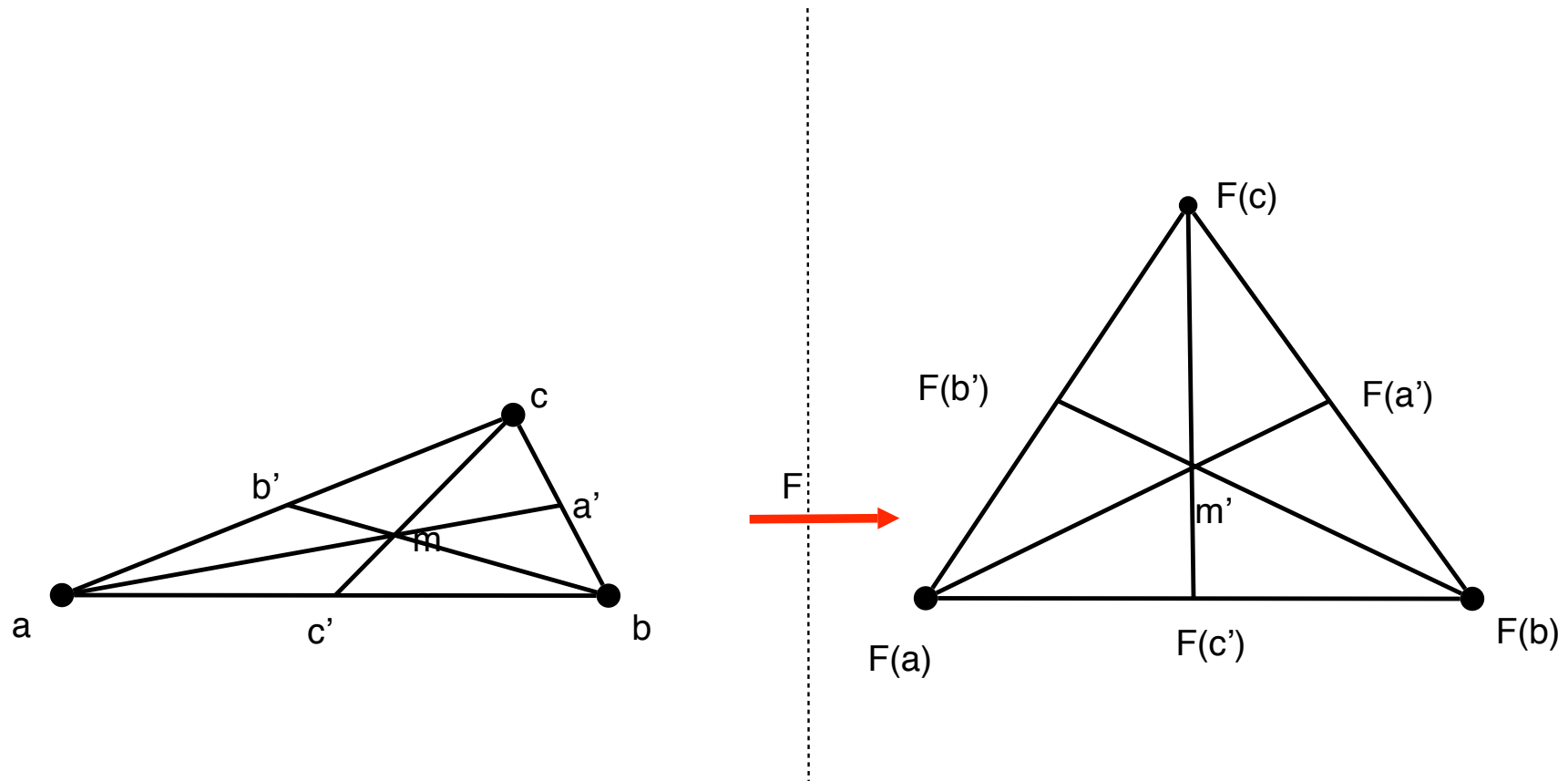
Um eine affine Aufgabe zu lösen, können Sie zuerst eine passende Affinität anwenden. Wenn Sie dies klug genug tun, vereinfacht dies die Aufgabe.

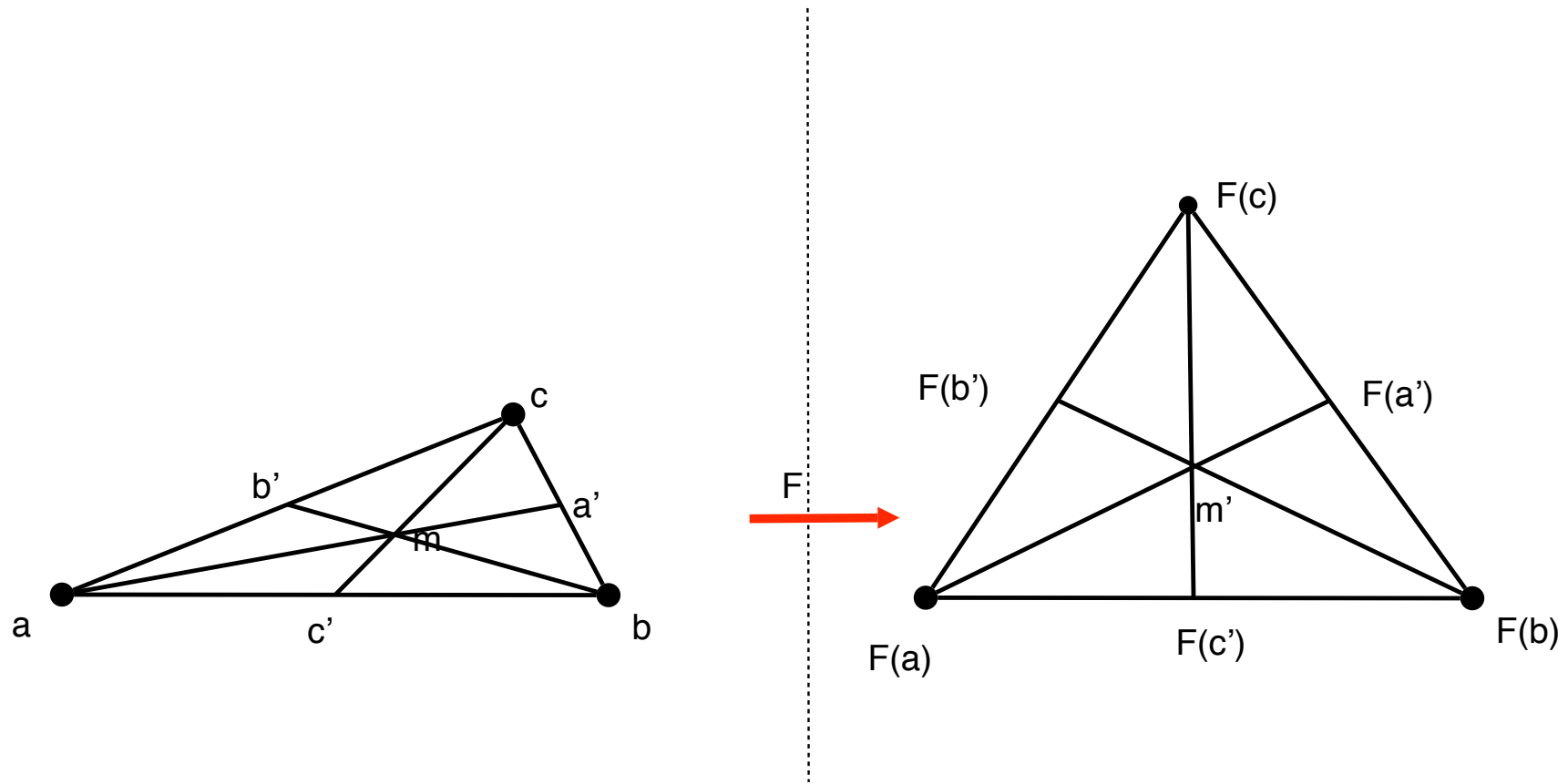
BspAufgabe. Beweisen Sie, dass die Seitenhalbierenden eines Dreiecks
(a) in einem Punkt schneiden
(b) dass der Schnittpunkt sie in Verhältnis $2 : 1$ teilt,
(c) dass die 6 Dreiecke, in denen die Seitenhalbierend das Dreieck teilen gleichen Flächeninhalt haben.

Bemerkung Hausaufgabe 4 ist die Verallgemeinerung von (a) und (c) für 3-dim Fall.



Man betrachte eine Affinität, die die Ecke des Dreiecks ABC in Ecken eines regelmäßigen Dreiecks überführt (existiert nach Satz 17). Die Abbildung führt die Seiten in Seiten über. Die Abbildung führt die Mittelpunkte der Seiten in die Mittelpunkte der Seiten über. Die Abbildung führt die Seitenhalbierenden in die Seitenhalbierenden über. Da in dem regelmäßigen Dreieck die Seitenhalbierende in einem Punkt schneiden, schneiden die Seitenhalbierende des ursprünglichen Dreiecks auch in einem Punkt. (a) ist bewiesen.





Da in einem regelmässigen Dreieck alle 6 Dreiecke in denen die Seitenhalbierende das Dreieck teilen gleich sind, sind ihre Flächeninhalte auch gleich, also

$$\frac{\text{Flächeninhalt eines kleinen Dreieck}}{\text{Flächeninhalt eines anderen kleinen Dreiecks}} = 1. \quad (*)$$

Da dies eine affine Eigenschaft ist, gilt (*) auch für das ursprünglichen Dreieck. (c) ist Bewiesen.

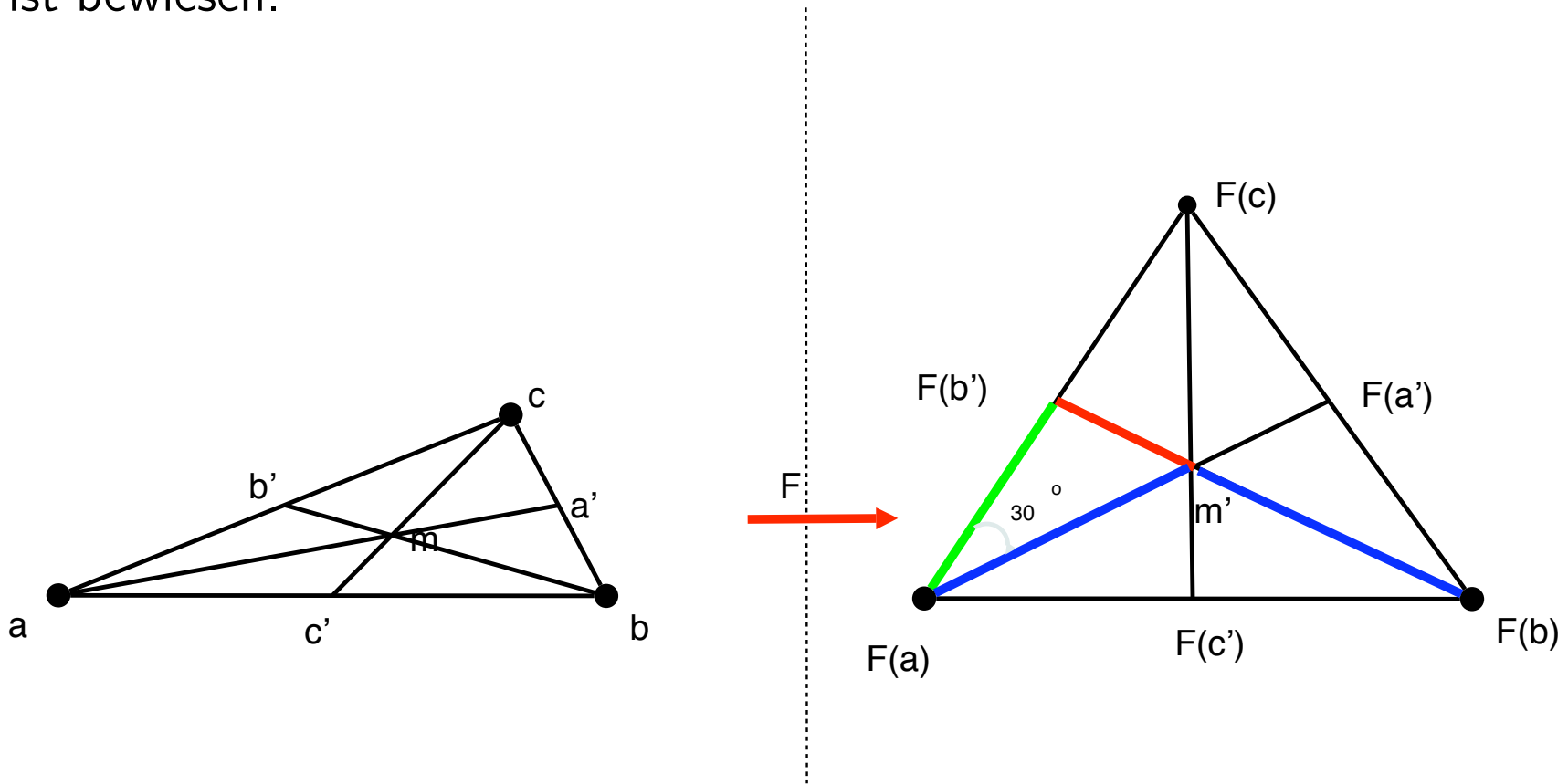
Man betrachte ein "kleines" Dreieck.

Da der Winkel $F(c)F(a)F(a')$ gleich 30° ist, ist

$|(\overrightarrow{m'}, \overrightarrow{F(b')})| = \sin(30^\circ) |(\overrightarrow{m'}, \overrightarrow{F(a)})| = \frac{1}{2} |(\overrightarrow{m'}, \overrightarrow{F(a)})|$. Da

$|(\overrightarrow{F(b)}, \overrightarrow{m'})| = |(\overrightarrow{m'}, \overrightarrow{F(a)})|$, teilt der Punkt m' die Seitenhalbierende $(F(b), F(b'))$ im Verhältnis $2 : 1$. Da Affinitäten die Teilverhältnis

erhalten, teilt m die Seitenhalbierende (b, b') im Verhältnis $2 : 1$. (b) ist bewiesen.

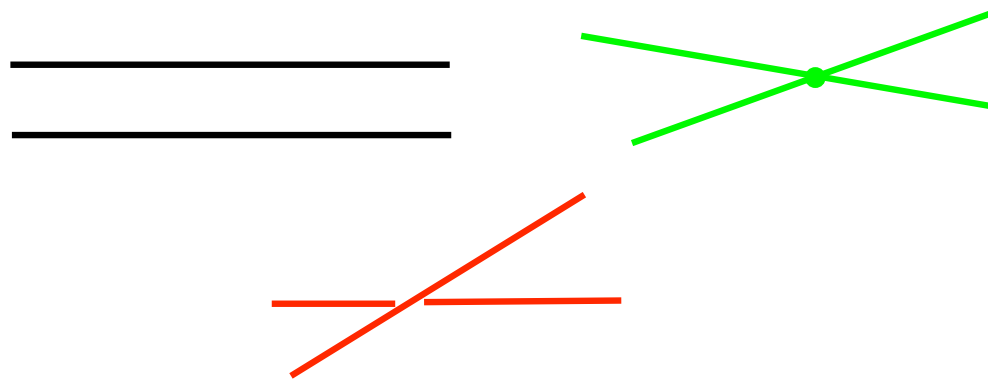


Lemma 10 Sei \mathcal{A} ein n -dimensionaler affiner Raum, \mathcal{U} affiner Unterraum und \mathcal{H} affine Hyperebene. Gilt $\mathcal{U} \cap \mathcal{H} = \emptyset$, so sind \mathcal{U} und \mathcal{H} parallel. (Körper ist \mathbb{K})

Bemerkung. Die Bedingung, dass \mathcal{H} Hyperebene ist, also dass $\dim(V_{\mathcal{H}}) = \dim(V) - 1$, ist wichtig:

Zwei verschiedene Gerade in \mathcal{E}_3 können

- ▶ parallel sein,
- ▶ windschief sein
- ▶ einander schneiden



Lemma 10 Sei \mathcal{A} ein n -dimensionaler affiner Raum, \mathcal{U} affiner Unterraum und \mathcal{H} affine Hyperebene. Gilt $\mathcal{U} \cap \mathcal{H} = \emptyset$, so sind \mathcal{U} und \mathcal{H} parallel. (Körper ist \mathbb{K})

Bemerkung. Die Bedingung, dass \mathcal{H} Hyperebene ist, also dass $\dim(V_{\mathcal{H}}) = \dim(V) - 1$, ist wichtig:

Zwei verschiedene Gerade in \mathcal{E}_3 können

- ▶ parallel sein,
- ▶ windschief sein
- ▶ einander schneiden

Windschiefe Geraden sind nicht parallel, haben aber keinen Schnittpunkt.

Lemma 10 Sei \mathcal{A} ein n -dimensionaler affiner Raum, \mathcal{U} affiner Unterraum und \mathcal{H} affine Hyperebene. Gilt $\mathcal{U} \cap \mathcal{H} = \emptyset$, so sind \mathcal{U} und \mathcal{H} parallel. (Körper ist \mathbb{K})

Widerspruchsbeweis. Angenommen sie sind nicht parallel. Man nehme ein $v \in V_{\mathcal{U}}$, $v \notin V_{\mathcal{H}}$. Man nehme eine Basis (h_1, \dots, h_{n-1}) in $V_{\mathcal{H}}$, und betrachte $\{h_1, \dots, h_{n-1}, v\}$. Die Menge ist linear unabhängig. Tatsächlich, sei

$$\lambda_1 h_1 + \dots + \lambda_{n-1} h_{n-1} + \lambda v = \vec{0} \quad (*)$$

Dann ist $\lambda = 0$, sonst kann man v als Linearkombination von h_1, \dots, h_{n-1} darstellen, was die Voraussetzungen widerspricht. Dann ist $(*)$ eine Linearkombination der Elemente aus $\{h_1, \dots, h_{n-1}\}$ und ist trivial, weil $\{h_1, \dots, h_{n-1}\}$ linear unabhängig ist. Dann ist $\{h_1, \dots, h_{n-1}, v\}$ linear unabhängig, und deswegen eine Basis.

Nehmen wir ein $a \in \mathcal{U}$ und ein $b \in \mathcal{H}$. Da (h_1, \dots, h_{n-1}, v) eine Basis ist, gilt $\vec{ab} = \lambda_1 h_1 + \dots + \lambda_{n-1} h_{n-1} + \lambda v$. Dann $b = a + \lambda_1 h_1 + \dots + \lambda_{n-1} h_{n-1} + \lambda v$, und deswegen

$$b - (\lambda_1 h_1 + \dots + \lambda_{n-1} h_{n-1}) = a + \lambda v. \quad (**)$$

Da $b \in \mathcal{H}$ und $(\lambda_1 h_1 + \dots + \lambda_{n-1} h_{n-1}) \in V_{\mathcal{H}}$, liegt die linke Seite von $(**)$ in \mathcal{H} . Da $b \in \mathcal{H}$ und $\lambda v \in V_{\mathcal{U}}$, liegt die rechte Seite von $(**)$ in \mathcal{U} . Also $\mathcal{U} \cap \mathcal{H} \neq \emptyset$. Widerspruch beweist das Lemma,

Ziel: Möglich viele Begriffe mit Hilfe von Begriff „Gerade“ zu definieren

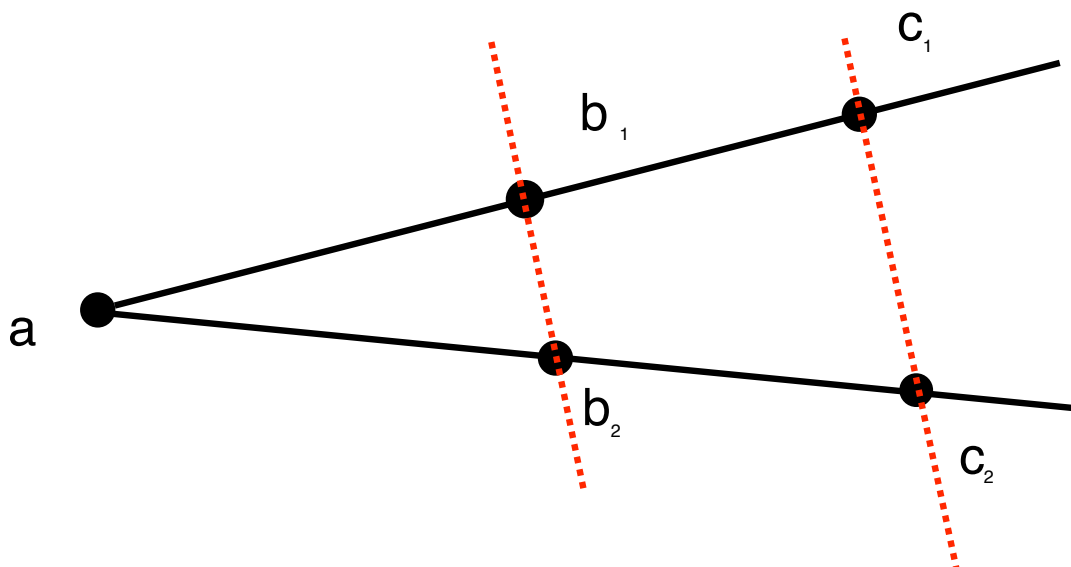
Plan

- ▶ Affine Unterräumen
- ▶ Affinität

Strahlensatz

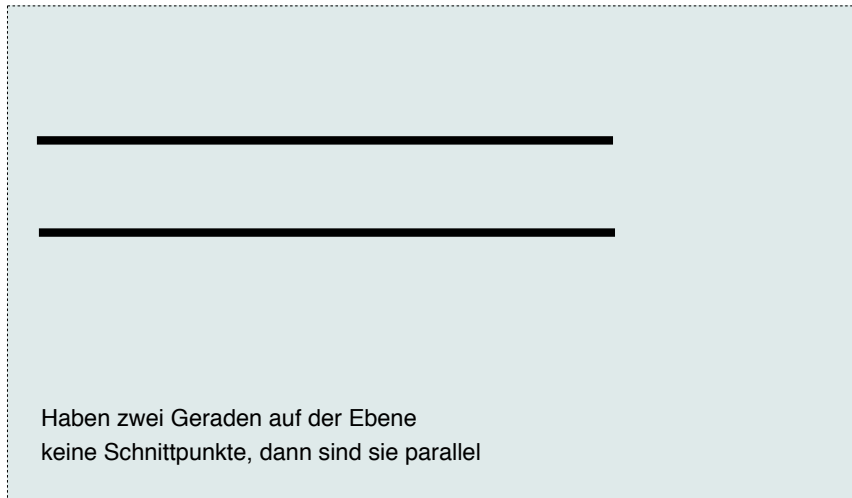
Bezeichnung: $\mathcal{G}_{b_1, b_2} := \{b_1 + \lambda \overrightarrow{b_1 b_2} \mid \lambda \in \mathbb{K}\} =$
(Eindeutige) Gerade, die die Punkte b_1, b_2 enthält.

Satz 20 (Strahlensatz) Seien (a, b_1, c_1) und (a, b_2, c_2) kollineare Punkttripel auf verschiedenen Geraden durch $a \notin \{b_1, b_2, c_1, c_2\}$. Dann sind die Geraden \mathcal{G}_{b_1, b_2} und \mathcal{G}_{c_1, c_2} genau dann parallel, wenn $TV(a, b_1, c_1) = TV(a, b_2, c_2)$.



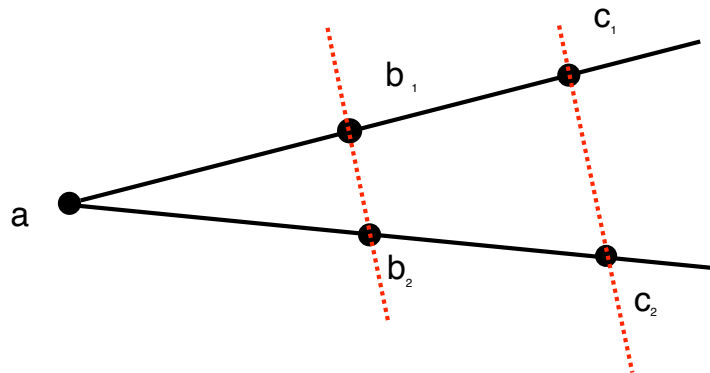
Def. 17 – Wiederholung Seien $\mathcal{U}_1, \mathcal{U}_2$ Unterräume von \mathcal{A} (man kann es über einem \mathbb{K} -Vektorraum V definieren). Sie heißen *parallel* zu einem anderen, falls $V_{\mathcal{U}_1} \subseteq V_{\mathcal{U}_2}$ oder $V_{\mathcal{U}_2} \subseteq V_{\mathcal{U}_1}$.

Lemma 10 – Wiederholung Sei \mathcal{A} ein n -dimensionaler affiner Raum, \mathcal{U} affiner Unterraum und \mathcal{H} affine Hyperebene. Gilt $\mathcal{U} \cap \mathcal{H} = \emptyset$, so sind \mathcal{U} und \mathcal{H} parallel. (Körper ist \mathbb{K})

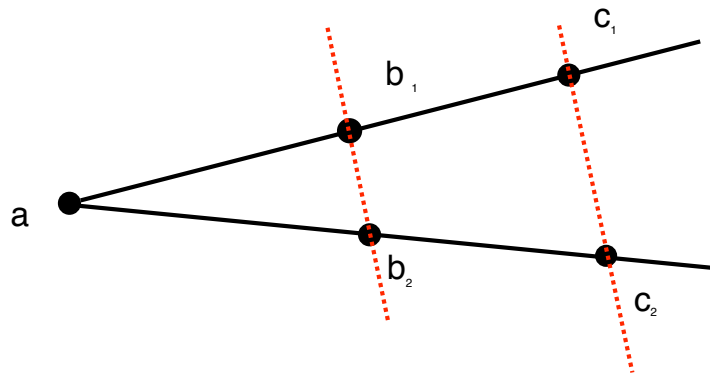


Strahlensatz

Satz 20 (Strahlensatz) Seien (a, b_1, c_1) und (a, b_2, c_2) kollineare Punkttripel auf verschiedenen Geraden durch $a \notin \{b_1, b_2, c_1, c_2\}$. Dann sind die Geraden \mathcal{G}_{b_1, b_2} und \mathcal{G}_{c_1, c_2} genau dann parallel, wenn $TV(a, b_1, c_1) = TV(a, b_2, c_2)$.



Wiederholung. Ist (a, b, c) ein kollineares Punkttripel und $a \neq b$, so heißt der durch die Gleichung $\vec{ac} = \lambda \vec{ab}$ eindeutig bestimmte Skalar λ das **Teilverhältnis** des kollinearen Punkttripels (a, b, c) , bezeichnet durch $TV(a, b, c)$.



Beweis. Nach Def. 18 gilt $TV(a, b_1, c_1) = \lambda \iff \overrightarrow{ac_1} = \lambda \overrightarrow{ab_1}$
 $TV(a, b_2, c_2) = \mu \iff \overrightarrow{ac_2} = \mu \overrightarrow{ab_2}$

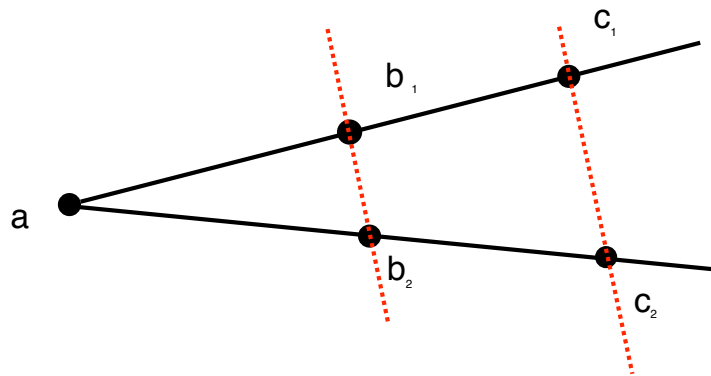
Nach Def. 17 die Gerade \mathcal{G}_{b_1, b_2} und \mathcal{G}_{c_1, c_2} sind parallel g.d.w. $\overrightarrow{c_1 c_2}$ und $\overrightarrow{b_1 b_2}$ linear abhängig sind. Aus $\overrightarrow{ac_1} + \overrightarrow{c_1 c_2} = \overrightarrow{ac_2}$ folgt

$\overrightarrow{c_1 c_2} = \overrightarrow{ac_2} - \overrightarrow{ac_1} = \mu \overrightarrow{ab_2} - \lambda \overrightarrow{ab_1}$ und ebenso $\overrightarrow{b_1 b_2} = \overrightarrow{ab_2} - \overrightarrow{ab_1}$. Dann

$$\alpha \overrightarrow{c_1 c_2} + \beta \overrightarrow{b_1 b_2} = 0 \iff$$

$\alpha(\mu \overrightarrow{ab_2} - \lambda \overrightarrow{ab_1}) + \beta(\overrightarrow{ab_2} - \overrightarrow{ab_1}) = (\alpha\mu + \beta)\overrightarrow{ab_2} - (\alpha\lambda + \beta)\overrightarrow{ab_1} = 0$. Da die Vektoren $\overrightarrow{ab_2}$ und $\overrightarrow{ab_1}$ linear unabhängig sind, ist dies äquivalent

zu dem linearen Gleichungssystem $\begin{cases} \alpha\mu + \beta = 0 \\ \alpha\lambda + \beta = 0 \end{cases}$ auf Unbekannten α, β .



Das System $\begin{cases} \alpha\mu + \beta = 0 \\ \alpha\lambda + \beta = 0 \end{cases}$ in Matrix Form:

$$\begin{pmatrix} \mu & 1 \\ \lambda & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Falls $\lambda \neq \mu$, ist die Matrix nichtausgeartet, weil

$$\det \begin{pmatrix} \mu & 1 \\ \lambda & 1 \end{pmatrix} = (\mu - \lambda) \neq 0.$$

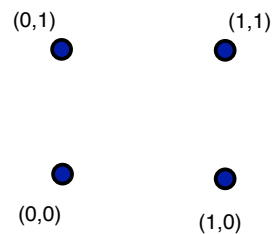
Also, falls $\lambda \neq \mu$ ist, gibt es nur triviale Lösung, was bedeutet dass $\overrightarrow{c_1 c_2}$ und $\overrightarrow{b_1 b_2}$ linear unabhängig sind und deswegen die Geraden \mathcal{G}_{b_1, b_2} und \mathcal{G}_{c_1, c_2} nicht parallel sind. Falls $\lambda = \mu$ ist, dann gibt es auch nichttriviale Lösungen (nach Satz 47(a) LAAG I) z.B. $\alpha = 1$, $\beta = \mu$, und deswegen sind $\overrightarrow{c_1 c_2}$ und $\overrightarrow{b_1 b_2}$ linear abhängig, und die Gerade parallel.

Def. 19 Sei $M \neq \emptyset$ eine Teilmenge des affinen Raums \mathcal{A} . Sie heißt **affin abgeschlossen**, falls M mit je zwei Punkte $a \neq b$ auch die Gerade $\mathcal{G}_{a,b} := \{a + \lambda \overrightarrow{ab} \mid \lambda \in \mathbb{R}\}$ durch a und b enthält.

Bsp. Ein affiner Unterraum ist eine affin abgeschlossene Teilmenge.

Bsp. einer affin abgeschlossene Teilmenge, die keinen Unterraum ist:

Man betrachte den \mathbb{Z}_2 -Vektorraum $(\mathbb{Z}_2)^2$ und einen affinen Raum über $(\mathbb{Z}_2)^2$. Nach Hauptsatz 17 können wir denken, dass der affine Raum auch $(\mathbb{Z}_2)^2$ ist, wie in StandardBsp.



Nach Definition 12 besteht die Gerade, die die Punkte a, b enthält aus allen Punkten der Form $a + \lambda \overrightarrow{ab}$. Da $\lambda \in \{0, 1\}$, besteht die ganze Gerade nur aus 2 Punkte a und b .

Deswegen ist jede Menge in $(\mathbb{Z}_2)^2$ affin abgeschlossen.

Aber die Menge $(\mathbb{Z}_2)^2 \setminus \{(0, 0)\}$ ist kein Unterraum (weil die Punkte $(1, 0), (1, 1), (0, 1)$ ein Koordinatensystem bilden; also jeder diesen Punkten enthaltenden Unterraum muss mit $(\mathbb{Z}_2)^2$ zusammenfallen.)

Def. 19 Sei $M \neq \emptyset$ eine Teilmenge des affinen Raums \mathcal{A} . Sie heißt **affin abgeschlossen**, falls M mit je zwei Punkte $a \neq b$ auch die Gerade $\mathcal{G}_{a,b} := \{a + \lambda \overrightarrow{ab} \text{ wobei } \lambda \in \mathbb{R}\}$ durch a und b enthält.

Satz 21 Angenommen, in der Körper \mathbb{K} gilt: $1 + 1 \neq 0$. Dann gilt: eine affin abgeschlossene Teilmenge M des affinen Raums \mathcal{A} (über \mathbb{K} -Vektorraum V) ist ein (affiner) Unterraum.

Beweis. Angenommen, M ist affin abgeschlossen und $a_0 \in M$.

Z.z.: $U := \{\overrightarrow{a_0 a} \text{ wobei } a \in M\}$ ein Unterraum ist. (Weil

$M := \{a_0 + u \text{ wobei } u \in U\}$).

Z.z.: (i) Ist $v \in U$, so ist $\lambda v \in U$. (ii) Sind $v, w \in U$, so ist $v + w \in U$.

Wir zeigen (i). Es gilt: $\vec{0} = \overrightarrow{a_0 a_0} \in U$. Sei $v \in U$, $v \neq \vec{0}$, d.h. $v = \overrightarrow{a_0 a}$ für $a \in M$, $a \neq a_0$. Dann liegen alle Punkte der Geraden

$\mathcal{G}_{a_0, a} := \{a_0 + \lambda \overrightarrow{a_0 a} \text{ wobei } \lambda \in \mathbb{R}\}$ in U , also $\lambda \overrightarrow{a_0 a} \in U$.

Wir zeigen (ii). Seien $v = \overrightarrow{a_0 a} \in U$, $w = \overrightarrow{a_0 b} \in U$ mit $a, b \in M$. Gilt

$v = w$, so folgt nach (i) $v + w = 2v \in U$. Wir können also annehmen,

dass $v \neq w$ und damit $a \neq b$. Nach Voraussetzungen gilt dann $\mathcal{G}_{a,b} \subseteq M$.

Wegen $\overrightarrow{ab} = \overrightarrow{aa_0} + \overrightarrow{a_0 b} = -\overrightarrow{a_0 a} + \overrightarrow{a_0 b} = w - v$ gilt dann

$\mathcal{G}_{a,b} = \{a + \lambda(w - v) \text{ wobei } \lambda \in \mathbb{R}\} = \{a_0 + (1 - \lambda)v + \lambda w \text{ wobei } \lambda \in \mathbb{R}\}$.

Für $\lambda = \frac{1}{2}$ erhalten wir $a_0 + (1 - \frac{1}{2})v + \frac{1}{2}w = a_0 + \frac{1}{2}w$

$= a_0 + \frac{1}{2}(v + w) \in M$, also $\frac{1}{2}(v + w) \in U$. Nach (i) folgt

$2 \cdot \frac{1}{2}(v + w) \in U$.

□

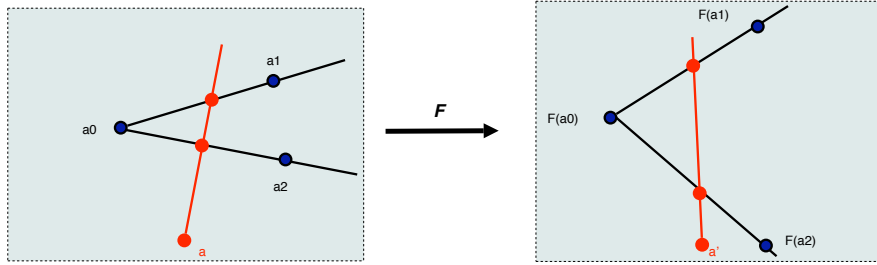
Satz 22 (Fundamentalsatz der affinen Geometrie (über \mathbb{R})) Seien $\mathcal{A}, \mathcal{A}_0$ affine Räume über den \mathbb{R} -Vektorräumen V, V_0 . Sei $\dim(\mathcal{A}) \geq 2$ und $F : \mathcal{A} \rightarrow \mathcal{A}_0$ eine Bijektion, die Geraden auf Geraden abbildet. Dann ist F eine Affinität.

Der nicht ganz einfache Beweis wird durch eine Reihe von Hilfslemmata (1–5) und einem Lemma (Lemma 11), das auch für Zukunft wichtig wird, erledigt. Wir nehmen stets an, dass F wie im Satz ist, also

- ▶ bijektiv ist, und
- ▶ bildet Gerade auf Geraden ab .

HilfsLemma 1. Sei $\mathcal{E} \subseteq \mathcal{A}$ eine affine Ebene. Dann ist $Bild_F(\mathcal{E}) \subseteq \mathcal{A}_0$ auch eine affine Ebene.

Beweis. Sei (a_0, a_1, a_2) ein Koordinatensystem für \mathcal{E} . Dann ist $\mathcal{G}_{a_0, a_1} \cap \mathcal{G}_{a_0, a_2} = \{a_0\}$.



Da F injektiv ist und $Bild_F(\mathcal{G}_{a_0, a_1}) = \mathcal{G}_{F(a_0), F(a_1)}$,
 $Bild_F(\mathcal{G}_{a_0, a_2}) = \mathcal{G}_{F(a_0), F(a_2)}$, folgt $\mathcal{G}_{F(a_0), F(a_1)} \cap \mathcal{G}_{F(a_0), F(a_2)} = \{F(a_0)\}$.

Deshalb existiert genau eine Ebene $\mathcal{E}_0 \subseteq \mathcal{A}_0$, die $\mathcal{G}_{F(a_0), F(a_1)}$ und $\mathcal{G}_{F(a_0), F(a_2)}$ enthält. Wir zeigen: $\mathcal{E}_0 \supseteq Bild_F(\mathcal{E})$.

Ist $a \in \mathcal{E} \setminus \mathcal{G}_{a_0, a_1} \cap \mathcal{G}_{a_0, a_2}$, so existiert eine Gerade $\mathcal{G} \subseteq \mathcal{E}$ mit $a \in \mathcal{G}$, die \mathcal{G}_{a_0, a_1} und \mathcal{G}_{a_0, a_2} schneidet.

Da F Geraden auf Geraden abbildet, folgt $F(a) \in Bild_F(\mathcal{E}_0)$. Da die Gerade $Bild_F(\mathcal{G})$ zwei Punkte der Ebene \mathcal{E} enthält, liegt sie auf der Ebene; deswegen liegt der Punkt $F(a) \in \mathcal{E}_0$.

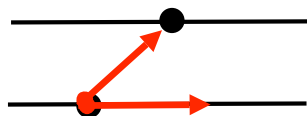
Damit haben wir $Bild_F(\mathcal{E}) \subseteq \mathcal{E}_0$ bewiesen. Da die Umkehrabbildung F^{-1} auch eine Bijektion ist, die Gerade auf Geraden abbildet, gilt $\mathcal{E}_0 \subseteq Bild_F(\mathcal{E})$.



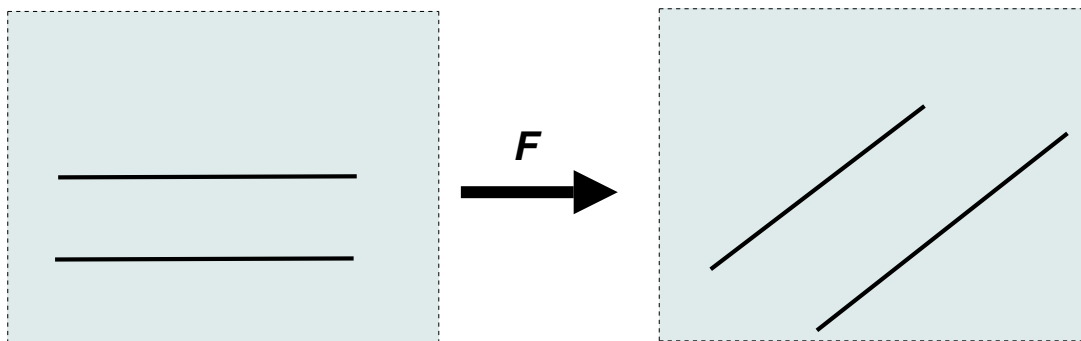
HilfsLemma 2. F bildet parallele Geraden auf parallele Geraden ab.

Im Beweis benutzen wir Lemma 10 : *in jeder Ebene \mathcal{E} sind zwei Geraden ohne Schnittpunkte parallel.*

Beweis. Seien $\mathcal{G}_1 \neq \mathcal{G}_2$ parallele Geraden in \mathcal{A} . Dann existiert eine Ebene $\mathcal{E} \subseteq \mathcal{A}$ mit $\mathcal{G}_i \subseteq \mathcal{E}$.



Nach HilfsLemma 1 ist dann $Bild_F(\mathcal{E})$ auch eine Ebene $\mathcal{E}_0 \subseteq \mathcal{A}_0$. Da F Gerade auf Geraden überführt, sind die Bilden von \mathcal{G}_i Geraden in \mathcal{E}_0 .



Da F bijektiv und deswegen injektiv ist, haben sie keine Schnittpunkte und sind deswegen nach Lemma 10 parallel,



Plan des Beweises des Satzes 22

Um den Fundamentalsatz 22 zu beweisen, wählen wir einen festen Punkt $a_0 \in \mathcal{A}$, setzen $a'_0 := F(a_0) \in \mathcal{A}_0$, und definieren $G : V \rightarrow V_0$ durch: Für alle $v \in V$ gelte

$$F(a_0 + v) = a'_0 + G(v)$$

(oder, was dasselbe ist:

$$G(\overrightarrow{a_0 a}) := \overrightarrow{a'_0 F(a)}).$$

Ziel: Wir werden beweisen, dass G eine lineare Abbildung ist; d.h., dass $G(v + w) = G(v) + G(w)$ **Wir werden zwei Fälle betrachten: v und w sind linear unabhängig (Hilfslemma 3), und $v = \lambda w$ (Hilfslemma 4).**

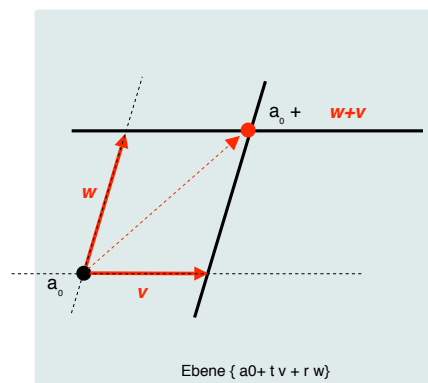
und $G(\lambda v) = \lambda G(v)$.

Dann wird F eine affine Abbildung nach Def. 13

HilfsLemma 3. Seien $v, w \in \mathbb{V}$ linear unabhängig. Dann gilt:

$$G(v + w) = G(v) + G(w)$$

Beweis. Wir betrachten die Geraden



$$\mathcal{G}_v = \{ a_0 + t v \mid t \in \mathbb{R} \},$$

$$\mathcal{G}_w = \{ a_0 + t w \mid t \in \mathbb{R} \},$$

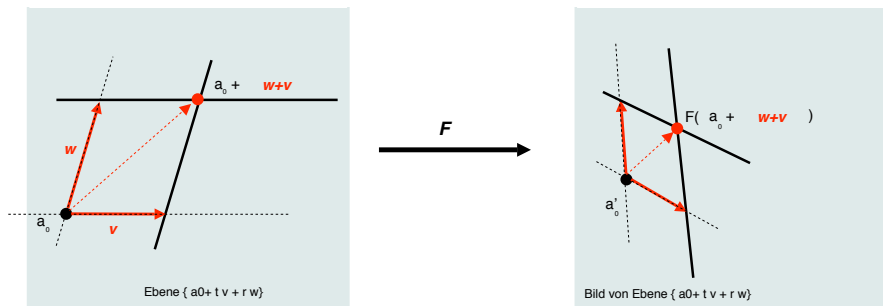
die einander im Punkt a_0 schneiden, und die Geraden

$$\mathcal{G}'_v = \{ a_0 + w + t v \mid t \in \mathbb{R} \}, \text{ und}$$

$$\mathcal{G}'_w = \{ a_0 + v + t w \mid t \in \mathbb{R} \}$$

die einander im Punkt $a_0 + v + w$ schneiden

Wir werden jetzt die Bilder von Geraden betrachten



Da \mathcal{G}'_v parallel zu $\mathcal{G}_v := \{a_0 + tv \mid t \in \mathbb{R}\}$ ist, ist nach HilfsLemma 2 auch $Bild_F(\mathcal{G}'_v)$ parallel zu $Bild_F(\mathcal{G}_v)$. Wegen

$F(a_0 + w) = a'_0 + G(w) \in Bild_F(\mathcal{G}'_v)$ und $a'_0 \in Bild_F(\mathcal{G}_v)$,

$F(a_0 + v) = a'_0 + G(v) \in Bild_F(\mathcal{G}_v)$ gilt

$Bild_F(\mathcal{G}'_v) = \{a'_0 + G(w) + tG(v) \mid t \in \mathbb{R}\}$ und analog

$Bild_F(\mathcal{G}'_w) = \{a'_0 + G(v) + tG(w) \mid t \in \mathbb{R}\}$

Da F injektiv ist und \mathcal{G}_v und \mathcal{G}_w einander genau in a_0 schneiden, gilt

$Bild_F(\mathcal{G}'_v) \cap Bild_F(\mathcal{G}'_w) = \{a'_0\}$. Deshalb sind $G(v) = F(a_0)F(a_0 + v)$ und

$G(w) = F(a_0)F(a_0 + w)$ linear unabhängig. Da $G(v)$ bzw. $G(w)$

Richtungsvektoren von $Bild_F(\mathcal{G}'_v)$ bzw. $Bild_F(\mathcal{G}'_w)$ sind, schneiden

$Bild_F(\mathcal{G}'_v)$ und $Bild_F(\mathcal{G}'_w)$ einander genau im Punkt

$F(a_0 + v + w) = a'_0 + G(v + w)$, dem Bild des Schnittpunktes von \mathcal{G}'_v

und \mathcal{G}'_w . Andererseits zeigen die obigen Formeln für $Bild_F(\mathcal{G}'_v)$ und

$Bild_F(\mathcal{G}'_w)$, daß auch $a'_0 + G(v) + G(w) \in Bild_F(\mathcal{G}'_v) \cap Bild_F(\mathcal{G}'_w)$ ist.

Daraus folgt $G(v + w) = G(v) + G(w)$,



- ▶ Probeklausur findet (voraussichtlich) am 16.06 von 8-10 Uhr in in HS1, HS4 Abbeanum statt.
- ▶ Alle sind zugelassen; keine Anmeldung. Die Teilnahme ist freiwillig. Bis zum 20% von Hausaufgabenpunkte. Die Erfahrung von letzten Jahren zeigt, dass es sich lohnt, teilzunehmen.
- ▶ Etwa 5 Aufgaben; davon 3 theoretische; eine davon wird bestehen, einen wichtigen Satz aus der Vorlesung zu beweise (die Liste von wichtigen Sätzen wird eine Woche vor der Klausur bekannt gegeben.)
- ▶ Keine Hilfsmittel zugelassen; Papier wird gegeben.
- ▶ Am Dienstag 17.06 werden die Lösungen besprochen.
- ▶ Sie bekommen Ihre Lösungen wieder, nachdem wir sie korrigieren.

Ziel: Fundamentalsatz der reellen affinen Geometrie

Satz 22 (Fundamentalsatz der affinen Geometrie (über \mathbb{R})) Seien $\mathcal{A}, \mathcal{A}_0$ affine Räume über den \mathbb{R} -Vektorräumen V, V_0 . Sei $\dim(\mathcal{A}) \geq 2$ und $F : \mathcal{A} \rightarrow \mathcal{A}_0$ eine Bijektion, die Geraden auf Geraden abbildet. Dann ist F eine Affinität.

Um den Fundamentalsatz 22 zu beweisen, wählen wir einen festen Punkt $\mathbf{a}_0 \in \mathcal{A}$, setzen $\mathbf{a}'_0 := F(\mathbf{a}_0) \in \mathcal{A}_0$, und definieren $G : V \rightarrow V_0$ durch: Für alle $\mathbf{v} \in V$ gelte

$$F(\mathbf{a}_0 + \mathbf{v}) = \mathbf{a}'_0 + G(\mathbf{v})$$

(oder, was dasselbe ist:

$$G(\overrightarrow{\mathbf{a}_0 \mathbf{a}}) := \overrightarrow{\mathbf{a}'_0 F(\mathbf{a})}.$$

Ziel: Wir werden beweisen, dass G eine lineare Abbildung ist; d.h., dass $G(\mathbf{v} + \mathbf{w}) = G(\mathbf{v}) + G(\mathbf{w})$. Den Fall \mathbf{v} und \mathbf{w} sind linear unabhängig haben wir gestern gemacht (HilfsLemma 3), und $\mathbf{v} = \lambda \mathbf{w}$ (HilfsLemma 4).

und $G(\lambda \mathbf{v}) = \lambda G(\mathbf{v})$.

Dann wird F eine affine Abbildung nach Def. 13

HilfsLemma 4. Für alle $v \in V$ und alle $\lambda, \mu \in \mathbb{R}$ gilt:

$$G(\lambda v) + G(\mu v) = G((\lambda + \mu)v).$$

Beweis. Wegen Injektivität von F gilt $G(\vec{0}) = \vec{0}$; deswegen genügt es, die Fälle zu betrachten, in denen $v \neq \vec{0}$, $\lambda \neq 0$ und $\mu \neq 0$ gilt. Aufgrund unserer Voraussetzung $\dim(\mathcal{A}) \geq 2$ existiert ein zu v linear unabhängiger Vektor $w \in V$.

1. Fall: $\lambda + \mu \neq 0$

Dann sind w und $(\lambda + \mu)v$ linear unabhängig und HilfsLemma 3 impliziert

$$G(w + (\lambda + \mu)v) = G(w) + G((\lambda + \mu)v).$$

Da $w + \lambda v$ und μv linear unabhängig sind, gilt nach HilfsLemma 3

$$G(w + (\lambda + \mu)v) = G(w + \lambda v) + G(\mu v) \stackrel{\text{analog}}{=} G(w) + G(\lambda v) + G(\mu v).$$

Dann gilt

$$G(w) + G((\lambda + \mu)v) = G(w) + G(\lambda v) + G(\mu v).$$

Die letzte Gleichung impliziert die Gleichung

$$G(\lambda v) + G(\mu v) = G((\lambda + \mu)v).$$

Bemerkung. Wir haben Voraussetzung $\dim(\mathcal{A}) \geq 2$ benutzt. Ohne diese Voraussetzung ist der Satz falsch – Hausaufgabe.

2. Fall: $\lambda = -\mu$

Da $\lambda v + w$ und $-\lambda v + w$ linear unabhängig sind, folgt aus HilfsLemma 3

$$\begin{aligned} G(2w) &= G((\lambda v + w) + (-\lambda v + w)) &= G(\lambda v + w) + G(-\lambda v + w) \\ & &= G(\lambda v) + G(w) + G(-\lambda v) + G(w) \\ & &= 2G(w) + G(\lambda v) + G(-\lambda v). \end{aligned}$$

Nach dem 1. Fall für $\lambda = \mu := 1$ gilt $G(2w) = 2G(w)$; daraus folgt die Behauptung $G(\lambda v) + G(-\lambda v) = 0$.

Ziel: $G(\lambda \cdot v) = \lambda \cdot G(v)$

HilfsLemma 5. Es existiert eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, so daß für alle $v \in V$ und alle $\lambda \in \mathbb{R}$ gilt: $G(\lambda v) = f(\lambda)G(v)$.

Beweis. Wegen $G(\vec{0}) = \vec{0}$ ist die Gleichung für $v = \vec{0}$ stets erfüllt, und sie ist für $\lambda = 0$ und alle $v \in V$ erfüllt, wenn wir $f(0) = 0$ setzen. Wir werden von jetzt $v \neq \vec{0}$ und $\lambda \neq 0$ annehmen.

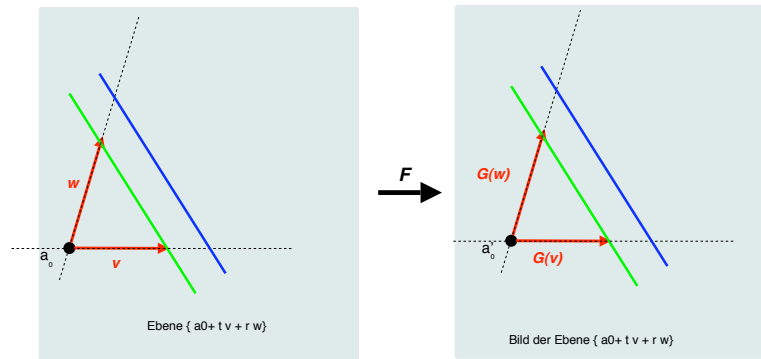
Da F injektiv ist, gilt $F(a_0 + v) = a'_0 + G(v) \neq a'_0$. Nach Voraussetzungen liegt $F(a_0 + \lambda v)$ für alle $\lambda \in \mathbb{R}$ auf der Geraden $\mathcal{G}_{a'_0, a'_0 + G(v)}$. Also existiert (genau ein) $f(\lambda, v) \in \mathbb{R}$, so daß $F(a_0 + \lambda v) = a'_0 + f(\lambda, v)G(v)$ gilt.

Wir bemerken, daß $TV(a'_0, F(a_0 + \lambda v), F(a_0 + v)) = \frac{1}{f(\lambda, v)}$ gilt. Um HilfsLemma 5 zu beweisen, müssen wir zeigen, daß $f(\lambda, v)$ unabhängig von v ist, d.h., daß für alle $\lambda \neq 0, v \neq \vec{0}, w \neq \vec{0}$ gilt:

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

1. Fall: v und w sind linear unabhängig.

Dann betrachten wir die Geraden $H := \{a_0 + tv \mid t \in \mathbb{R}\}$ und $J := \{a_0 + tw \mid t \in \mathbb{R}\}$ durch a_0



und die Geraden $G_1 := G_{a_0+v, a_0+w}$, $G_2 := G_{a_0+\lambda v, a_0+\lambda w}$, die nach Strahlensatz 20 parallel sind. Nach HilfsLemma 2 sind dann $Bild_F(G_1)$ und $Bild_F(G_2)$ auch parallel. Die andere Richtung des Strahlensatzes impliziert nun, dass

$$TV(a'_0, F(a_0 + \lambda v), F(a_0 + v)) = TV(a'_0, F(a_0 + \lambda w), F(a_0 + w))$$

gilt. Daraus folgt $\frac{1}{1f(\lambda, v)} = \frac{1}{f(\lambda, w)}$. Dann gilt $f(\lambda, v) = f(\lambda, w)$.

2. Fall: $w = \mu v$.

In dem Fall wählen wir ein von v linear unabhängiges $z \in V$. Dann gilt mit zweifacher Anwendung des 1. Falls:

$$f(\lambda, v) = f(\lambda, z) = f(\lambda, w).$$

HilfsLemma 5 ist bewiesen.

Folgerung. $f : \mathbb{R} \rightarrow \mathbb{R}$ ist ein Körperautomorphismus, d.h., es gilt

- (i) $f(1) = 1$, und für alle $\lambda, \mu \in \mathbb{R}$
- (ii) $f(\lambda) + f(\mu) = f(\lambda + \mu)$,
- (iii) $f(\lambda \cdot \mu) = f(\lambda) \cdot f(\mu)$.

Beweis. Wähle $v \in V \setminus \{\vec{0}\}$. Da F injektiv ist, folgt
 $G(v) = \overrightarrow{F(a_0)F(a_0 + v)} \neq 0$.

- (i) $G(v) = G(1 \cdot v) = f(1) \cdot G(v) \implies (f(1) - 1)G(v) = \vec{0} \implies f(1) = 1$.
- (ii) $G((\lambda + \mu)v) = f(\lambda + \mu)G(v)$
 $G((\lambda + \mu)v) = G(\lambda v + \mu v) \stackrel{HL 4}{=} G(\lambda v) + G(\mu v) = (f(\lambda) + f(\mu))G(v)$.
Diese Gleichungen beweisen $f(\lambda + \mu) = f(\lambda) + f(\mu)$.
- (iii) $G((\lambda\mu)v) = f(\lambda\mu)G(v)$
 $G((\lambda\mu)v) = G(\lambda(\mu v)) = f(\lambda)G(\mu v) = f(\lambda)f(\mu)G(v)$.
Diese Gleichungen beweisen $f(\lambda \cdot \mu) = f(\lambda) \cdot f(\mu)$, □

Bemerkung. Nach dem Beweis von Satz 22 werden wir die Definition von Körperautomorphismus nochmal ansehen.

Bisher wurde noch nicht benutzt, daß \mathcal{A} , \mathcal{A}_0 **reelle** affine Räume sind.
(Wir haben nur benutzt, dass $1 + 1 \neq 2$ ist, weil wir den Satz 21 benutzt haben.)

Das folgende Lemma aber wäre etwa für den Körper \mathbb{C} (statt \mathbb{R}) nicht wahr – sieh Hausaufgabe 1.

Lemma 11. Die Identität $f = Id$ ist der einzige Körperautomorphismus von \mathbb{R} .

Lemma 11. Die Identität $f = Id$ ist der einzige Körperautomorphismus von \mathbb{R} .

Beweis. Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ ein Körperautomorphismus. Die Additivität von f und $f(1) = 1$ zeigen, daß $f(2) = f(1 + 1) = f(1) + f(1) = 1 + 1 = 2$ und – induktiv – daß $f(n) = n$ für alle $n \in \mathbb{N}$ gilt. Aus $f(0) + f(0) = f(0 + 0) = f(0)$ folgt $f(0) = 0$ und aus $f(n + (-n)) = 0 = f(n) + f(-n) = n + f(-n)$ für $n \in \mathbb{N}$ folgt $f(n) = n$ für alle $n \in \mathbb{Z}$. Ist $n \in \mathbb{Z} \setminus \{0\}$, so gilt

$$1 = f(1) = f\left(n \cdot \frac{1}{n}\right) = f(n)f\left(\frac{1}{n}\right) = nf\left(\frac{1}{n}\right) \implies f\left(\frac{1}{n}\right) = \frac{1}{n}.$$

Verwenden wir nochmals die Multiplikativität von f , so erhalten wir

$f\left(\frac{p}{q}\right) = \frac{p}{q}$ für alle $p, q \in \mathbb{Q}$. Weiter gilt $\text{Bild}_f(\mathbb{R}_{\geq 0}) \subseteq \mathbb{R}_{\geq 0}$, denn aus

$r \geq 0$ folgt $f(r) = f(\sqrt{r^2}) = f(\sqrt{r})^2 \geq 0$.

Zusammen mit der Additivität von f ergibt das $r \geq s \implies f(r) \geq f(s)$,

d.h. f ist monoton wachsend. Ist schließlich $r \in \mathbb{R}$ beliebig, so wählen wir

$x_i, y_i \in \mathbb{Q}$ mit $\lim x_i = r = \lim y_i$ und $x_i < r < y_i$. Dann $f(r) = r$ aus

$x_i = f(x_i) \leq f(r) \leq f(y_i) = y_i$ und $\lim x_i = r = \lim y_i$, □

Zusammenfassung des Beweises des Satzes 22

Satz 22 Seien $\mathcal{A}, \mathcal{A}_0$ affine Räume über den \mathbb{R} -Vektorräumen V, V_0 . Sei $\dim(\mathcal{A}) \geq 2$ und $F : \mathcal{A} \rightarrow \mathcal{A}_0$ eine Bijektion, die Geraden auf Geraden abbildet. Dann ist F eine Affinität.

Wir wählen einen festen Punkt $a_0 \in \mathcal{A}$, setzen $a'_0 := F(a_0) \in \mathcal{A}_0$, und definieren $G : V \rightarrow V_0$ durch: Für alle $v \in V$ gelte

$$F(a_0 + v) = a'_0 + G(v)$$

Wir zeigen, dass G eine lineare Abbildung ist:

HL 3,4: $G(v + w) = G(v) + G(w)$.

HL 5: $G(\lambda v) = f(\lambda)G(v)$, wobei $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ein Körperautomorphismus ist.

Lemma 11: f ist Identität (d.h., $G(\lambda v) = \lambda G(v)$.)

Dann ist G linear, und deswegen F eine Affinität.

Körperautomorphismus $f : \mathbb{K} \rightarrow \mathbb{K}$

In der Folgerung aus HilfsLemma 5 haben wir einen Körperautomorphismus durch die folgende Bedingungen definiert:

- (i) $f(1) = 1$, und für alle $\lambda, \mu \in \mathbb{K}$
- (ii) $f(\lambda) + f(\mu) = f(\lambda + \mu)$,
- (iii) $f(\lambda \cdot \mu) = f(\lambda) \cdot f(\mu)$.

Wir zeigen, dass jedes f mit diesen Eigenschaften ein Körperisomorphismus auf dem Bild von f ist: nach Definition (siehe Vorl. 1) ist f ein **Körperisomorphismus**, falls $f : \mathbb{K} \rightarrow \text{Bild}_f(\mathbb{K})$ bijektiv ist, und Verknüpfungen erhält: $f(\lambda) + f(\mu) = f(\lambda + \mu)$, $f(\lambda \cdot \mu) = f(\lambda) \cdot f(\mu)$.

Also, um zu zeigen dass Körperautomorphismus ein Körperisomorphismus ist, müssen wir nur zeigen, dass f mit Eigenschaften (i), (ii), (iii) bijektiv ist (als Abbildung $f : \mathbb{K} \rightarrow \text{Bild}_f(\mathbb{K})$)

Wir zeigen zuerst, dass $f(\lambda) = 0 \iff \lambda = 0$.

In der Tat, sonst liefert $1 = f(1) = f(\lambda \cdot \lambda^{-1}) = \underbrace{f(\lambda)}_{=0} \cdot f(\lambda^{-1}) = 0$

ein Widerspruch.

Wir können jetzt beweisen, dass die Abbildung f bijektiv ist.

Injektivität: angenommen, $f(\lambda) = f(\mu)$. Dann ist

$0 = f(\lambda) - f(\mu) \stackrel{(ii)}{=} f(\lambda - \mu)$; dann ist $\lambda - \mu = 0$.

Subjektivität ist automatisch erfüllt, da der Bildmenge von $f : \mathbb{K} \rightarrow \text{Bild}_f(\mathbb{K})$ das ganze $\text{Bild}_f(\mathbb{K})$ ist.

Schwerpunkt

Sei x_1, \dots, x_k die Punkte eines affinen Raums \mathcal{A} über \mathbb{R} , und $m_1, \dots, m_k \in \mathbb{R}_{\geq 0}$ so dass mind. ein $m_i > 0$.

Def. Der Schwerpunkt von x_1, \dots, x_k mit Massen m_1, \dots, m_k ist der Punkt $a + \frac{1}{\sum_{i=1}^k m_i} \sum_{i=1}^k m_i \vec{ax}_i$.

Lemma 12 Schwerpunkt hängt nicht von der Wahl von Punkt a ab.

Beweis.

$$\begin{aligned} b + \frac{1}{\sum_{i=1}^k m_i} \sum_{i=1}^k m_i \vec{bx}_i &= a + \vec{ab} + \frac{1}{\sum_{i=1}^k m_i} \sum_{i=1}^k m_i \vec{bx}_i \\ &= a + \frac{1}{\sum_{i=1}^k m_i} \left(\sum_{i=1}^k m_i \vec{ab} + \sum_{i=1}^k m_i \vec{bx}_i \right) \\ &= a + \sum_{i=1}^k m_i (\vec{ab} + \vec{bx}_i) \\ &= a + \sum_{i=1}^k m_i \vec{ax}_i, \end{aligned}$$

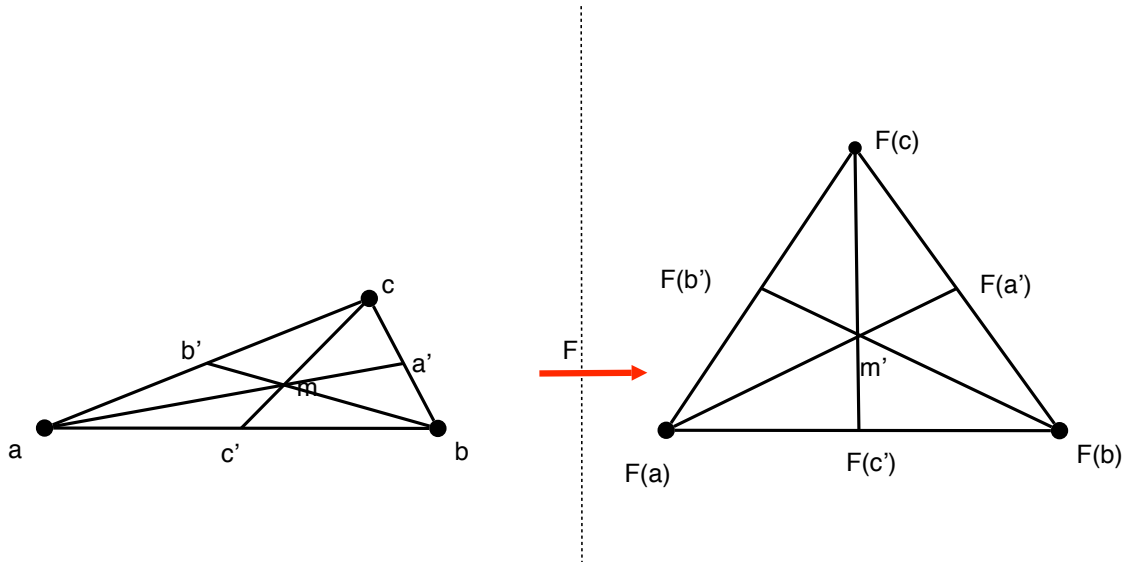
Bemerkung. Eigenschaft „Schwerpunkt zu sein“ ist eine affine Eigenschaft: ist $F : \mathcal{A} \rightarrow \mathcal{A}_0$ eine affine Abbildung, dann ist Schwerpunkt von $F(x_1), \dots, F(x_k)$ (mit Massen m_1, \dots, m_k) ist Bild des Schwerpunktes von x_1, \dots, x_k (mit denselben Massen m_1, \dots, m_k)

$$\begin{aligned}
 & \text{In der Tat, } F(a) + \frac{1}{\sum_{i=1}^k m_i} \sum_{i=1}^k m_i \overrightarrow{F(a)F(x_i)} \\
 &= F(a) + \frac{1}{\sum_{i=1}^k m_i} \sum_{i=1}^k m_i \overrightarrow{ax_i} \\
 &= F \left(\underbrace{a + \frac{1}{\sum_{i=1}^k m_i} \sum_{i=1}^k m_i \overrightarrow{ax_i}}_{\text{Schwerpunkt von } x_1, \dots, x_k} \right)
 \end{aligned}$$

Schwerpunkt eines Dreiecks

Folgerung. Man betrachte die Punkte a, b, c eines Dreiecks mit Massen $m_a = 1, m_b = 1, m_c = 1$. Dann ist der Schwerpunkt der Schnittpunkt der Seitenhalbierenden.

Beweis. Man betrachte eine Affinität F , die die Ecke des Dreiecks abc in Ecken eines regelmäßigen Dreiecks überführt



Lemma 13. Seien x_1, \dots, x_k und y_1, \dots, y_r Punkte; $m_{x_1}, \dots, m_{x_k}, m_{y_1}, \dots, m_{y_r}$ die Massen. Schwerpunkt von x_1, \dots, x_k sei S_x ; Schwerpunkt von y_1, \dots, y_r sei S_y . Dann gilt: Schwerpunkt vom Paar von Punkten S_x und S_y mit Massen $\sum_{i=1}^k m_{x_i}$ und $\sum_{i=1}^r m_{y_i}$ ist gleich Schwerpunkt vom $x_1, \dots, x_k, y_1, \dots, y_r$ mit Massen $m_{x_1}, \dots, m_{x_k}, m_{y_1}, \dots, m_{y_r}$.

Beweis. Wir berechnen die beide Schwerpunkte, und stellen fest, dass sie zusammenfallen. Nach Lemma 12 können wir einen beliebigen Punkt als den Punkt a in der Formeln

$$a + \frac{1}{\sum_{i=1}^k m_{x_i} + \sum_{j=1}^r m_{y_j}} \left(\sum_{j=1}^k m_{x_i} \overrightarrow{ax_i} + \sum_{j=1}^r m_{y_j} \overrightarrow{ay_j} \right) (*)$$

wählen; wir nehmen $a = S_x$. Da S_x Schwerpunkt von x_1, \dots, x_k ist, ist $S_x + \frac{1}{\sum_{i=1}^k m_{x_i}} \sum_{j=1}^k m_{x_i} \overrightarrow{S_x x_i} = S_x$ folglich $\sum_{j=1}^k m_{x_i} \overrightarrow{S_x x_i} = \vec{0}$.
Dann ist (*) gleich

$$S_x + \frac{1}{\sum_{i=1}^k m_{x_i} + \sum_{j=1}^r m_{y_j}} \left(\sum_{j=1}^r m_{y_j} \overrightarrow{S_x y_j} \right).$$

Wenn wir Schwerpunkt von y_1, \dots, y_r berechnen bekommen wir

$$S_y = S_x + \frac{1}{\sum_{j=1}^r m_{y_j}} \sum_{j=1}^r m_{y_j} \overrightarrow{S_x y_j}, \quad \text{also} \quad \overrightarrow{S_y S_x} = \frac{1}{\sum_{j=1}^r m_{y_j}} \sum_{j=1}^r m_{y_j} \overrightarrow{S_x y_j},$$

und deswegen

$$\frac{1}{\sum_{i=1}^k m_{x_i} + \sum_{j=1}^r m_{y_j}} \left(\sum_{j=1}^r m_{y_j} \overrightarrow{S_x y_j} \right) = \frac{\sum_{j=1}^r m_{y_j}}{\sum_{i=1}^k m_{x_i} + \sum_{j=1}^r m_{y_j}} \overrightarrow{S_y S_x}.$$

Dann ist die Formel (*) gleich

$$S_x + \frac{\sum_{i=1}^k m_{x_i}}{\sum_{i=1}^k m_{x_i} + \sum_{j=1}^r m_{y_j}} \vec{0} + \frac{\sum_{j=1}^r m_{y_j}}{\sum_{i=1}^k m_{x_i} + \sum_{j=1}^r m_{y_j}} \overrightarrow{S_y S_x},$$

und dies ist die Formel für den Schwerpunkt von zwei Punkte S_x, S_y mit Massen $\sum_{i=1}^k m_{x_i}, \sum_{j=1}^r m_{y_j}$, □

Neues Thema: Konvexe Geometrie

Konvexität ist ein Begriff, der in verschiedenen Zusammenhängen eine Rolle spielt und sich als nützlich erwiesen hat.

Wir werden im n-dimensionalen **reellen** affinen Raum arbeiten; nach Hauptsatz der affiner Geometrie kann man annehmen, dass wir in \mathbb{R}^n (über $(\mathbb{R}^n, +, \bullet)$) arbeiten.

In dem Fall gilt

$$\vec{xy} = \underbrace{y - x}_{\text{als Elemente von } (\mathbb{R}^n, +, \bullet)} .$$

Manchmal werden wir auch annehmen, dass der Vektorraum $(\mathbb{R}^n, +, \bullet)$ Euklidisch ist, d.h., wir auf $(\mathbb{R}^n, +, \bullet)$ ein Skalarprodukt \langle , \rangle gewählt haben.

Affine Hülle

Wiederholung. Der Vektor $x \in \mathbb{R}^n$ ist eine lineare Kombination der Vektoren $x_1, \dots, x_k \in \mathbb{R}^n$, wenn es Zahlen $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ gibt mit $x = \lambda_1 x_1 + \dots + \lambda_k x_k$.

Def. 21 Gibt es solche Zahlen $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ mit $\lambda_1 + \dots + \lambda_k = 1$, so ist der Punkt x eine **affine Kombination** der Punkte x_1, \dots, x_k .

Bemerkung. Dies ist äquivalent zu „ $\overrightarrow{x_1 x}$ ist lineare Kombination der Vektoren $\overrightarrow{x_1 x_2} = x_2 - x_1, \overrightarrow{x_1 x_3} = x_3 - x_1, \dots, \overrightarrow{x_1 x_k} = x_k - x_1$.“

Tatsächlich, in diesem Fall ist $\lambda_1 = 1 - \lambda_2 - \lambda_3 - \dots - \lambda_k$, und

$$\begin{aligned}x &= \lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_k x_k \\&= x_1 + (-\lambda_2 - \dots - \lambda_k) x_1 + \lambda_2 x_2 + \dots + \lambda_k x_k \\&= x_1 + \lambda_2 (x_2 - x_1) + \lambda_3 (x_3 - x_1) + \dots + \lambda_k (x_k - x_1)\end{aligned}$$

Wiederholung. Für $A \subseteq \mathbb{R}^n$ ist die **lineare Hülle** $\text{Span}(A)$ die Menge aller linearen Kombinationen von Elementen von A . $\text{Span}(A)$ ist zugleich der kleinste lineare Untervektorraum von \mathbb{R}^n , der A enthält, und die Schnittmenge von allen linearen Untervektorräumen von \mathbb{R}^n , die A enthalten.

Def. 22 Für $A \subseteq \mathbb{R}^n$ ist die affine Hülle $\text{aff}(A)$ die Menge aller affinen Kombinationen von Elementen von A .

Lemma 14 $\text{aff}(A)$ ist der kleinste affine Unterraum von \mathbb{R}^n , der A enthält (d.h., gilt $A \subseteq \mathcal{U}$ für einen affinen Unterraum $\mathcal{U} \subseteq \mathbb{R}^n$, so gilt $\text{aff}(A) \subseteq \mathcal{U}$).

Bemerkung. Daraus folgt insbesondere, dass der kleinste affine Unterraum von \mathbb{R}^n , der A enthält, existiert. Sie haben es auch in Hausaufgabe 2 bewiesen: Man nehme Schnitt aller affinen Räumen, die A enthalten. Wie in Lemma 7 kann man beweisen, dass der Schnitt ein affiner Raum ist; nach Konstruktion enthält er alle Punkte von A , und ist in jedem affinen Raum, der A enthält, enthalten.

Lemma 14 $\text{aff}(A)$ ist der kleinste affine Unterraum von \mathbb{R}^n , der A enthält (d.h., gilt $A \subseteq \mathcal{U}$ für einen affinen Unterraum $\mathcal{U} \subseteq \mathbb{R}^n$, so gilt $\text{aff}(A) \subseteq \mathcal{U}$).

Beweis. Z.z.: (i) $\text{aff}(A) \subseteq \mathcal{U}$ und (ii) $\text{aff}(A)$ ist ein affiner Raum.

(i): Liegt $x \in \text{aff}(A)$, so ist $\overrightarrow{x_1 x}$ nach Bemerkung nach der Def. 21 lineare Kombination der Vektoren

$\overrightarrow{x_1 x_2} = x_2 - x_1, \overrightarrow{x_1 x_3} = x_3 - x_1, \dots, \overrightarrow{x_1 x_k} = x_k - x_1$. Da diese Vektoren in $V_{\mathcal{U}}$ liegen, liegt auch deren Linearkombination in $V_{\mathcal{U}}$, folglich auch $x_1 + \overrightarrow{x_1 x} = x$ in \mathcal{U} .

Lemma 14 $\text{aff}(A)$ ist der kleinste affine Unterraum von \mathbb{R}^n , der A enthält (d.h., gilt $A \subseteq \mathcal{U}$ für einen affinen Unterraum $\mathcal{U} \subseteq \mathbb{R}^n$, so gilt $\text{aff}(A) \subseteq \mathcal{U}$).

Beweis von (ii): $\text{aff}(A)$ ist ein affiner Raum.

Wir zeigen, dass $\text{aff}(A)$ affine abgeschlossen ist; dann muß es nach Satz 21 ein affiner Raum sein.

Angenommen $x, y \in \text{aff}(A)$. Dann sind $x = \lambda_1 x_1 + \dots + \lambda_k x_k$ und $y = \mu_1 x_1 + \dots + \mu_k x_k$ für irgendwelche $\lambda_1 + \dots + \lambda_k = 1$, $\mu_1 + \dots + \mu_k = 1$. Wir betrachten die Gerade

$$\mathcal{G}_{x,y} = \{x + t \cdot \vec{x\bar{y}} \mid t \in \mathbb{R}\}$$

$$= \{\lambda_1 x_1 + \dots + \lambda_k x_k + t \cdot ((\mu_1 - \lambda_1)x_1 + \dots + (\mu_k - \lambda_k)x_k) \mid t \in \mathbb{R}\}.$$

Wir sehen, dass jeder Punkt der Geraden das Aussehen

$$\underbrace{(\lambda_1 + t \cdot (\mu_1 - \lambda_1))}_{\eta_1} x_1 + \dots + \underbrace{(\lambda_k + t \cdot (\mu_k - \lambda_k))}_{\eta_k} x_k$$

hat. Da $\sum \lambda_i = \sum \mu_i (= 1)$ ist, gilt $\sum \eta_i = \sum \lambda_i = 1$, und deswegen liegt jeder Punkt der Geraden in $\text{aff}(A)$. Dann ist $\text{aff}(A)$ affin abgeschlossen, □

Affin-unabhängige Punktenmenge

Def. Eine Teilmenge $A \subseteq \mathbb{R}^n$ heißt affin abhängig, falls es $k \in \mathbb{N}$ und $x_0, x_1, \dots, x_k \in A$ sodass $\{\overrightarrow{x_0x_1}, \overrightarrow{x_0x_2}, \dots, \overrightarrow{x_0x_k}\}$ linear unabhängig ist.

Nach Bemerkung nach Def. 21 die Menge ist genau dann affin abhängig, wenn es Zahlen $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ und die Punkte $x_1, \dots, x_k \in A$ gibt, so dass $\sum_{i=1}^k \alpha_i x_i = \vec{0}$ und $\sum_{i=1}^k \alpha_i = 0$ gilt.

Anzahl von Elemente in einer affin unabhängigen Menge $A \subseteq \mathbb{R}^n$ ist gleich $\dim(\text{aff}(A)) + 1$ und ist höchstens $n + 1$.

Def. 23 Eine Teilmenge $A \subseteq \mathbb{R}^n$ heißt konvex, wenn sie mit je zwei Punkten x, y auch stets deren **Verbindungsstrecke** $\{x + t \cdot \overrightarrow{xy} \mid 0 \leq t \leq 1\}$ enthält.

Bemerkung. Da $\overrightarrow{xy} = y - x$, gilt $x + t \cdot \overrightarrow{xy} = (1 - t)x + ty$. Die Menge A ist also genau dann konvex, wenn

$$\forall x, y \in A \quad \forall t \in [0, 1] \text{ gilt } (1 - t)x + ty \in A.$$

Bsp. Jeder Punkt, jede Strecke, jeder affine Unterraum ist eine konvexe Menge.

Bemerkung. \emptyset ist auch eine konvexe Menge.

Konvexe Hülle (von $A \subseteq \mathbb{R}^n$)

Def. 24(a) Die **konvexe Hülle** von A ist die Menge

$$\text{conv}(A) := \{\lambda_1 x_1 + \dots + \lambda_k x_k \mid \lambda_1 + \dots + \lambda_k = 1, \lambda_i \geq 0, x_i \in A\}.$$

(Der Punkt $\lambda_1 x_1 + \dots + \lambda_k x_k$, wobei $\sum_i \lambda_i = 1$, $\lambda_i \geq 0$ heißt **konvexe Kombination**. Konvexe Hülle von A ist die Menge aller konvexen Kombinationen der Elemente von A .)

Def. 24(b) Die **konvexe Hülle** von A ist die kleinste konvexe Menge $\text{conv}(A)$, die A enthält. (d.h., gilt $A \subseteq C$ für eine konvexe Menge C , so gilt $\text{conv}(A) \subseteq C$)

Def. 24(c) Die **konvexe Hülle** von A ist der Durchschnitt von allen konvexen Mengen, die A enthalten:

$$\text{conv}(A) = \bigcap_{\substack{C \supseteq A \\ C \text{ ist konvex}}} C.$$

Satz 23. Die Definitionen 24(a), (b), (c) sind äquivalent: ist eine Menge konvexe Hülle nach einer Definition, so ist sie auch die konvexe Hülle nach andere Definition.

Schema des Beweises: Zuerst $(c) \iff (b)$ und dann $(a) \iff (b)$.



(c) \iff (b)

Sei $\text{conv}_{(c)}(A)$ die konvexe Hülle in Sinne der Definition 24 (c), d.h., der Durchschnitt von allen konvexen Mengen, die A enthalten. Da liegt sie in jeder konvexen Menge C , die A enthält.

(a) \iff (b)

Lemma 16. Menge $A \subseteq \mathbb{R}^n$ ist genau dann konvex, wenn für alle $k \in \mathbb{N}$, für alle $x_1, \dots, x_k \in A$ und für alle $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ mit $\sum_i \lambda_i = 1$, $\lambda_i \geq 0$ gilt:

$$\lambda_1 x_1 + \dots + \lambda_k x_k \in A \quad (*)$$

Beweis \Leftarrow : Angenommen die Punkte $x_1, x_2 \in A$. Wir nehmen $k = 2$ und $\lambda_1 = (1 - t)$, $\lambda_2 = t$. Dann sind (für $t \in [0, 1]$) $\lambda_1 x_1 + \lambda_2 x_2$ genau die Punkte der Strecke, die x_1 und x_2 verbindet. Falls (*) erfüllt ist, liegen sie in A , und die Menge ist konvex.

Lemma 16. Menge $A \subseteq \mathbb{R}^n$ ist genau dann konvex, wenn für alle $k \in \mathbb{N}$, für alle $x_1, \dots, x_k \in A$ und für alle $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ mit $\sum_i \lambda_i = 1$, $\lambda_i \geq 0$ gilt:

$$\lambda_1 x_1 + \dots + \lambda_k x_k \in A \quad (*)$$

Beweis \implies : Angenommen A ist konvex.

Zuerst bemerken wir, dass $\lambda_1 x_1 + \dots + \lambda_k x_k$ genau der Schwerpunkt von Punkte x_1, \dots, x_k mit Massen $\lambda_1, \dots, \lambda_k$ ist (weil $\sum_i \lambda_i = 1$). Falls wir alle Massen mit einer Konstante multiplizieren, ändern wir den Schwerpunkt nicht, weil diese Konstante in der Formel

$$S := \vec{0} + \sum_i \frac{m_i}{\sum_j m_j} x_i$$

in Zähler und in Nenner erscheint. Also, wir müssen beweisen, dass für alle $k \in \mathbb{N}$ und für alle Massen m_1, \dots, m_k der Schwerpunkt von $x_1, \dots, x_k \in A$ mit Massen m_1, \dots, m_k wieder in A liegt.

Induktion nach k . [Induktionsanfang](#) für $k = 2$ ist offensichtlich: in dem Fall fällt die Verbindungsstrecke xy mit der Menge $\{(1-t)x + ty \mid t \in [0, 1]\}$ zusammen.

Induktionsschritt $k - 1 \mapsto k$. Es seien $x_1, \dots, x_k \in A$ und $m_1, \dots, m_k \geq 0$ mit $k > 2$, $m_i \geq 0$, $\sum_i m_i > 0$.

Sei S Schwerpunkt von $\{x_1, \dots, x_{k-1}\}$ mit Massen m_1, \dots, m_{k-1} . Nach Induktionsvoraussetzung ist $S \in A$.

Nach Lemma 13 ist Schwerpunkt von $\{x_1, \dots, x_k\}$ (mit Massen m_1, \dots, m_k) der Schwerpunkt von $S \in A$, $x_k \in A$ mit Massen jeweils $m_1 + \dots + m_{k-1}$ und m_k .

Nach Induktionsvoraussetzung liegt Schwerpunkt dann in A , □

(a) \iff (b)

Nach Definition 24(b) ist $\text{conv}_{(b)}(A)$ konvex. Dann liegen nach Lemma 16 alle Punkte der Form $\sum_i \lambda_i x_i$ (wobei $\lambda_i \geq 0$, $\sum_i \lambda_i = 1$, $x_i \in A$) in $\text{conv}_{(b)}(A)$. Also $\text{conv}_{(a)}(A) \subseteq \text{conv}_{(b)}(A)$. Also, bleibt uns nur zu zeigen, dass $\text{conv}_{(a)}(A)$ konvex ist.

Angenommen die Punkte x und y sind konvexe Kombinationen von Punkten $x_1, \dots, x_k \in A$ (OBdA können wir denken, dass die Punkte in der Linearkombination für x und y gleich sind, weil wir die fehlende mit Koeffizient $\lambda = 0$ nehmen können.)

$$x = \sum_i \lambda_i x_i \quad y = \sum_i \mu_i x_i.$$

Dann besteht die Verbindungsstrecke aus den Punkten der Form (wobei $t \in [0, 1]$) $(1 - t) \sum_i \lambda_i x_i + t \sum_i \mu_i x_i = \sum_i \underbrace{((1 - t)\lambda_i + t\mu_i)}_{\eta_i} x_i$.

Die Koeffizienten η_i erfüllen $\eta_i \geq 0$ und $\sum_i \eta_i = \sum_i ((1 - t)\lambda_i + t\mu_i) = (1 - t) + t = 1$. Deswegen liegt nach Lemma 16 jeder Punkt der Verbindungsstrecke in A , □

Satz 24 (von Carathéodory). Ist $A \in \mathbb{R}^n$ und $x \in \text{conv}(A)$, so ist x konvexe Kombination von affin unabhängigen Punkten von A . Insbesondere ist x konvexe Kombination von $n + 1$ oder weniger Punkten von A .

Beweis. Der Punkt $x \in \text{conv}(A)$ hat nach Satz 23 eine Darstellung $x = \sum_{i=1}^k \lambda_i x_i$ mit $x_i \in A$, $\lambda_i > 0$, $\sum_{i=1}^k \lambda_i = 1$, mit einem $k \in \mathbb{N}$, und wir können ODDA annehmen, dass hierbei k minimal gewählt ist. Angenommen, x_1, \dots, x_k wären affin abhängig. Dann gibt es Zahlen $\alpha_1, \dots, \alpha_k \in \mathbb{R}$, die nicht alle Null sind, so dass $\sum_{i=1}^k \alpha_i x_i = \vec{0}$ und $\sum_{i=1}^k \alpha_i = 1$ gilt. Wir können eine Zahl m wählen derart, dass $\frac{\lambda_m}{\alpha_m}$ positiv und dabei so klein wie möglich ist (denn alle λ_i sind positiv, und wenigstens ein α_i ist positiv). In der affinen Darstellung

$$x = \sum_{i=1}^k \left(\lambda_i - \frac{\lambda_m}{\alpha_m} \alpha_i \right) x_i$$

sind alle Koeffizienten nichtnegativ (trivialerweise, wenn $\alpha_i \leq 0$ ist, und andernfall wegen der Wahl von m), und wenigstens ein Koeffizient ist Null. Dies ist ein Widerspruch zur Minimalität von k . Also sind x_1, \dots, x_k affin unabhängig, woraus insbesondere $k \leq n + 1$ folgt.

Def. 25 Die konvexe Hülle von endlich vielen Punkten wird als **Polytop** bezeichnet. Ein k -Simplex ist die konvexe Hülle von $k + 1$ affin unabhängigen Punkten, und diese Punkte heißen dann die Ecken des Simplex.

Man kann den Satz von Carathéodory also auch so formulieren, dass $\text{conv}(A)$ die Vereinigung aller Simplizes mit Ecken in A ist.

Satz 25 (von Radon). Jede nichtleere Menge von affin abhängigen Punkten (insbesondere also jede Menge von mindestens $n + 2$ Punkten) im \mathbb{R}^n kann zerlegt werden in zwei Mengen, deren konvexe Hüllen nichtleeren Durchschnitt haben.

Beweis. Seien $x_1, \dots, x_k \in \mathbb{R}^n$ affin abhängig. Dann gibt es Zahlen $\alpha_1, \dots, \alpha_k \in \mathbb{R}$, die nicht alle Null sind, so dass

$$\sum_i \alpha_i = 0 \quad \text{und} \quad \sum_i \alpha_i x_i = \vec{0}$$

gilt. Nach eventueller Umnummerierung können wir annehmen, dass $\alpha_i > 0$ genau für $i = 1, \dots, j$ gilt; dann ist $1 \leq j < k$ (denn wenigstens ein α_i ist > 0 , aber nicht alle α_i sind > 0). Mit $\alpha := \alpha_1 + \dots + \alpha_j = -(\alpha_{j+1} + \dots + \alpha_k) > 0$ gilt

$$x := \sum_{i=1}^j \frac{\alpha_i}{\alpha} x_i = \sum_{i=j+1}^k \left(-\frac{\alpha_i}{\alpha} \right) x_i$$

und daher $x \in \text{conv}(\{x_1, \dots, x_j\}) \cup \text{conv}(\{x_{j+1}, \dots, x_k\})$. Daraus folgt die Behauptung.

Satz 26 (von Helly). Seien $A_1, \dots, A_k \subset \mathbb{R}^n$ konvexe Mengen. Haben je $n + 1$ dieser Mengen nichtleeren Durchschnitt, so haben alle dieser Mengen nichtleeren Durchschnitt.

Induktionsbeweis. Für $k < n + 1$ ist nichts zu zeigen, und für $k = n + 1$ ist die Behauptung trivial – **Induktionsanfang**.

Sei $k \geq n + 2$, und sei die Behauptung bewiesen für $k - 1$ konvexe Mengen. Für jedes $i \in \{1, \dots, k\}$ gibt es dann einen Punkt

$$x_i \in A_1 \cap \dots \cap \check{A}_i \cap \dots \cap A_k,$$

wobei \check{A}_i bedeutet, dass die Menge A_i weggelassen wird. Die $k - n + 2$ Punkte x_1, \dots, x_k sind affin abhängig. Nach Satz 25 (von Radon) können wir daher nach geeigneter Umnummerierung annehmen, dass es einen Punkt

$$x \in \text{conv}(\{x_1, \dots, x_j\}) \cap \text{conv}(\{x_{j+1}, \dots, x_k\})$$

gibt, wo $j \in \{1, \dots, k - 1\}$ geeignet gewählt ist. Wegen $x_1, \dots, x_j \in A_{j+1}, \dots, A_k$ gilt

$$x \in \text{conv}(\{x_1, \dots, x_j\}) \subset A_{j+1}, \dots, A_k \quad \text{und analog}$$

$$x \in \text{conv}(\{x_{j+1}, \dots, x_k\}) \subset A_1, \dots, A_j,$$

also $x \in A_1 \cap \dots \cap A_k$,



Seien $\alpha, \beta \in \mathbb{R}$. Wir definieren

$$\alpha A + \beta B := \{\alpha x + \beta y \mid x \in A, y \in B\} \subseteq \mathbb{R}^n.$$

(die Menge $\alpha A + \beta B$ hängt von der Wahl von $\vec{0}$ ab); falls wir den Punkt $\vec{0}$ verschieben, wird auch $\alpha A + \beta B$ verschoben.

Die Menge $\alpha A + \beta B$ heißt eine **Linearkombination** von Mengen A und B ; die Operation „+“ heißt **Vector-Addition**.

Spezialfälle haben auch besondere Namen:

$A + B$	die Summe
$A + \{v\}$	ein Translat von A
αA	eine Vielfachheit von A
$-A := (-1)A$	Spiegelung von A
$A - B := A + (-B)$	Difference von A und B

Lemma 17. Ist A und B konvex, so ist $\alpha A + \beta B$ auch konvex.

Beweis. Wir nehmen $\alpha \underbrace{a_i}_{\in A} + \beta \underbrace{b_i}_{\in B} \in \alpha A + \beta B$, $i = 1, 2$. Dann ist die

Verbindungsstrecke

$$\begin{aligned} & \{(1-t)(\alpha a_1 + \beta b_1) + t(\alpha a_2 + \beta b_2) \mid t \in [0, 1]\} = \\ & = \left\{ \alpha \underbrace{((1-t)a_1 + ta_2)}_{\in A} + \beta \underbrace{((1-t)b_1 + tb_2)}_{\in B} \mid t \in [0, 1] \right\} \subseteq \alpha A + \beta B. \end{aligned}$$

Aus dem Satz von Helly kann man mehrere Aussagen der kombinatorischen Geometrie herleiten; hierfür ein kleines Beispiel. Wir sagen, dass die Mengen A_1, \dots, A_m von K getroffen werden, wenn $A_i \cap K \neq \emptyset$ für $i = 1, \dots, m$ gilt.

Folgerung. Sei M eine endliche Familie konvexer Mengen im \mathbb{R}^n , sei $K \in \mathbb{R}^n$ konvex. Werden je $n + 1$ Elemente von M von einem Translat von K getroffen, so werden alle Elemente von M von einem Translat von K getroffen.

Wiederholung. **Translat** von K ist die Menge $K + \{v\} = \{x + v \mid x \in K\}$.

Beweis. Sei $M = \{A_1, \dots, A_k\}$. Zu je $n + 1$ der Zahlen $\{1, \dots, k\}$, etwa zu $1, \dots, n + 1$, gibt es ein $v \in \mathbb{R}^n$ und Punkte $a_i \in A_i \cap (K + \{v\})$. Es ist also $a_i = k_i + v$ mit $a_i \in A_i$ und $k_i \in K$. Es folgt $v = a_i - k_i \in A_i - K$ für $i = 1, \dots, n + 1$. Also haben je $n + 1$ Elemente der Familie $\{A_1 - K, \dots, A_k - K\}$ nichtleeren Durchschnitt. Da $A_i - K$ konvex ist ($i = 1, \dots, k$), haben nach dem Satz von Helly alle Elemente dieser Familie nichtleeren Durchschnitt. Es gibt also ein $v \in \mathbb{R}^n$ mit $v \in A_i - K$, also $A_i \cap (K + v) \neq \emptyset$ für $i = 1, \dots, k$, □

Topologische Begriffe in der Konvexe Geometrie

Wiederholung. **Offener** Ball vom Radius $r \geq 0$ um x_0

$$B(r, x_0) := \{x \in \mathbb{R}^n \mid |x - x_0| < r\}.$$

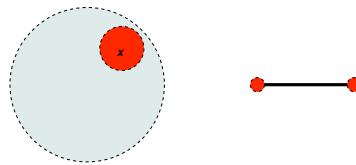
Def. 26 (Analysis- Vorlesung) Sei A eine Teilmenge $A \in \mathbb{R}^n$. Das **Innere** von A (Bezeichnung: $\text{int}(A)$) ist die Vereinigung von aller offene Bälle, die ganz in A enthalten sind.

Eine Menge ist **offen**, falls $A = \text{int}(A)$.

Das **relative Innere** einer Menge A (Bezeichnung: $\text{relint}(A)$) ist das Innere in $\text{aff}(A)$.

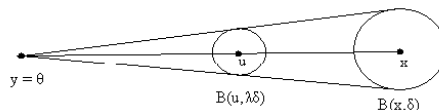
Bsp. Ein offener Ball ist offen.

$$\text{relint}(\{(1-t)x + ty \mid t \in [0, 1]\}) = \{(1-t)x + ty \mid t \in (0, 1)\}.$$



Lemma 18. Sei A eine konvexe Menge. Falls $x \in \text{int}(A)$ und $y \in A$, dann ist $\text{relint}(xy) \subset \text{int}(A)$, wobei xy die Verbindungsstrecke ist.

Beweis. Sei $y \in A$ und o.B.d.A. sei $y = \vec{0}$ der Ursprung. Sei $x \in \text{int}(A)$.
^{Def. 26}
 \exists eine offene Kugel $B(x, \delta) \subset A$. Für beliebiges $u \in \text{relint}(xy)$ existiert ein λ , ($0 < \lambda < 1$), so dass $u = \lambda x$. Es gilt:
 $B(\lambda\delta, \lambda x) = \lambda B(\delta, x)$



Weil A konvex ist und $\vec{0} \in A$, gilt: $\lambda B(\delta, x) \subset A$
 $B(\lambda\delta, \lambda x) = B(\lambda\delta, u) \subset A$.

Da die Kugel um u mit Radius $\lambda\delta$ in A enthalten ist, folgt daraus, dass $u \in \text{int}(A)$, □

Folgerung. Falls A konvex ist, dann ist das Innere $\text{int}(A)$ konvex.

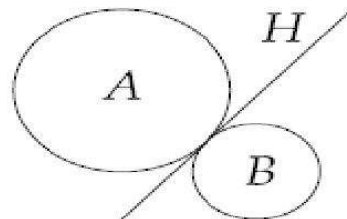
Beweis. Seien $x, y \in \text{int}(A)$. Nach Lemma 17 gilt: $\text{relint}(xy) \subset \text{int}(A)$.
 Nun ist aber $xy = \text{relint}(xy) \cup \{x, y\}$, daher ist $xy \subset \text{int}(A)$. Damit gilt, daß $\text{int}(A)$ konvex ist, □

Separierende Hyperebenen

Hyperebenen = affiner Unterraum von \mathbb{R}^n der Dimension $n - 1$.

Nach Satz 19 ist jede Hyperebene die Lösungsmenge einer linearen Gleichung $\underbrace{l_1x_1 + \dots + l_nx_n}_{\langle l, x \rangle} = c$.

Def. 27 Zwei Mengen A, B werden von der Hyperebene $\langle l, x \rangle = c$ **separiert** (man sagt auch: **schwach separiert**), falls für alle $a \in A$, $b \in B$ gilt: $\langle l, a \rangle \leq c$ und $\langle l, b \rangle \geq c$ bzw. $\langle l, a \rangle \geq c$ und $\langle l, b \rangle \leq c$. Zwei Mengen A, B werden von der Hyperebene $\langle l, x \rangle = b$ **streng separiert**, falls für alle $a \in A$, $b \in B$ gilt: $\langle l, a \rangle < c$ und $\langle l, b \rangle > c$ bzw. $\langle l, a \rangle < c$ und $\langle l, b \rangle > c$.



Aufgrund dieser Definitionen können zwei Mengen überhaupt nur dann streng separiert werden, wenn sie disjunkt sind.

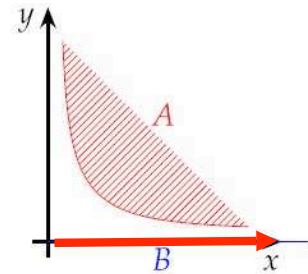
Trotzdem ist die Bedingung, daß zwei Mengen streng separierbar sind, wenn sie disjunkt sind, lediglich notwendig und nicht hinreichend. Dies gilt insbesondere für abgeschlossene konvexe Mengen, wie das folgende Beispiel zeigt:

Die Mengen

$$A := \{(x, y) \in \mathbb{R}^2 \mid x > 0, y \geq \frac{1}{x}\}$$

$$B := \{(x, y) \in \mathbb{R}^2 \mid x \geq 0, y = 0\}$$

sind disjunkte, abgeschlossene konvexe Mengen. Es existiert aber keine Hyperebene (Gerade im \mathbb{R}^2), die diese Mengen streng separiert, da die Menge B auf der Geraden liegt, die waagerechte Asymptote an den Rand der Menge A ist.



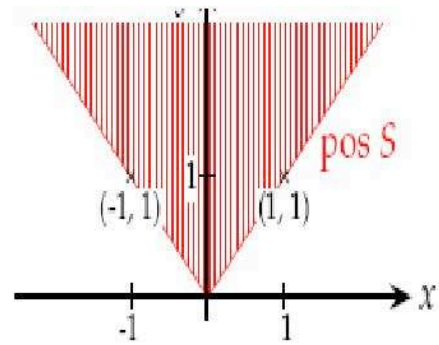
Def. 27 Der Punkt $\lambda_1 x_1 + \dots + \lambda_k x_k$, $\lambda_i \geq 0$, $x_i \in A$ heißt **positive Kombination** von Elementen von A . **Positive Hülle** von A $\text{pos}(A)$ ist die Menge aller konvexen Kombinationen der Elemente von A .

① Die Menge

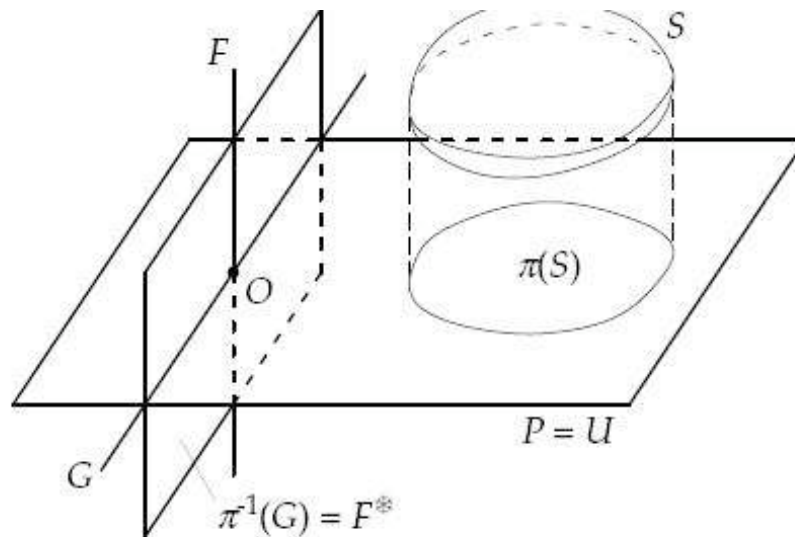
$$S := \{(-1, 1), (1, 1)\} \subset \mathbb{R}^2$$

besitzt die positive Hülle

$$\text{pos } S := \{(x, y) \in \mathbb{R}^2 \mid y \geq |x|\}.$$

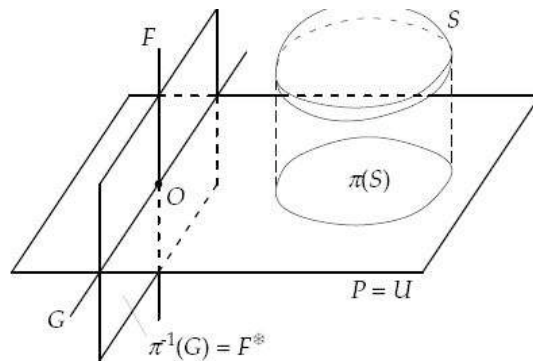


Satz 27. Sei \mathcal{F} ein affiner Unterraum von \mathbb{R}^n mit $\dim(\mathcal{F}) = k$. Sei $S \in \mathbb{R}^n$ offen und konvex, so daß $\mathcal{F} \cap S = \emptyset$. Falls $0 \leq k \leq n - 2$ ist, existiert ein $(k + 1)$ -dimensionaler affiner Unterraum \mathcal{F}^* mit $\mathcal{F} \subseteq \mathcal{F}^*$ und $\mathcal{F}^* \cap S = \emptyset$.



Beweis. Sei e_1, \dots, e_n die Standardbasis des \mathbb{R}^n . OBdA ist $\vec{0} \in \mathcal{F}$. Für $0 < k \leq n - 2$ sei $\mathcal{F} = \text{aff}(\{\vec{0}, e_1, \dots, e_k\})$ und $\mathcal{U} = \text{aff}(\{\vec{0}, e_{k+1}, \dots, e_n\})$. Sei $\pi : \mathbb{R}^n \rightarrow \mathcal{U}$ die orthogonale Projektion auf \mathcal{U} , $\pi(a_1, \dots, a_n) := (0, \dots, 0, a_{k+1}, \dots, a_n)$. Da π offensichtlich eine affine Abbildung ist, ist $\text{Bild}_\pi(S)$ konvexe Teilmenge von \mathcal{U} .

Wir betrachten die Ebene $\mathcal{P} := \text{aff}(\{\vec{0}, e_n, e_{n-1}\})$. Die Schnittmenge $\text{Bild}_\pi(S) \cap \mathcal{P}$ auch eine konvexe Menge ist; sie ist offen als Teilmenge von \mathcal{P} .



Wegen $\mathcal{F} \cap S = \emptyset$ gilt: $\vec{0} \notin \text{Bild}_\pi(S)$, deswegen $\vec{0} \notin \text{Bild}_\pi(S) \cap \mathcal{P}$.
 Deswegen existiert nach Lemma 19 eine Gerade $\mathcal{G} \subseteq \mathcal{P}$ mit $x := 0 \in \mathcal{G}$ und $\mathcal{G} \cap \text{Bild}_\pi(S) = \emptyset$.

Für $\mathcal{F}^* := \text{Urbild}_\pi \mathcal{F} = \mathcal{F} + \mathcal{G}$ gilt: $\dim(\mathcal{F}^*) = \dim(\mathcal{F}) + 1$ und $\mathcal{F}^* \cap S = \emptyset$, □

Folgerung. Sei $\mathcal{F} \in \mathbb{R}^n$ ein affiner Unterraum mit $\dim(\mathcal{F}) = k$. Sei $S \in \mathbb{R}^n$ offen und konvex, so daß $\mathcal{F} \cap S = \emptyset$. Falls $0 \leq k < n$ ist, existiert eine Hyperebene \mathcal{H} , mit $\mathcal{F} \subseteq \mathcal{H}$ und $\mathcal{H} \cap S = \emptyset$.

Beweis. $((n - 1) - k)$ -maliges Anwenden von Satz 27 liefert die Behauptung, da die Dimension des affinen Raumes jeweils um 1 wächst, bis der Raum eine Hyperebene ist, □

Satz 28 Seien $A, B \subseteq \mathbb{R}^n$ konvex, so daß $\text{int}(A) \neq \emptyset$ und $B \cap \text{int}(A) = \emptyset$ ist. Dann existiert eine Hyperebene \mathcal{H} , die die Mengen A und B schwach separiert.

Beweis. Sei A zunächst zusätzlich offen. Es gilt nach Lemma 17: $A - B$ ist offen und konvex. Da A offen ist, ist $\text{int}(A) = A$, weshalb wegen $B \cap A = B \cap \text{int}(A) = \emptyset$ gilt: $\vec{0} \notin A - B$. Damit gibt es nach Folgerung oben eine Hyperebene $\mathcal{H}_0 := \{\langle l, x \rangle = 0\}$ durch $\vec{0}$ mit $\mathcal{H}_0 \cap (A - B) = \emptyset$. Da $A - B$ konvex ist, können wir oBdA annehmen, dass $\langle l, x \rangle > 0$ für alle $x \in (A - B)$. Sonst existieren $x, y \in (A - B)$ so dass $\langle l, x \rangle > 0$ und $\langle l, y \rangle < 0$. Dann gibt es ein $t \in [0, 1]$ so dass $\langle l, (1 - t)x + ty \rangle = 0$, weil die Funktion $f(t) := \langle l, (1 - t)x + ty \rangle$ stetig ist. Dies widerspricht $\mathcal{H}_0 \cap (A - B) = \emptyset$. Dann gilt für alle $a \in A$ und $b \in B$:

$$0 < \langle l, a - b \rangle = \langle l, a \rangle - \langle l, b \rangle \implies \langle l, a \rangle > \langle l, b \rangle.$$

Folglich ist die Menge $\{\langle l, a \rangle \mid a \in A\}$ nach unten beschränkt. Deswegen existiert $\alpha := \inf\{\langle l, a \rangle \mid a \in A\}$. Daher werden A und B von der Hyperebene $\mathcal{H} := \{\langle l, x \rangle = \alpha\}$ separiert.

Ist A nicht offen, existiert gemäß der vorangegangenen Argumentation für die offene Menge $\text{int}(A)$ eine Hyperebene $\mathcal{H} := \{\langle l, x \rangle = \alpha\}$, die $\text{int}(A)$ und $\text{int}(B)$ separiert:

$\langle l, x \rangle \leq \alpha$ für $\forall x \in \text{int}(A)$ und $\langle l, x \rangle \geq \alpha$ für $x \in \text{int}(B)$. Da

$f(x) = \langle l, x \rangle$ stetig ist (als Abbildung von $\mathbb{R}^n \rightarrow \mathbb{R}$), folgt aus Lemma 18:

$\langle l, x \rangle \leq \alpha$ für $\forall x \in \text{int}(A)$, weshalb die Hyperebene $\mathcal{H} := \{\langle l, x \rangle = \alpha\}$

auch A und B separiert, □

Sätze zur Klausur

1. Klausur am Montag 16.06.
2. Während der Klausur werden Sie eine Aussage aus der Liste beweisen (oder eine Teilaussage aus dem Beweis der Aussage).

- ▶ Mengenlehre

- ▶ Satz 2 (Wann sind zwei Menge gleichmächtig)
- ▶ Satz 4 $|2^{\mathbb{N}}| = |\mathbb{R}|$

- ▶ Jordan-Form

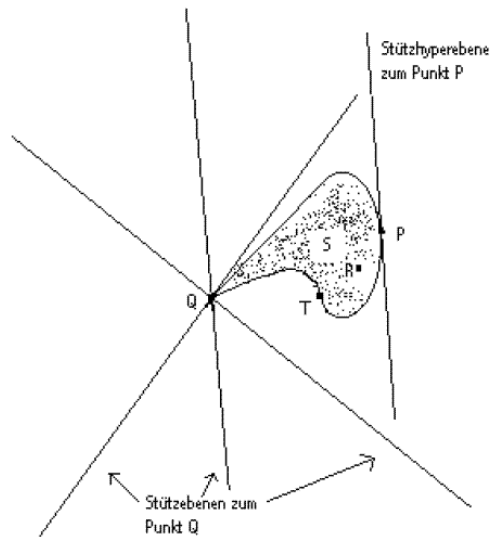
- ▶ Satz 5 (Normalform für die Endomorphismen, die über \mathbb{C} diagonalisierbar sind.)
- ▶ Satz 11 (Zerlegung in Produkt von verallgemeinerten Eigenräumen)
- ▶ Lemma 4 (Zerlegungslemma)
- ▶ Satz 13 (Jordan-Normalform)

- ▶ Affine und konvexe Geometrie

- ▶ Satz 20 (Stralensatz)
- ▶ Satz 22 (Fundamentalsatz der affinen Geometrie)
- ▶ Satz 23 (Konvexe Hülle)

Stützende Hyperebenen

Def. 28 Eine Hyperebene \mathcal{H} **stützt** eine Menge $S \in \mathbb{R}^n$ in einem Punkt $x \in S$, wenn $x \in \mathcal{H}$ und \mathcal{H} die Menge S **beschränkt**, d.h. S liegt ganz auf einer Seite von \mathcal{H} (d.h. für $\mathcal{H} = \{y \in \mathbb{R}^n \mid \langle l, y \rangle = c\}$ gilt dann $\langle l, y \rangle \geq c$ für alle $y \in S$, oder $\langle l, y \rangle \leq c$ für alle $y \in S$.) \mathcal{H} heißt **Stützhyperebene**.



Der Punkt Q besitzt mehrere Stützebenen, der Punkt P genau eine, die Tangente. Die Punkte T und R besitzen gar keine Stützebenen.

Def. (Analysis-Vorlesung) Die Menge $Rand(S) := S \setminus int(S)$ heißt die **Randmenge** von S .

Satz 29. Sei $S \subseteq \mathbb{R}^n$, $S \neq \mathbb{R}^n$ abgeschlossen und konvex. Es gilt: Ist x ein Randpunkt von S , so existiert mindestens eine Hyperebene, die S in x stützt.

Beweis. Ist $dim(aff(S)) < n$, so ist jede Hyperebene, die S enthält, eine stützende Hyperebene von S in jedem Punkt $x \in S$ nach Definition.

Ist $dim(aff(S)) = n$, so ist $int(S) \neq \emptyset$. In der Tat, falls $x_0, \dots, x_n \in S$ affin unabhängig sind, dann liegt das ganze Simplex $conv(\{x_0, \dots, x_n\})$ auch in S , und es hat innere Punkte.

Sei x Randpunkt von S . Wir haben:

$\left. \begin{array}{l} int(S) \text{ ist offen und } \neq \emptyset \\ x \notin int(S) \text{ ist ein} \\ \text{0-dim affiner Raum} \end{array} \right\}$	$\text{Folg. aus Satz 27} \implies$	$\exists \text{ eine Hyperebene } \mathcal{H} \text{ durch } x, \\ \text{die } int(S) \text{ nicht schneidet.}$
---	-------------------------------------	---

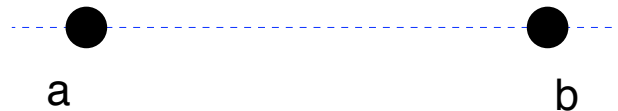
Die Hyperebene \mathcal{H} sei gegeben durch $\mathcal{H} = \{y \in \mathbb{R}^n \mid \langle l, y \rangle = c\}$. Wir zeigen, dass die Hyperebene $int(S)$ beschränkt ist. In der Tat, nach Konstruktion für jeden $y \in int(S)$ gilt: $\langle l, y \rangle \neq c$. Dann ist entweder $\langle l, y \rangle > c$ für alle $y \in int(S)$, oder $\langle l, y \rangle < c$ für alle $y \in int(S)$ (Wir benutzen, dass $int(S)$ nach Folg. aus Lemma 18 auch konvex ist. Falls es $y_1, y_2 \in S$ mit $\langle l, y_1 \rangle < c$ und $\langle l, y_2 \rangle > c$ gibt, dann gibt es einen Punkt auf der Verbindungsstrecke $y_1 y_2 = \{ty_1 + (1-t)y_2 \mid t \in [0, 1]\}$ mit $\langle l, ty_1 + (1-t)y_2 \rangle = c$ (Zwischenwertsatz), was Voraussetzungen widerspricht.)

Also, beschränkt \mathcal{H} die Menge $\text{int}(S)$. OBdA können wir annehmen, dass $\langle l, y \rangle > c$ für alle $y \in \text{int}(S)$. Wir zeigen, dass $\langle l, y \rangle \geq c$ für alle $y \in S$. Man nehme einen $z \in \text{rand}(S)$ und betrachte die Verbindungsstrecke zy , wobei y ein innerer Punkt ist. Nach Lemma 18 sind alle Punkte von $zy \setminus \{z\}$ innere Punkte von S , also $\langle l, * \rangle > c$ für alle diese Punkte. Da die Funktion $\langle l, (1-t)z + ty \rangle$ stetig (als Funktion der Variable t) ist, ist dann $\langle l, z \rangle \geq c$. Dann beschränkt die Hyperebene \mathcal{H} die Menge S . Da $x \in \mathcal{H}$, stützt \mathcal{H} die Menge S in x , □

Umkehrung von Satz 29

Satz 30. Sei $S \subseteq \mathbb{R}^n$ abgeschlossen und $\text{int}(S) \neq \emptyset$. Es gilt: Wenn durch jeden Randpunkt von S eine Hyperebene durchgeht, die S stützt, dann ist S konvex.

Bemerkung. Dies ist nur eine partielle Umkehrung von Satz 29, da nun zusätzlich gefordert wird, dass $\text{int}(S) \neq \emptyset$. Dass diese zusätzliche Bedingung notwendig ist, zeigt dieses Beispiel in \mathbb{R}^1 :



Sei $S = \{a, b\}$ mit $a \neq b$. Dann ist S abgeschlossen (ein Punkt ist abgeschlossen, ebenso die endliche Vereinigung) und die Stützhyperebenen, die in diesem Fall null-dimensional sind, sind die beiden Punkte selber. Also geht durch jeden Randpunkt von S eine Stützebene. Trotzdem ist S sicherlich nicht konvex.

Satz 30. Sei $S \subseteq \mathbb{R}^n$ abgeschlossen und $\text{int}(S) \neq \emptyset$. Es gilt: Wenn durch jeden Randpunkt von S eine Hyperebene durchgeht, die S stützt, dann ist S konvex.

Beweis für $n = 1$. Sei $S \subseteq \mathbb{R}$ abgeschlossen und $\text{int}(S) \neq \emptyset$. Ist $S = \mathbb{R}$, so ist S konvex. Wir nehmen an, dass $S \neq \mathbb{R}$ nicht konvex ist (obwohl durch jeden Randpunkt von S eine Stützhyperebene durchgeht), und bekommen ein Widerspruch.

Da nach Voraussetzung $\text{int}(S) \neq \emptyset$ ist, gibt es ein $u \in \text{int}(S)$. Sei I ein abgeschlossenes Intervall, $I \subseteq S$, wobei $I = [a, b]$ mit

$$a := \inf\{x \in S \mid [x, u] \subseteq S\} \quad \text{und} \quad b := \sup\{y \in S \mid [u, y] \subseteq S\}.$$

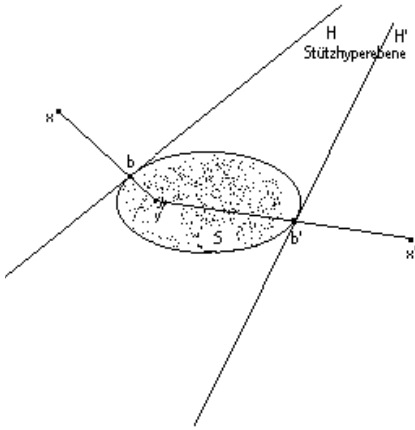
Da S nach unserer Annahme nicht konvex ist, gilt $[a, b] \not\subseteq S$. Dies bedeutet jedoch, dass

- Fall 1.** $\exists y_0 \in S, y_0 > b$. Dann ist aber $b \in S$ ein Randpunkt, durch den keine Stützebene geht – Widerspruch.
- Fall 2.** $\exists x_0 \in S, x_0 < a$. Dann ist aber $a \in S$ ein Randpunkt, durch den keine Stützebene geht – Widerspruch.

Also, Satz 30 ist in dimension 1 bewiesen.

Satz 30. Sei $S \subseteq \mathbb{R}^n$ abgeschlossen und $\text{int}(S) \neq \emptyset$. Es gilt: Wenn durch jeden Randpunkt von S eine Hyperebene durchgeht, die S stützt, dann ist S konvex.

Beweis für $n \geq 2$. Sei wieder $S \subseteq \mathbb{R}^n$ abgeschlossen und $\text{int}(S) \neq \emptyset$. Ist $S = \mathbb{R}^n$, so ist S konvex. Wir nehmen an, dass $S \neq \mathbb{R}^n$. Wir nehmen einen Punkt $x \notin S$. Außerdem gibt es, da $\text{int}(S) \neq \emptyset$, ein $y \in \text{int}(S)$ und einen Randpunkt b von S , mit $b \in \text{relint}(xy)$.



Die nach Voraussetzung existierende stützende Hyperebene \mathcal{H} von S durch b enthält nicht x (denn würde sie b und x enthalten, so enthielte sie auch die Strecke bx und damit auch die Strecke xy , was im Widerspruch dazu steht, dass y ein innerer Punkt ist). Folglich gilt für den abgeschlossenen Halbraum, der von \mathcal{H} begrenzt wird und y enthält, dass er S einschließt, aber nicht x enthält.

Da x ein beliebiger Punkt nicht in S ist, kann man dieses Verfahren für alle $x \notin S$ anwenden und so schließen, dass S gleich dem Schnitt aller so entstandenen Halbräume, welche S enthalten, ist. Also ist S ein Schnitt von konvexen Mengen und selbst konvex,

Sätze 29, 30 in Worten

Die Sätze 29 und 30 kombiniert ergeben die folgende Charakterisierung von abgeschlossenen konvexen Mengen mit nicht leerem Inneren:

Sei $S \subseteq \mathbb{R}^n$ abgeschlossen und $\text{int}(S) \neq \emptyset$. Dann gilt:

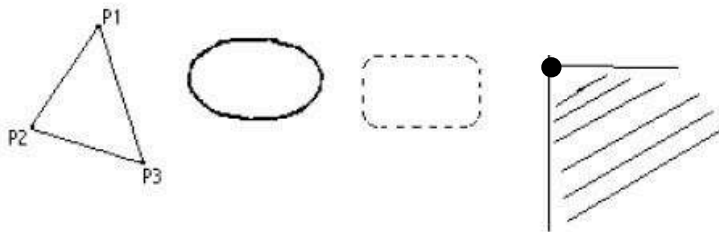
S ist konvex genau dann, wenn durch jeden Randpunkt von S eine die Menge S stützende Hyperebene geht.

\implies Gilt nach Satz 29

\impliedby Gilt nach Satz 30

Bemerkung. Das interessante an diesem wichtigen Satz ist, dass er die globale paarweise Definition von Konvexität (Die Strecke zweier beliebiger Punkte muß ganz in S liegen) mit einer Eigenschaft von einigen einzelnen Punkten (jeder Randpunkt muss in einer Stützebene liegen) gleichsetzt.

Def. 29 Sei $S \subseteq \mathbb{R}^n$ eine konvexe Menge. Ein Punkt $x \in S$ heißt **Extrempunkt** von S , wenn es kein Intervall ganz in S liegend gibt, die x in ihrem relativen Inneren enthält. Die Menge aller Extrempunkte von S heißt **Profil** von S .



Das Profil eines Dreiecks sind seine drei Eckpunkte, das Profil einer Ellipse ist der ganze Rand. Offene Mengen haben kein Profil, genauso wenig der ganze Raum und Halbräume. Viertelräume hingegen haben einen Eckpunkt als Profil.

Wiederholung – Vorl. Analysis. Eine abgeschlossene beschränkte Menge in \mathbb{R}^n ist **kompakt**.

Satz 31. Sei $S \subseteq \mathbb{R}^n$ kompakte und konvexe Menge. Es gilt: S ist die konvexe Hülle ihres Profils P .

In anderen Worten, Jedes $x \in S$ ist Konvexkombination von Elementen des Profils S .

Zuerst eine Hilfsaussage:

Lemma 20 Sei $S \in \mathbb{R}^n$ abgeschlossen und konvex. Sei \mathcal{H} eine Stützhyperebene von S in einem Punkt $x \in S$. Es gilt: Ist x ein Extrempunkt von $\mathcal{H} \cap S$, so ist x ein Extrempunkt von S .

Beweis. Sei $\mathcal{H} := \{y \mid \langle l, y \rangle = c\}$. Sei x aus dem Profil von $S \cap \mathcal{H}$. Dann gilt $\langle l, x \rangle = c$. OBdA ist $\langle l, z \rangle \leq c$ für alle $z \in S$.

Wir nehmen an, dass x nicht aus dem Profil von S ist. Das heißt, x liegt im Inneren eines Intervals, also $x = \lambda \cdot u + (1 - \lambda) \cdot v$ mit

$\lambda \in (0, 1), u, v \in S$ und entweder $u \notin S \cap \mathcal{H}$ oder $v \notin S \cap \mathcal{H}$ (sonst wäre x kein Extrempunkt in $S \cap \mathcal{H}$.) Folglich gilt $\langle l, u \rangle < c$ oder $\langle l, v \rangle > c$.

OBdA können wir annehmen, dass $\langle l, u \rangle < c$.

Dann gilt $c = \langle l, x \rangle = \langle l, \lambda \cdot u + (1 - \lambda) \cdot v \rangle \stackrel{\text{Linearität}}{=} \lambda \cdot \langle l, u \rangle + (1 - \lambda) \cdot \langle l, v \rangle$

$\lambda \cdot \langle l, u \rangle + (1 - \lambda) \cdot \langle l, v \rangle < \lambda \cdot c + (1 - \lambda) \cdot c < c$. Widerspruch zeigt, dass $x \in P$, □

Satz 31. Sei $S \subseteq \mathbb{R}^n$ kompakte und konvexe Menge. Es gilt: S ist die konvexe Hülle ihres Profils P .

Induktionsbeweis nach $k := \dim(\text{aff}(S))$.

InduktionsAnfang: Ist $k = 0$, so ist S nur ein Punkt. Da dieser Punkt trivialerweise auch das Profil P von S ist, gilt $\text{conv}(P) = S$.

InduktionsVoraussetzung: Gelte der Satz für jede kompakte, konvexe Menge der Dimension **höchstens** $k - 1$.

Sätze zur Klausur

1. Klausur am Montag 16.06.
2. Während der Klausur werden Sie eine Aussage aus der Liste beweisen (oder eine Teilaussage aus dem Beweis der Aussage).

- ▶ Mengenlehre

- ▶ Satz 2 (Wann sind zwei Menge gleichmächtig)
- ▶ Satz 4 $|2^{\mathbb{N}}| = |\mathbb{R}|$

- ▶ Jordan-Form

- ▶ Satz 5 (Normalform für die Endomorphismen, die über \mathbb{C} diagonalisierbar sind.)
- ▶ Satz 11 (Zerlegung in Produkt von verallgemeinerten Eigenräumen)
- ▶ Lemma 4 (Zerlegungslemma)
- ▶ Satz 13 (Jordan-Normalform)

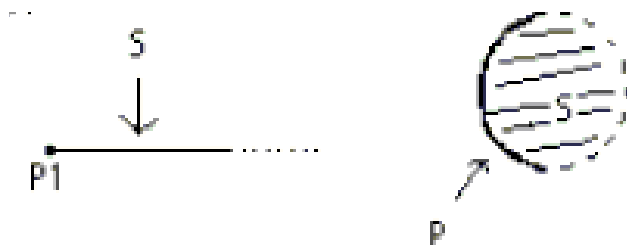
- ▶ Affine und konvexe Geometrie

- ▶ Satz 20 (Stralensatz)

Folgerung. Jede kompakte, konvexe Menge S besitzt mindestens einen Extrempunkt.

Beweis. Weil S nach Satz 31 mit der konvexen Hülle von Extrempunkten übereinstimmt, □

Beispiele, die zeigen, dass die Annahmen wichtig sind:



Links: S ist eine Halbgerade, d.h. S ist konvex aber unbeschränkt und somit nicht kompakt. Das Profil von S ist der Punkt P_1 . Es gilt hier: $\text{conv}(P) \neq S$.

Rechts: S ist ein halboffener, halbgeschlossener Kreis, d.h. S ist konvex und beschränkt, aber nicht abgeschlossen, also nicht kompakt. Das Profil von S ist der linke Halbkreis. Folglich gilt $\text{conv}(P) \neq S$.

Es gibt Ähnlichkeiten zwischen dem Profil einer kompakten, p konvexen Menge und einer Basis eines linearen Untervektorraums bzw.

Koordinatensystem eines affinen Unterraums:

- Eine Basis B eines linearen Unterraums U ist eine linear unabhängige Teilmenge von U , welche den Unterraum aufspannt, in dem Sinn, dass jedes Element von U eine Linearkombination der Elemente in B ist. Jeder Unterraum des \mathbb{R}^n besitzt eine Basis, und obwohl die Basis nicht eindeutig ist, hat sie immer die gleiche Anzahl an Elementen.
- Eine Teilmenge F des \mathbb{R}^n besitzt eine affin unabhängige Teilmenge A von F mit einer endlichen Anzahl von Elementen, so dass die affine Hülle von A gleich affine Hülle von F ist. Obwohl diese Teilmenge nicht eindeutig ist, hat sie immer die gleiche Anzahl an Elementen.
- Wenn wir eine Menge als **konvex unabhängig** definieren, das heißt, dass kein Element der Menge eine Konvexkombination der anderen Elemente ist, dann ist das Profil P von S eine konvex unabhängige Teilmenge von S , wobei S wieder kompakt und konvex, so dass die konvexe Hülle von P gleich S ist. In der Analogie gesprochen heißt das, dass P die Menge S konvex aufspannt. Unglücklicherweise, obwohl das Profil eindeutig ist, kann es unendlich viele Elemente haben.

Wie die Beispiele oben zeigen, falls die Menge S nicht kompakt ist, bricht diese Analogie leider völlig zusammen.

Satz 32. Sei $f \in \mathbb{R}^{n*}$ ein lineare Form, sei $S \in \mathbb{R}^n$ kompakt und konvex. Es gilt: Es existieren Extrempunkte \bar{x} und $\bar{y} \in S$ so dass $f(\bar{x}) = \max_{x \in S} f(x)$ und $f(\bar{y}) = \min_{y \in S} f(x)$.

Wiederholung (Vorl. 17 LAAG I; Def. 39 dort.) Dualraum \mathbb{R}^{n*} ist die Menge der linearen Abbildungen von \mathbb{R}^n nach $\mathbb{R} \equiv \mathbb{R}^1$. Die Elemente des Dualraums heißen die **Linearformen**. Die Menge aller Linearformen ist ein Vektorraum der gleichen Dimension bzgl. natürlichen Operationen. In Koordinaten, kann man jede Linearform in der Form $f(x) = l_1 x_1 + \dots + l_n x_n$ darstellen; also in der Form $f(x) = \langle l, x \rangle$.

Bemerkung. Linearform ist offensichtlich eine stetige Abbildung. Dann nimmt sie (Analysis-Vorlesung) auf kompakten Mengen ihr Minimum (z.B, im Punkt \bar{y}) und Maximum z.B, (im Punkt \bar{x}) an. Also, wir müssen nur zeigen, dass diese Punkte Extrempunkte sind.

Satz 32. Sei $f \in \mathbb{R}^{n*}$ ein lineare Form, sei $S \in \mathbb{R}^n$ kompakt und konvex. Es gilt: Es existieren Extrempunkte \bar{x} und $\bar{y} \in S$ so dass $f(\bar{x}) = \max_{x \in S} f(x)$ und $f(\bar{y}) = \min_{y \in S} f(x)$.

Beweis. Sei $\gamma := \max_{x \in S} f(x)$.

Nach Satz 31 ist $\text{conv}(P) = S$; deswegen gibt es Extrempunkte $x_1, \dots, x_k \in P \subseteq S$ und nicht negative $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ mit $\sum_{i=1}^k \lambda_i = 1$ so dass $\bar{x} = \sum_i \lambda_i x_i$.

Dann ist

$$f(\bar{x}) = f(\lambda_1 x_1 + \dots + \lambda_k x_k) = \lambda_1 \underbrace{f(x_1)}_{\leq \gamma} + \dots + \lambda_k \underbrace{f(x_k)}_{\leq \gamma} \leq \gamma \text{ und ist}$$

$= \gamma$ nur wenn alle $\lambda_i f(x_i) = \lambda_i \gamma$ sind. Da ein von λ nicht 0 ist, impliziert dies, dass ein $f(x_i) = \gamma$, also die Linearform nimmt den maximalen Wert im Extrempunkt x_i an, □

Nachtrag zu separierten Hyperebenen

Def. 27 – Wiederholung Zwei Mengen A, B werden von der Hyperebene $\langle l, x \rangle = c$ **streng separiert**, falls für alle $a \in A, b \in B$ gilt: $\langle l, a \rangle < c$ und $\langle l, b \rangle > c$ bzw. $\langle l, a \rangle < c$ und $\langle l, b \rangle > c$.

Satz 33. Seien $A, B \in \mathbb{R}^n$ nichtleer und konvex. Zudem sei A kompakt und B abgeschlossen. Die Mengen A und B werden genau dann von einer Hyperebene \mathcal{H} streng separiert, wenn $A \cap B = \emptyset$ ist.

Beweis. „ \implies “ Offensichtlich (zu zeigen wäre: \mathcal{H} separiert A und B streng $\implies A \cap B = \emptyset$).

Beweis in „ \impliedby “ wird noch ein Hilfsaussage benötigen.

Def. 30 Unter dem **Abstand** der Mengen A und B versteht man:

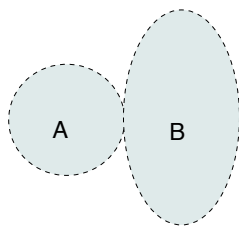
$$d(A, B) := \inf \{d(a, b) \mid a \in A, b \in B\},$$

wobei $d(a, b) := |b - a| := \sqrt{\langle b - a, b - a \rangle}$.

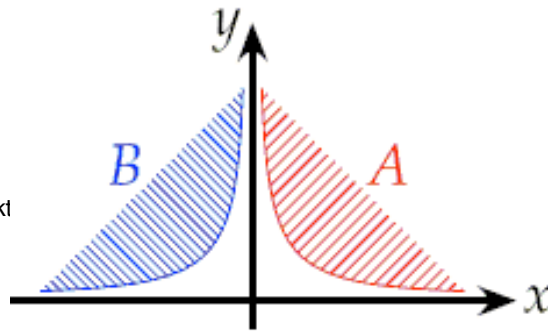
Lemma 21. Sei A kompakt und B abgeschlossen. Dann gilt:

$$d(A, B) = 0 \iff A \cap B \neq \emptyset.$$

Bsp: Kompaktheit und abgeschlossenheit sind wichtig:



$d(A, B) = 0$, obwohl
die Mengen disjunkt
sind



Unter dem **Abstand** der Mengen A und B versteht man:

$$d(A, B) := \inf \{d(a, b) \mid a \in A, b \in B\},$$

wobei $d(a, b) := |b - a| := \sqrt{\langle b - a, b - a \rangle}$.

Lemma 21. Sei A kompakt und B abgeschlossen. Dann gilt:

$$d(A, B) = 0 \iff A \cap B \neq \emptyset.$$

Beweis. „ \Leftarrow “ ist offensichtlich. Wir zeigen „ \Rightarrow “. Angenommen $d(A, B) = 0$. Dann gibt es Folgen $a_k, k \in \mathbb{N}$ in A und $b_k, k \in \mathbb{N}$ in B mit $0 \leq d(a_k, b_k) < \frac{1}{k}$. Da A kompakt ist, hat die Folge a_k eine konvergente Teilfolge. OBdA können wir annehmen, dass

$$a_k \xrightarrow{\text{konvergiert}} a \in A.$$

Weil $d : A \times B \rightarrow \mathbb{R}$ stetig ist, folgt:

$$0 = \lim_{k \rightarrow \infty} \frac{1}{k} \geq \lim_{k \rightarrow \infty} d(a_k, b_k) \stackrel{\text{weil } d \text{ stetig}}{=} \lim_{k \rightarrow \infty} d(a, b_k) \geq 0.$$

Daraus folgt:

- ▶ $\lim_{k \rightarrow \infty} d(a, b_k) = 0$, und
- ▶ dass die Folge b_k beschränkt ist.

Dann hat sie eine konvergente Teilfolge. OBdA können wir annehmen,

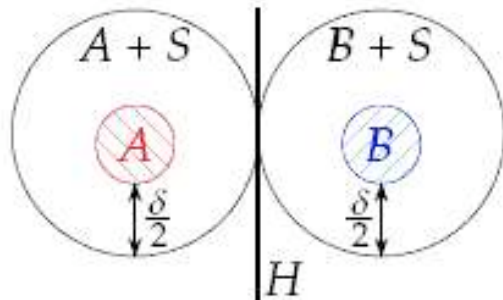
dass $b_k \xrightarrow{\text{konvergiert}} b \in B$. Dann ist $\lim_{k \rightarrow \infty} d(a, b_k) \stackrel{\text{weil } d \text{ stetig}}{=} d(a, b)$.

Also, $d(a, b) = 0$, deswegen $a = b$, deswegen $A \cap B \neq \emptyset$.



Satz 33. Seien $A, B \in \mathbb{R}^n$ nichtleer und konvex. Zudem sei A kompakt und B abgeschlossen. Die Mengen A und B werden genau dann von einer Hyperebene \mathcal{H} streng separiert, wenn $A \cap B = \emptyset$ ist.

Beweis „ \implies “. Sei $\delta := d(A, B) \stackrel{\text{Lemma 21}}{>} 0$.



Sei $S := B\left(\vec{0}, \frac{\delta}{2}\right)$ die offene Kugel um $\vec{0}$ mit Radius $\frac{\delta}{2}$. Dann sind $A + S, B + S \subseteq \mathbb{R}^n$

- ▶ disjunkt, weil

$$a + s_1 = b + s_2 \implies$$

$$a - b = s_2 - s_1 \implies$$

$$\underbrace{|a - b|}_{\geq \delta} = \underbrace{|s_2 - s_1|}_{< \delta}$$
 Widerspruch!

- ▶ offen, weil $A + S$, Vereinigung von der offenen Mengen der Form $a + S$ ist,
- ▶ und konvex nach Lemma 17.

Satz 33. Seien $A, B \in \mathbb{R}^n$ nichtleer und konvex. Zudem sei A kompakt und B abgeschlossen. Die Mengen A und B werden genau dann von einer Hyperebene \mathcal{H} streng separiert, wenn $A \cap B = \emptyset$ ist.

Folgerung. Seien A, \mathbb{R}^n nichtleer, konvex, und abgeschlossen. Dann gilt für jeden Punkt $b \in \mathbb{R} \setminus A$: A und $\{b\}$ werden von einer Hyperebene streng separiert.

Konvexe Polytope

Wiederholung. Eine Teilmenge $S \subseteq \mathbb{R}^n$ heißt **(konvexes) Polytop**, falls $P = \{x_1, x_2, \dots, x_k\} \subseteq \mathbb{R}^n$ existiert, so daß $S = \text{conv}(P)$.

Bemerkung. Konvexes Polytop ist abgeschlossen und beschränkt, also kompakt.

Def. 31 Sei $K \subseteq \mathbb{R}^n$ konvex und abgeschlossen. Sei $F \subseteq K$, und sei $k \in \{0, 1, \dots, \dim(\text{aff}(K))\}$. F heißt **k -Seite** von K , falls $\dim(\text{aff}(F)) = k$ und eine Stützhyperebene \mathcal{H} an K existiert, so daß $F = K \cap \mathcal{H}$.

Triv. Bsp. \emptyset ist n -Seite von K . Man nimmt oft an, daß \emptyset auch eine Seite von K ist (obwohl nach Definition dies falsch ist – stützende Hyperebene haben immer nichtleeren Durchschnitt mit der Menge.)

Def. 31 – Vortsetzung. \emptyset (und \emptyset , falls wir \emptyset künstlich als eine Seite definieren) heißen **uneigentliche** Seiten von K , alle übrigen heißen **eigentliche** Seiten von K . Jede $\dim(K) - 1$ dimensionale Seite von K heißt **Facette**, jede 1-Seite von K **Kante** und jede 0-Seite **Ecke**.

Wir wissen bereits, daß jede kompakte konvexe Menge durch ihr Profil erzeugt wird Satz 32. Im Folgenden werden wir sehen, daß das Profil eines Polytopes gleich der Menge der Ecken von S ist. Desweiteren ist diese Menge die im Sinne der folgenden Definition eindeutig bestimmte minimale Darstellung des Polytopes P .

Def. 32 Die Menge $S := \{x_1, x_2, \dots, x_k\}$ ist eine **minimale Darstellung des Polytopes P** , wenn $P = \text{conv}(\{x_1, x_2, \dots, x_k\})$ und $\forall i \in \{1, 2, \dots, k\}$ gilt: $x_i \notin \text{conv}(S \setminus \{x_i\})$.

Bemerkung. Jedes Polytopes $P = \text{conv}(S)$ mit $S = x_1, x_2, \dots, x_k$ besitzt eine minimale Darstellung. Falls nämlich $\{x_1, x_2, \dots, x_k\}$ nicht minimal ist, existiert $x_i \in S$ mit $x_i \notin \text{conv}(\{x_j \in S \mid j \neq i\})$. Es folgt $P = \text{conv}(\{x_j \in S \mid j \neq i\})$.

Auf diese Weise kann man stets eine echt kleinere Erzeugermenge des Polytopes finden, bis die Erzeugermenge minimal ist.

Satz 34. Sei $M := \{x_1, x_2, \dots, x_k\}$ eine minimale Darstellung des Polytopes P . Dann sind folgende Aussagen äquivalent:

(i) $x \in M$.

(ii) x ist Ecke von P .

(iii) x ist Extrempunkt von P .

(i) \implies (ii): Sei $x \in M$. Man betrachte $\text{conv}(M \setminus \{x\})$. Wegen M minimal und $x \in M$ folgt $x \notin \text{conv}(M \setminus \{x\}) =: Q$.

Da $\{x\}$ und Q kompakt und konvex sind, folgt aus Satz 33 die Existenz einer Hyperebene $\mathcal{H}_0 = \{y \mid \langle l, y \rangle = c_0\}$ die $\{x\}$ und Q streng trennt.

OBdA sei $\langle l, q \rangle > c_0$ für alle $q \in Q$, und $\langle l, x \rangle < c_0$. Sei $c = \langle l, x \rangle$.

Wir zeigen, dass $\mathcal{H} = \{y \mid \langle l, y \rangle = c\}$ wie in Definition von Ecke ist, d.h., \mathcal{H} eine Stützhyperebene an P im Punkt x ist, und $P \cap \mathcal{H} = \{x\}$. Wir benutzen:

$$\langle l, q \rangle > c_0 > \langle l, x \rangle \quad (**)$$

Für jedes $y \in P$, $y \neq x$ gilt:

$$y = \lambda x + \sum_i \lambda_i x_i \quad \text{mit} \quad 0 \leq \lambda < 1, \quad \lambda + \sum_i \lambda_i = 1, \quad x_i \in Q.$$

Dann gilt

$$\langle l, y \rangle = \lambda \langle l, x \rangle + \sum_i \lambda_i \langle l, x_i \rangle \quad \text{mit} \quad 0 \leq \lambda < 1, \quad \lambda + \sum_i \lambda_i = 1, \quad x_i \in Q.$$

Wegen $0 \leq \lambda < 1$ und $\lambda + \sum_i \lambda_i = 1$ existiert ein $i \in \{1, 2, \dots, n\}$ mit $\lambda_i \neq 0$. Es folgt mit (**):

$$\langle l, y \rangle > \lambda \langle l, x \rangle + \sum_i \lambda_i \langle l, x \rangle = \langle l, x \rangle.$$

Wir sehen, dass alle $y \in P$ von einer Seite von \mathcal{H} liegen, und $P \cap \mathcal{H} = \{x\}$. Der Punkt x ist also Ecke von P .



(ii) \implies (iii):

(ii) x ist Ecke von P .

(iii) x ist Extrempunkt von P .

Im Folgenden genügt die Eigenschaft, daß P kompakt und konvex ist.

Widerspruchsbeweis. Sei x eine Ecke von P . Angenommen x ist kein Extrempunkt von P . Dann gilt: \exists eine Strecke $yz \in P$ mit $x \in \text{relint}(yz)$.

Da x Ecke von P ist, \exists eine Hyperebene \mathcal{H} sodass $\mathcal{H} \cap P = \{x\}$ Sei

$\mathcal{H} = \{\langle l, y \rangle = c\}$ und $P \subseteq \{\langle l, y \rangle \geq c\}$. Wegen $yz \not\subseteq \mathcal{H}$ folgt

$\exists \tilde{x} \in yz \subseteq P$ sodass $\langle l, \tilde{x} \rangle > c$. Daraus folgt $\langle l, \tilde{x} - x \rangle > 0$. Da

$x \in \text{relint}(yz)$, kann λ so gewählt werden, daß $x - \lambda(\tilde{x} - x) \in yz \subseteq P$.

Deswegen $\langle l, x - \lambda(\tilde{x} - x) \rangle = \langle l, x \rangle - \lambda \langle l, \tilde{x} - x \rangle < \langle l, x \rangle = c$.

Widerspruch zu $\langle l, P \rangle \geq c$, □

(iii) \implies (i):

(iii) x ist Extrempunkt von P .

(i) $x \in M$.

Sei x ein Extrempunkt von P und M minimale Darstellung von P . Dann ist $P \setminus \{x\}$ konvex. Wäre $x \notin M$, folgte $\text{conv}(M \setminus \{x\}) \subseteq P \setminus \{x\}$, \square

Folgerung. Jedes Polytop P hat genau eine minimale Darstellung.

Beweis. Nach Satz 34 besteht die Minimaldarstellung genau aus der Ecken. \square

Folgerung. Sei $P = \text{conv}(\{x_1, x_2, \dots, x_{k+1}\})$ ein k -dimensionales Simplex (die Punkte x_1, \dots, x_{k+1} sind affin unabhängig.)

Dann sind x_1, x_2, \dots, x_{k+1} die Ecken von P .

Beweis. Dies ist die minimale Darstellung, da $\dim(\text{aff}(\text{conv}(\{x_1, x_2, \dots, x_r\}))) := \dim(\text{aff}(\{x_1, x_2, \dots, x_r\})) < r \leq k$ \square

Satz 35. Sei $P \in \mathbb{R}^n$ ein konvexes Polytop. Dann ist jede Seite von P selbst ein Polytop, und es gibt nur endlich viele Seiten.

Beweis. Sei $\{x_1, x_2, \dots, x_k\}$ die minimale Darstellung von P und sei $F \subseteq P$ eine echte Seite von P . Sei $\mathcal{H} := \{y \mid \langle l, y \rangle = c\}$ eine stützende Hyperebene von P mit $\mathcal{H} \cap P = F$.

OBdA sei $\{x_1, x_2, \dots, x_r\} \subseteq \mathcal{H}$ und $\langle l, x_i \rangle = c + \varepsilon_i$ mit $\varepsilon_i > 0$ für $i \in \{r+1, \dots, k\}$. Da $P \cap \mathcal{H} \neq \emptyset$, folgt $r > 0$. Sei nun $x \in P$. D.h.

$$x = \sum_{i=0}^k \lambda_i x_i, \quad \lambda_i \geq 0 \quad \sum_i \lambda_i = 1$$

$$\implies \langle l, x \rangle = \sum_{i=0}^r \lambda_i \langle l, x_i \rangle + \sum_{i=r+1}^k \lambda_i \langle l, x_i \rangle =$$

$$\sum_{i=0}^r \lambda_i c + \sum_{i=r+1}^k \lambda_i (c + \varepsilon_i) = \sum_{i=0}^k \lambda_i c + \sum_{i=r+1}^k \lambda_i \varepsilon_i.$$

Daraus folgt:

$$x \in \mathcal{H} \iff \lambda_i = 0 \quad \text{für } i \in \{r+1, \dots, k\} \iff x \in \text{conv}(\{x_{r+1}, \dots, x_k\}).$$

D.h. F ist ein Polytop. Wegen $\#\{x_1, x_2, \dots, x_k\} = 2^k < \infty$ hat P nur endliche viele verschiedene Seiten.

Satz 36. Seien F_1, \dots, F_m Seiten einer kompakten, konvexen Menge $S \in \mathbb{R}^n$. Dann ist $F := F_1 \cap \dots \cap F_m$ eine Seite von S .

Beweis. OBdA sei $F \neq \emptyset$, sei $\vec{0} \in F$ und sei F_i echte Seite von S für alle $i \in \{1, 2, \dots, m\}$.

Nach Voraussetzung existieren Stützhyperbenen

$\mathcal{H}_i := \{y \in \mathbb{R}^n \mid \langle l_i, y \rangle = 0\}$, so daß $F_i \subseteq \mathcal{H}_i$ und oBdA $\langle l_i, S \rangle \geq 0$. Wir definieren $l = l_1 + \dots + l_m$ und betrachten $\mathcal{H} = \{y \in \mathbb{R}^n \mid \langle l, y \rangle = 0\}$.

Wir zeigen: $F = \mathcal{H} \cap S$.

Wegen $\langle l_i, S \rangle \geq 0$ gelten folgende Äquivalenzen:

$x \in F \cap S \iff x \in \bigcap_{i=1}^m F_i \cap S \iff x \in \bigcap_{i=1}^m \mathcal{H}_i \cap S \iff x \in \mathcal{H} \cap S$.

Außerdem gilt $\vec{0} \in \mathcal{H} \cap S$ und $\langle l, S \rangle \geq 0$, d.h., F eine Seite von S ist, \square

Wiederholung. Eine Teilmenge $S \subseteq \mathbb{R}^n$ heißt **(konvexes) Polytop**, falls $P = \{x_1, x_2, \dots, x_k\} \subseteq \mathbb{R}^n$ existiert, so daß $S = \text{conv}(P)$.

Bemerkung. Konvexes Polytop ist abgeschlossen und beschränkt, also kompakt.

Def. 31 Sei $K \subseteq \mathbb{R}^n$ konvex und abgeschlossen. Sei $F \subseteq K$, und sei $k \in \{0, 1, \dots, \dim(\text{aff}(K))\}$. F heißt **k -Seite** von K , falls $\dim(\text{aff}(F)) = k$ und eine Stützhyperebene \mathcal{H} an K existiert, so daß $F = K \cap \mathcal{H}$.

K (und \emptyset , falls wir \emptyset künstlich als eine Seite definieren) heißen **uneigentliche** Seiten von K , alle übrigen heißen **eigentliche** Seiten von K . Jede $\dim(K) - 1$ dimensionale Seite von K heißt **Facette**, jede 1–Seite von K **Kante** und jede 0–Seite **Ecke**.

Wir wissen bereits, daß jede konvexe kompakte Menge der Durchschnitt abgeschlossener Halbräume ist (Satz 30). Im Folgenden werden wir feststellen, daß jedes Polytop Durchschnitt endlich vieler abgeschlossener Halbräume ist. Polytope lassen sich sogar als genau die beschränkten Mengen beschreiben, die Durchschnitt endlich vieler abgeschlossener Halbräume sind.

Def. 33 Eine **polyedrische Menge** ist Durchschnitt endlicher vieler abgeschlossener Halbräume.

Satz 37. Jedes Polytop $P \subseteq \mathbb{R}^n$ ist eine beschränkte polyedrische Menge.

Satz 37. Jedes Polytop $P \subseteq \mathbb{R}^n$ ist eine beschränkte polyedrische Menge.

Beweis. Sei $\{x_1, x_2, \dots, x_k\}$ die minimale Darstellung von P . O.B.d.A. sei P n -dimensional. Seien F_1, \dots, F_m die Facetten von P , $m > 0$. Die Existenz von Facetten zeigen wir im Folgenden auch. Zu F_i sei \mathcal{H}_i die entsprechende stützende Hyperebene von P und \mathcal{H}_i^+ der P enthaltende abgeschlossene Halbraum. Es gilt also:

$F_i = \mathcal{H}_i \cap P$ und $P \subseteq \mathcal{H}_i^+$ für $i \in \{1, \dots, m\}$. Wir zeigen, daß $P = \bigcap_{i=1}^m \mathcal{H}_i^+$. Es gilt: $P \subseteq \bigcap_{i=1}^m \mathcal{H}_i^+$.

Für die andere Inklusion $P \supseteq \bigcap_{i=1}^m \mathcal{H}_i^+$ machen wir ein

Widerspruchsbeweis: (für $P \supseteq \bigcap_{i=1}^m \mathcal{H}_i^+$) Angenommen, es existierte $x \in \bigcap_{i=1}^m \mathcal{H}_i^+$ mit $x \notin P$. Definiere:

$$D := \bigcup_{\substack{B \in \{x_1, x_2, \dots, x_k\} \\ \text{mit } \#b \leq n-1}} \text{aff}(\{x\} \cup B).$$

Da $\dim(\text{aff}(\{x\} \cup B)) \leq n-1$, und D endliche Vereinigung solcher höchstens $n-1$ -dimensionalen affinen Unterräume ist, folgt $\dim(D) < n = \dim(P)$. Da außerdem D und P abgeschlossen sind, folgt $\text{int}(P) \not\subseteq D$. Sei $y \in \text{int}(P) \setminus D$, und sei $z \in \text{relint}(xy)$ mit $z \in \text{Rand}(P)$. Da $xy \cap P$ kompakt ist, kann z als dasjenige Element aus $xy \cap P$ definiert werden, das von y maximalen Abstand hat. Es genügt jetzt zu zeigen, daß z auf einer Facette von P liegt und damit $z \in \mathcal{H}_j$ für $j \in \{1, \dots, k\}$ gilt. Dann folgt mit $z \in \mathcal{H}_j$ und $y \in \text{int}(P) \subseteq \mathcal{H}_j^+$, daß $x \notin \bigcap_{i=1}^m \mathcal{H}_i^+$. Widerspruch zur Voraussetzung!

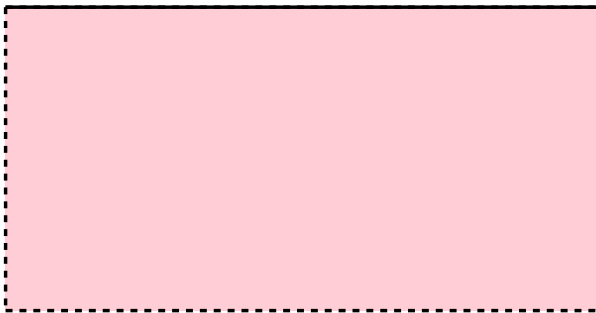
Es bleibt zu zeigen, daß z in einer $n - 1$ -dimensionalen Seite, einer Facette, von P enthalten ist. Wir verwenden hier die oben konstruierte Menge D . Man beachte, daß dafür nur $\dim(P) = n$ und die Existenz eines $x \notin P$ vorausgesetzt werden muß. Mit Satz 29 folgt, daß z als Randpunkt einer kompakten konvexen Menge P mit $\text{int}(P) \neq \emptyset$ auf einer stützenden Hyperebene von P liegt. z liegt also auf einer k -Seite von P . Wir zeigen, daß $k > n - 2$. Angenommen, es wäre $k \leq n - 2$. Da jede k -Seite von P nach Satz 34 ein von Elementen der minimalen Darstellung erzeugtes Polytop ist und da nach dem Satz 24 (Caratheodory) jedes Element der Seite Konvexkombination von höchstens $(n - 2) + 1 = n - 1$ Elementen der Erzeugermenge ist, folgte $z \in \text{conv}(B_0)$ mit $B_0 \subseteq \{x_1, x_2, \dots, x_k\}$ mit $\#B_0 \leq n - 1$. Also wäre $z \in D$ und damit auch $y \in D$. Widerspruch! Also $k = n - 1$, □

Satz 38. Sei $P \subseteq \mathbb{R}^n$ eine polyedrische Menge, die keine Geraden enthält. Sei $f \in \mathbb{R}^{n*}$ eine Linearform (d.h. eine lineare Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}$), die auf P nach oben beschränkt ist. (D.h. $\sup_{x \in P} f(x) < \infty$.) Dann existiert ein Extrempunkt $\bar{x} \in P$, so daß $f(\bar{x}) = \sup_{x \in P} f(x)$.

Bemerkung. Falls P zusätzlich beschränkt ist (in dem Fall enthält P sicher keine Gerade), ist die Aussage zum **Satz 32** äquivalent:

Sei $f \in \mathbb{R}^{n*}$ ein lineare Form, sei $P \in \mathbb{R}^n$ kompakt und konvex. Es gilt: Es existiert einen Extrempunkt $\bar{x} \in S$ so dass $f(\bar{x}) = \sup_{x \in S} f(x)$.

Bemerkung. Die Annahme, dass P keine Gerade enthält, ist wichtig, siehe das Bild



die Linearform $f(x,y)=x$
ist auf Halbebene $x \leq 1$
beschränkt, hat aber
keine Extrempunkte

Bemerkung. Falls konvexe abgeschlossene Menge P eine Gerade enthält (z.B. $\mathcal{G} = \{x + t\vec{v} \mid t \in \mathbb{R}\}$), dann enthält sie mit jedem $y \in P$ die ganze Gerade $\mathcal{G}_{y,\vec{v}} := \{y + t\vec{v} \mid t \in \mathbb{R}\}$, die parallel zu \mathcal{G} ist und den Punkt y enthält.

Bemerkung. Insbesondere gilt: enthält eine abgeschlossene konvexe Menge eine Gerade, so hat sie keine Ecken.

Satz 38. Sei $P \subseteq \mathbb{R}^n$ eine polyedrische Menge, die keine Geraden enthält. Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine Linearform, die auf P nach oben beschränkt ist. (D.h. $\sup_{x \in P} f(x) < \infty$.) Dann existiert ein Extrempunkt $\bar{x} \in P$, so daß $f(\bar{x}) = \sup_{x \in P} f(x)$.

Beweis. OBdA sei $\dim(\text{aff}(P)) = n$. Induktion über n :

InduktionsAnfang $n = 1$: Gilt, da P keine Gerade ist.

InduktionsVoraussetzung: Gelte die Behauptung für $P \subseteq \mathbb{R}^k$, $k \leq n - 1$, $n > 0$.

InduktionsSchritt : Da f linear ist, gilt:

$a := \sup_{x \in P} f(x) = \sup_{x \in \text{Rand}(P)} f(x)$. Da P polyedrisch, folgt:

$\text{Rand}(P) = (P \cap \mathcal{H}_1) \cup (P \cap \mathcal{H}_2) \cup \dots \cup (P \cap \mathcal{H}_m)$ wobei $\mathcal{H}_1, \dots, \mathcal{H}_m$ die P begrenzenden Hyperebenen. Also folgt: $a := \sup_{x \in P \cap \mathcal{H}_i} f(x)$ für

$i \in \{1, \dots, m\}$. Da $P \cap \mathcal{H}_i$ polyedrisch ist und $\dim(P \cap \mathcal{H}_i) < n$ und

außerdem keine Geraden enthält, folgt mit der Induktionsvoraussetzung,

daß ein Extrempunkt $\bar{x} \in P \cap \mathcal{H}_i$ existiert, so daß $a = f(\bar{x})$. Da jeder

Extrempunkt von $P \cap \mathcal{H}_i$ ein Extrempunkt von P ist, folgt die

Behauptung, □

Satz 39. Seien S_2 und S_1 kompakte, konvexe Mengen mit $S_2 \subseteq S_1$. Dann gilt: F Seite von $S_1 \implies F \cap S_2$ Seite von S_2 .

Beweis. Falls F unechte Seite ist, so auch $F \cap S_2$. Sei also F eine echte Seite von S_1 . Dann existiert eine stützende Hyperebene H von S_1 , so daß $F = H \cap S_1$. Dann ist $H \cap S_2 = H \cap S_1 \cap S_2 = F \cap S_2$ Seite von S_2 .

Dualität

Konvexen Mengen kann man (unter schwachen Zusatzvoraussetzungen) duale (oder polare) konvexe Mengen zuordnen. Diese Dualität ist oft von Nutzen und soll hier untersucht werden.

Def. **Konvexe Körper** = konvexe kompakte Menge K in \mathbb{R}^n mit $\text{int}(K) \neq \emptyset$.

Bsp. Abgeschlossener Ball, Polyeder sind konvexe Körper. Polyedrische Menge ist nur dann ein konvexer Körper, wenn sie ein Polyeder ist.

Def. 34 Zunächst sei $K \in \mathbb{R}^n$ eine Menge mit $\vec{0} \in \text{int}(K)$. Wir definieren


$$K^* := \{x \in \mathbb{R}^n \mid \langle x, y \rangle \leq 1 \text{ für alle } y \in K\}.$$

K heißt der **Polarkörper** oder **duale** Körper von K .

Bemerkung. Um duale Körper zu definieren, braucht man nicht vorauszusetzen, dass die ursprüngliche Menge konvex ist

Frage. Ist Polarkörper konvex? Was ist Polarkörper zu Polarkörper?

Bsp. Für die Kugel $B(\vec{0}, \varepsilon)$ mit Radius $\varepsilon > 0$ gilt $B(\vec{0}, \varepsilon)^* = B(\vec{0}, 1/\varepsilon)$, wie unmittelbar aus der Definition folgt.

Der nachfolgende Satz zeigt, dass (falls K ein konvexer Körper ist) ist K^* wieder ein konvexer Körper ist, und dass in der Tat eine Dualität vorliegt. 

Satz 40. Sei $K \in \mathbb{R}^n$ ein konvexer Körper mit $\vec{0} \in \text{int}(K)$. Dann gilt K ist konvex, $\vec{0} \in \text{int}(K^*)$ und $K^{**} = K$.

Beweis. Zuerst **konvexität und Kompaktheit**. Für $x_1, x_2 \in K^*$ und $t \in [0, 1]$ gilt $\langle (1-t)x_1 + tx_2, y \rangle \leq 1$ für alle $y \in K$, also $(1-t)x_1 + tx_2 \in K^*$. Also ist K^* konvex. Trivialerweise ist K^* **abgeschlossen** (konvergiert eine Folge x_k gegen x , so konvergiert die Folge $\langle x_k, y \rangle$ gegen $\langle x, y \rangle$). Falls alle Elemente $\langle x_k, y \rangle$ der Folge ≤ 1 ist, dann ist der Grenzwert $\langle x, y \rangle$ auch ≤ 1).

Für die Kugel $B(\vec{0}, \varepsilon)$ mit Radius $\varepsilon > 0$ gilt $B(\vec{0}, \varepsilon)^* = B(\vec{0}, 1/\varepsilon)$, siehe Bsp. oben. Aus der Definition ergibt sich, dass aus $K_1 \subseteq K_2$ stets $K_1^* \supseteq K_2^*$ folgt (in der Tat, falls für alle $y \in K_2$ die Ungleichung $\langle x, y \rangle \leq 1$ erfüllt ist, dann ist diese Ungleichung auch für alle $y \in K_1$ erfüllt, wenn $K_1 \subseteq K_2$ ist.)

Wir können $\varepsilon, \rho > 0$ wählen mit $B(\vec{0}, \varepsilon) \subseteq K \subseteq B(\vec{0}, \rho)$; dann ist $\underbrace{B(\vec{0}, 1/\rho)}_{B(\vec{0}, \rho)^*} \subseteq K^* \subseteq \underbrace{B(\vec{0}, 1/\varepsilon)}_{B(\vec{0}, \varepsilon)^*}$, also gilt $\vec{0} \in \text{int}(K)$, und K ist beschränkt und daher **kompakt**.

$$K = K^{**}$$

Zuerst $K \subseteq K^{**}$:

Sei $y \in K$. Für beliebiges $x \in K^*$ gilt $\langle x, y \rangle \leq 1$, also ist $y \in K^{**}$. Somit gilt $K \subseteq K^{**}$ (übrigens haben wir bis jetzt weder die Konvexität noch die Abgeschlossenheit von K benutzt).

Jetzt $K \supseteq K^{**}$:

Sei $z \in \mathbb{R}^n \setminus K$. Wir zeigen, dass $z \notin K^{**}$, was äquivalent zu $K \supseteq K^{**}$ ist.

Da K konvex und abgeschlossen ist, gibt es nach Satz 33 eine streng separierte Hyperebene $\mathcal{H}_{l,a} = \{x \in \mathbb{R}^n \mid \langle l, x \rangle = a\}$ mit

$K \in \mathcal{H}_{l,a}^- = \{x \in \mathbb{R}^n \mid \langle l, x \rangle < a\}$ und $\langle l, z \rangle > a$; wegen $\vec{0} \in \text{int}(K)$ ist hier $a > 0$.

Für alle $y \in K$ gilt $\langle l, y \rangle \leq 1$, also ist $l/a \in K^*$. Wegen $\langle \frac{l}{a}, z \rangle > 1$ folgt jetzt $z \notin K^{**}$. Damit ist der Beweis von $K^{**} = K$ erbracht, \square

Wiederholung — Def. 29 Sei $S \subseteq \mathbb{R}^n$ eine konvexe Menge. Ein Punkt $x \in S$ heißt **Extrempunkt** von S , wenn es kein Intervall ganz in S liegend gibt, die x in ihrem relativen Inneren enthält.

Wiederholung—Satz 32. Sei $f \in \mathbb{R}^{n*}$ ein lineare Form, sei $S \in \mathbb{R}^n$ kompakt und konvex. Es gilt: Es existieren Extrempunkte \bar{x} und $\bar{y} \in S$ so dass $f(\bar{x}) = \max_{x \in S} f(x)$ und $f(\bar{y}) = \min_{y \in S} f(x)$.

Wiederholung — Def. 32 Die Menge $S := \{x_1, x_2, \dots, x_k\}$ ist eine **minimale Darstellung des Polytopes P** , wenn $P = \text{conv}(\{x_1, x_2, \dots, x_k\})$ und $\forall i \in \{1, 2, \dots, k\}$ gilt: $x_i \notin \text{conv}(S \setminus \{x_i\})$.

Wiederholung — Satz 34. Sei $M := \{x_1, x_2, \dots, x_k\}$ eine minimale Darstellung des Polytopes P . Dann sind folgende Aussagen äquivalent:

- (i) $x \in M$.
- (ii) x ist Ecke von P .
- (iii) x ist Extrempunkt von P .

Der folgende Satz beschreibt, inwiefern beim Übergang zu Polarkörpern Durchschnitte und Vereinigungen vertauscht werden.

Satz 41. Seien K_1, K_2 konvexe Körper mit $\vec{0} \in \text{int}(K_i)$ ($i = 1, 2$).

Dann gilt:

$$(K_1 \cap K_2)^* = \text{conv}(K_1^* \cup K_2^*),$$

$$\text{conv}(K_1 \cup K_2)^* = K_1^* \cap K_2^*.$$

Ist außerdem $K_1 \cup K_2$ konvex, so ist $K_1^* \cup K_2^*$ auch konvex, also gilt

$$(K_1 \cap K_2)^* = K_1^* \cup K_2^*,$$

$$(K_1 \cup K_2)^* = K_1^* \cap K_2^*.$$

Beweis. Aus $K_1 \cap K_2 \subseteq K_i$ folgt $K_i^* \subseteq (K_1 \cap K_2)^*$, also $\text{conv}(K_1^* \cup K_2^*) \subseteq (K_1 \cap K_2)^*$. Aus $K_i \subseteq \text{conv}(K_1 \cup K_2)$ folgt $\text{conv}(K_1 \cup K_2)^* \subseteq K_i^*$, also $\text{conv}(K_1 \cup K_2)^* \subseteq K_1^* \cap K_2^*$. Wendet man beide Inklusionen auf K_i^* statt K_i an und benutzt $K^{**} = K$, so erhält man die ersten beiden Gleichungen des Satzes.

Sei jetzt $K_1 \cup K_2$ konvex. Sei $x \in \mathbb{R}^n \setminus (K_1^* \cup K_2^*)$. Nach Satz 33 gibt es $l_i \in \mathbb{R}^n \setminus \{\vec{0}\}$ und $a_i \in \mathbb{R}$ mit $\langle l_i, x \rangle > a_i$ und $\langle l_i, y \rangle \leq a_i$ für alle $y \in K_i^*$ (also ist $a_i > 0$), woraus $l_i/a_i \in K_i^{**} = K_i$ folgt ($i = 1, 2$). Da $K_1 \cup K_2$ konvex ist, gibt es eine Konvexe Kombination $z := t * l_1/a_1 + (1 - t)l_2/a_2 \in K_1 \cup K_2$. Wegen $\langle l_i, x \rangle > a_i$ gilt $\langle z, x \rangle > 1$ und somit $x \notin (K_1 \cap K_2)^*$. Damit ist $(K_1 \cap K_2)^* \subseteq K_1^* \cup K_2^*$ gezeigt. Da die entgegengesetzte Inklusion trivial ist, folgt $(K_1 \cup K_2)^* = K_1^* \cap K_2^*$ und damit die Konvexität von $K_1^* \cup K_2^*$, □

Anwendung: lineare Optimierungsproblem (Simplex-Verfahren, George Dantzig 1947) .

Seien $A \in \text{Mat}(m, n, \mathbb{R})$, $b \in \mathbb{R}^m$, und $c \in \mathbb{R}^{n*}$ gegeben.

Eine **zulässige Lösung** ist ein Vektor $x \in \mathbb{R}^n$, der die linearen Bedingungen

$$a_{11}x_1 + \dots + a_{1n}x_n \leq b_1$$

$$a_{21}x_1 + \dots + a_{2n}x_n \leq b_2$$

$$\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots$$

$$a_{m1}x_1 + \dots + a_{mn}x_n \leq b_m$$

erfüllt. Ziel ist es, unter allen zulässigen Vektoren x einen zu finden, der die Linearform

$$c(x) = c_1x_1 + \dots + c_nx_n$$

maximiert. Dieses Optimierungsproblem in der sogenannten Standardform wird oft abkürzend als

$$\max\{c^T x \mid Ax \leq b, \}$$

geschrieben, wobei die Bedingungen $Ax \leq b$ komponentenweise zu verstehen sind.

Darüber hinaus gibt es noch weitere äquivalente Formulierungen, die sich durch einfache Operationen in diese Standardform bringen lassen:

* Minimierungsproblem statt Maximierungsproblem: Multiplikation des Zielsform c mit (-1)

* Größer-gleich- statt Kleiner-gleich-Bedingungen: Multiplikation der entsprechenden Ungleichungen mit (-1)

* Gleichheitsbedingungen statt Ungleichheitsbedingungen: Ersetzung von $a_i x = b_i$ durch $a_i x \leq b_i$ und $-a_i x \leq -b_i$

(Eine ähnliche Aufgabe ist Hausaufgabe.)

Eine Firma stellt zwei verschiedene Produkte her, für deren Fertigung drei Maschinen A, B, C zur Verfügung stehen. Diese Maschinen haben eine maximale monatliche Laufzeit (Kapazität) von 170 Stunden (A), 150 Stunden (B) bzw. 180 Stunden (C). Eine Mengeneinheit (ME) von Produkt 1 liefert einen Deckungsbeitrag von 300 Euro, eine ME von Produkt 2 dagegen 500 Euro. Fertigt man eine ME von Produkt 1, dann benötigt man dafür eine Stunde die Maschine A und eine Stunde die Maschine B . Eine Einheit von Produkt 2 belegt zwei Stunden lang Maschine A , eine Stunde Maschine B und drei Stunden Maschine C . Ziel ist es, Produktionsmengen zu bestimmen, die den Deckungsbeitrag der Firma maximieren, ohne die Maschinenkapazitäten zu überschreiten.

Auf mathematische Sprache umformulieren: Angenommen, der Betrieb fertigt pro Monat x_1 ME von Produkt 1 und x_2 ME von Produkt 2. Dann beträgt der Gesamtdeckungsbeitrag $G(x_1, x_2) = 300x_1 + 500x_2$.

Diesen Wert möchte die Firma maximieren. Da die Maschinenkapazitäten eingehalten werden müssen, ergeben sich die Nebenbedingungen:

$$x_1 + 2x_2 \leq 170$$

$$x_1 + x_2 \leq 150$$

$$3x_2 \leq 180$$

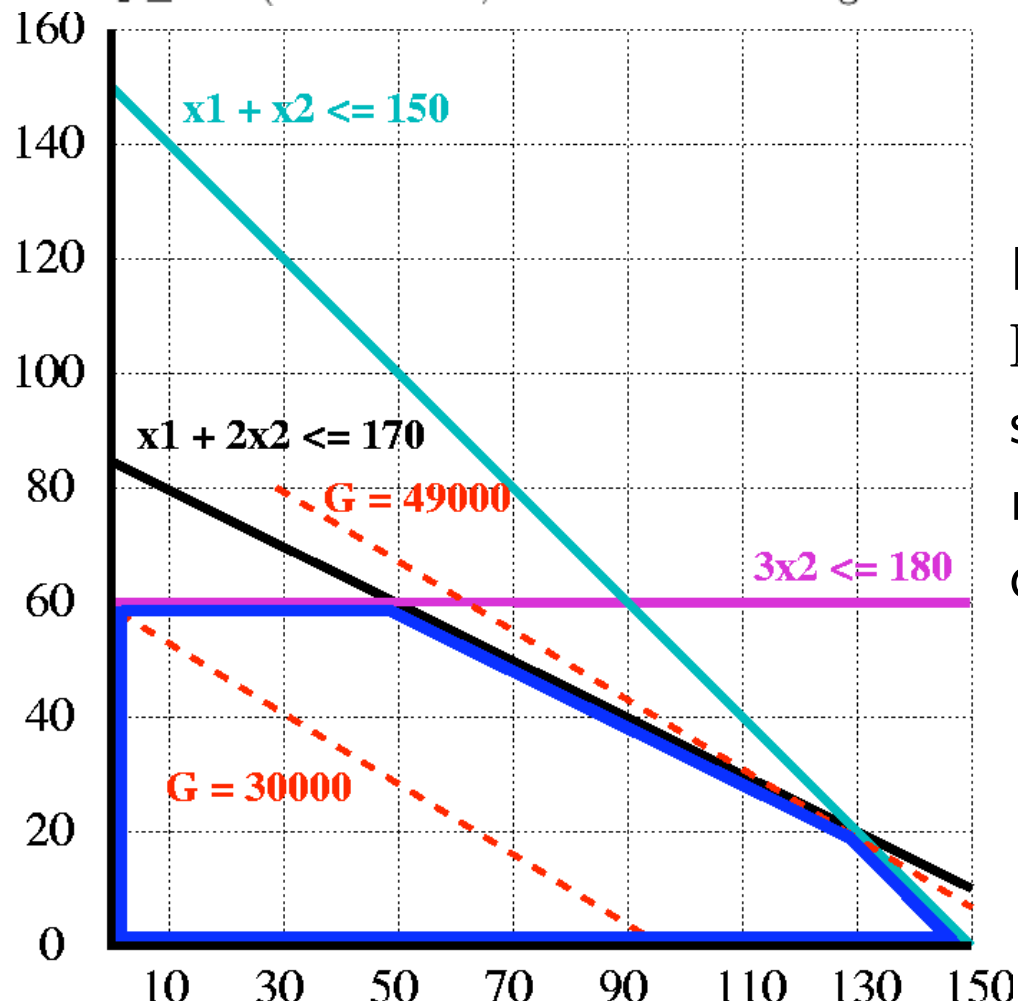
Da außerdem keine negativen Produktionsmengen möglich sind, muss $x_1, x_2 \geq 0$ gelten (Nichtnegativitätsbedingung).

Wenn die Dimensionen n klein ist (z.B., $n = 2$), kann man solche Aufgaben graphisch lösen: wir minimieren $G(x_1, x_2) = 300x_1 + 500x_2$ mit

$x_1 + 2x_2 \leq 170$ (Maschine A, rechts in schwarz eingezeichnet)

$x_1 + x_2 \leq 150$ (Maschine B, rechts in tuerkis eingezeichnet)

$3x_2 \leq 180$ (Maschine C, rechts in violett eingezeichnet)



Die Geraden $\{(x_1, x_2) \in \mathbb{R}^2 \mid G(x_1, x_2) = a\}$ sind rot auf dem Bild; wir müssen solche Gerade finden, sodass a maximal ist.

Umformulierung auf der Sprache der Konvexgeometrie:

Gegeben ist eine polyedrische Menge (= die Menge von zulässigen Lösungen $P := \{x \in \mathbb{R}^n \mid Ax \leq b\}$) und ein $f \in \mathbb{R}$ (oben: $f(x_1, \dots, x_n) = c_1x_1 + \dots + c_nx_n$). Man soll den Punkt $\bar{x} \in P$ finden, sodass $f(\bar{x}) = \sup_{x \in P} f(x)$.

Bemerkung. Die Lösung analytisch zu finden ist nicht möglich. Man ist mit einem Algorithmus zufrieden, das die Lösung in Vernünftige Zeit algorithmisch findet. Gegenstand der lineare Optimierung.

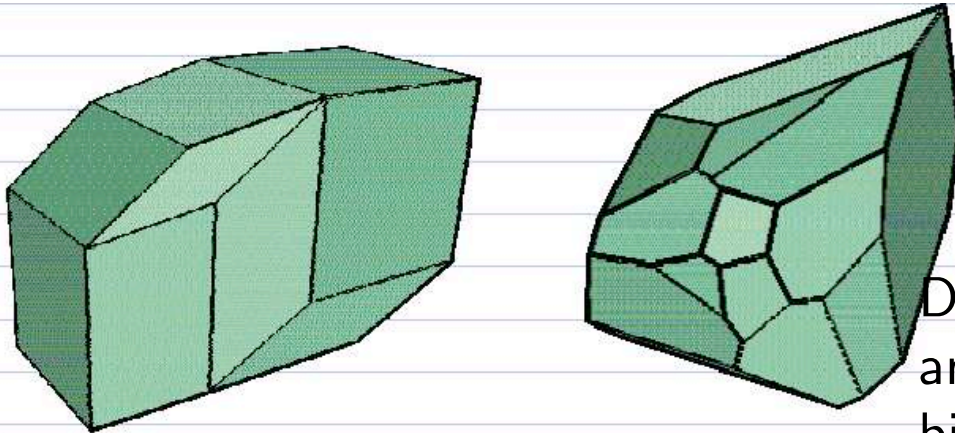
Fallunterscheidung:

1. f sei nach oben beschränkt, P enthalte keine Gerade. In dem Fall gibt es einen **EXTREMPUNKT** \bar{x} , so dass $f(\bar{x}) = \sup_{x \in P} f(x)$ (Satz 38). Man muss den Extrempunkt finden.
2. Es könnte sein, dass die Linearform f nach oben nicht beschränkt ist. Dann gibt es keine Lösung. (es wäre nicht schlecht, zu verstehn, ob der Fall tatsächlich zutrifft.)
3. Es könnte sein, dass die Menge P eine Gerade enthält. Man soll verstehen, dass der Fall tatsächlich zutrifft, und was man weiter machen kann

Bemerkung. Falls die polyedrische Menge P beschränkt ist (was öfter in praktikbezogenen Aufgaben der Fall ist), dann ist die polyedrische Menge P ein Polytop (Satz 37), und es gibt einen **EXTREMPUNKT=ECKE** \bar{x} , so dass $f(\bar{x}) = \sup_{x \in P} f(x)$ (Satz 32).

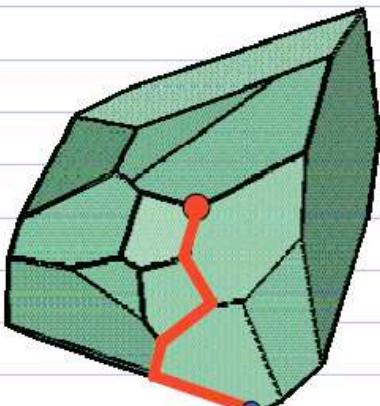
Fall 1. Simplex-Algorithmus (besonders einfach, falls P beschränkt ist)

Die Idee des Simplexalgorithmus ist es, an einer Ecke zu starten und den Zielfunktionswert mit den Zielfunktionswerten alle benachbarten Ecken zu vergleichen. Wird eine Ecke gefunden die einen höheren Zielfunktionswert aufweist, so wird die Prozedur wiederholt.

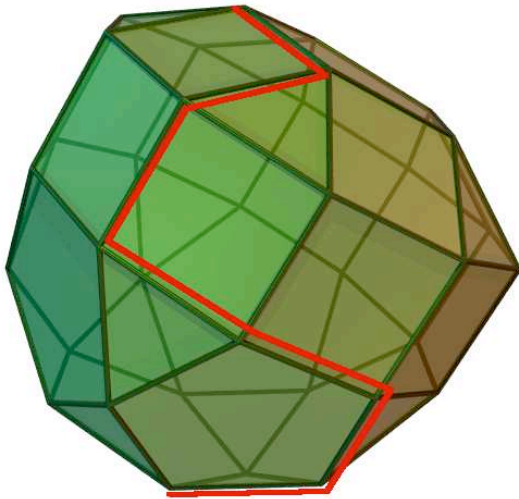


Der Simplexalgorithmus läuft dabei an den Kanten des Polyeders entlang bis er eine optimale Lösung gefunden hat. Der Vorteil liegt auf der Hand. Der Simplexalgorithmus braucht im optimalen Fall nur wenige Ecken zu evaluieren, bis er eine Lösung findet.

- Optimum wird auf Ecke angenommen
- Simplexalgorithmus durchläuft Ecken
-



Simplexalgorithmus



1. Bestimme einen beliebigen Ecke v von P .
2. Falls es keine verbessernde Kante (1-Seite) inzident zu v gibt, stopp. v ist optimal.
3. Folge einer beliebigen verbessernden Kante e von v . Falls e unbeschränkt ist, d.h. keinen anderen Endpunkt hat, stopp. Die Linearform f ist ebenfalls auf P unbeschränkt.
4. Sei u der andere Endpunkt von e . Setze $v = u$. Gehe zurück zu Schritt 2.

Um das algorithmisch zu machen, braucht man:

1. Methode, eine Ecke zu finden

(Theoretisch Einfach: Man betrachtet eine Gerade, läuft entlang der Geraden bis zum Rand. Weil Rand aus polyedren Mengen der kleineren Dimension besteht, wiederholt man das Prozedur, bis zum Man in einer Ecke kommt)

2. Methode, alle Kanten aus der gegebenen Ecke v zu finden:

(Theoretisch Einfach: Man betrachtet alle Hyperebenen, $a_{i1}x_1 + \dots + a_{in}x_n = b_i$ die den Punkt v enthalten. Dann untersucht man, ob Schnittmenge von irgendwelchen $n - 1$ (unabhängigen) davon mit P nicht leer ist)

3. Methode, um zu entscheiden, ob eine von Kanten verbessernde Kante ist: Alle Kanten auszuprobieren

Die Menge von zulässigen Lösungen enthält eine Gerade.

Folgerung Sei $P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$ ein nichtleere polyedrische Menge.

Dann sind äquivalent:

1. P hat mindestens einen Eckpunkt.
2. P enthält keine Gerade.
3. $\text{rang}(A) = n$.

Beweis. (1) \iff (2) haben wir in Bemerkung nach Satz 38 bewiesen.

Wir zeigen (3) \implies (2) durch Widerspruch.

Falls P eine Gerade enthält, dann enthält jeder Halbraum

$a_{i1}x_1 + \dots + a_{in}x_n \leq b_i$ eine Gerade (z.B. $\mathcal{G}_{x,\vec{v}}$); dann muss diese Gerade nach Lemma 10 parallel zur Hyperebene $\{a_{i1}x_1 + \dots + a_{in}x_n = b_i\}$ sein;

also muss $a_{i1}v_1 + \dots + a_{in}v_n = 0$ sein. Weil dass für alle i erfüllt ist, liegt \vec{v} in $\text{Kern}(A)$, und $\text{Rank}(A) < n$.

Wir zeigen (2) \implies (3). Angenommen, P enthält keine Gerade. Dann ist $\text{Kern}(A) = \{\vec{0}\}$ nach Satz 27 (a) LAAG 1 (weil ein Punkt von P plus alle Vektoren aus Kern wieder in P liegen soll.) Dann ist

$$\text{rk}(A) \stackrel{\text{LAAG I}}{=} \dim(\text{Bild}_A) = n - \dim(\text{Kern}(A)) = n.$$

Das dritte Hilbertsche Problem



David Hilbert

- 23.1.1862 (Königsberg)
– 14.2.1943 (Göttingen)
- Gründungsmitglied der DMV (1890)

Hat in 1900 eine Liste aus 23 wichtigste Probleme der Mathematik veröffentlicht.

Das dritte Hilbertsche Problem (1900) und dessen Motivation.

Erstes gelöstes Problem:

3. Die Volumengleichheit zweier Tetraeder von gleicher Grundfläche und Höhe.

..., zwei Tetraeder mit gleicher Grundfläche und von gleicher Höhe anzugeben, die sich auf keine Weise in congruente Tetraeder zerlegen lassen und die sich auch durch Hinzufügung congruenter Tetraeder nicht zu solchen Polyedern ergänzen lassen, für die ihrerseits eine Zerlegung in congruente Tetraeder möglich ist.

Man kann das Problem auch wie folgt umformulieren: Sind je zwei Polyeder mit gleichem Volumen zerlegungsgleich, das heißt kann man immer ein Polyeder in polyedrische Teile schneiden und das andere Polyeder aus diesen Teilen zusammensetzen?

Motivation.

- ▶ In dim 2 ist hat das Problem keine Lösung (Elementare Geometrie; Gauß).
- ▶ Falls die Problem keine Lösung hat, kann man viel einfacher den Begriff „Volum“ in dim 3 definieren.

Dieses Problem wurde von Max Dehn im selben Jahr 1900 gelöst. Der Beweis von Dehn war sehr kompliziert. Später stellte es sich heraus, dass man eine einfachere Lösung mittels Lineare Algebra konstruieren kann. Wir machen dies in der nächste Woche. Dazu brauchen wir **Das Auswahlaxiom**

Exkurs in Mengenlehre: Das Auswahlaxiom

Def. 35 Sei A eine Menge von nichtleeren Mengen. Dann heißt $F : A \rightarrow \bigcup_{\alpha \in A} \alpha$ eine **Auswahlfunktion** für A (mit dem Definitionsbereich A und Wertebereich Vereinigung von allen Elementen aus A), falls gilt:
 $\forall X \in A : F(X) \in X$.

F wählt also aus jeder Menge X in A genau ein Element aus.

Das **Auswahlaxiom** lautet dann wie folgt:

Zu jeder Menge nichtleerer Mengen gibt es mindestens eine Auswahlfunktion.

Das Auswahlaxiom postuliert die Existenz einer Auswahlfunktion, gibt jedoch kein Verfahren an, wie man eine solche konstruieren könnte.

Beispielsweise ist es nicht allgemein möglich, für eine beliebige Menge von Teilmengen \mathbb{R} eine Auswahlfunktion explizit anzugeben.

Das Auswahlaxiom ist von der überwiegenden Mehrheit der Mathematiker akzeptiert. Es folgt nicht von anderen Axiomen der Mathematik. Es gibt Zweigen der Mathematik (z.B. Die Konstruktivistische Mathematik), die auf das Auswahlaxiom verzichtet.

Beispiele

Für welche Fälle das Auswahlaxiom relevant ist, sei an den folgenden Beispielen verdeutlicht:

* Für eine endliche Menge $A = \{A_1, \dots, A_n\}$ von nichtleeren Mengen ist es trivial, eine Auswahlfunktion anzugeben: Man wählt von jeder Menge irgendein bestimmtes Element aus, was problemlos möglich ist. Man braucht das Auswahlaxiom hierfür nicht. Ein formaler Beweis würde Induktion über die Größe der endlichen Menge verwenden.

* Für Mengen von nichtleeren Teilmengen der natürlichen Zahlen ist es ebenfalls problemlos möglich: Man wählt von jeder Teilmenge das kleinste Element aus. Ähnlich kann man für eine Menge von abgeschlossenen Teilmengen der reellen Zahlen eine explizite Auswahlfunktion (ohne Verwendung des Auswahlaxioms) angeben, indem man etwa aus jeder Menge das (wenn möglich positive) Element mit kleinstem Absolutbetrag wählt.

Beispiele

- * Selbst für Mengen von Intervallen reeller Zahlen ist eine Auswahlfunktion definierbar: Man wählt von jedem Intervall den Mittelpunkt aus.
- * Für Mengen von beliebigen nichtleeren Teilmengen der reellen Zahlen gibt es jedoch keine offensichtliche Definition einer Auswahlfunktion. In diesem Fall ist das Auswahlaxiom relevant. Es postuliert die Existenz einer Auswahlfunktion, ohne sie anzugeben. Man kann sogar beweisen, dass man sie nicht für allen Mengen von beliebigen nichtleeren Teilmengen der reellen Zahlen konstruieren kann. Beweisidee: Es gibt nur abzählbare viel ($= |\mathbb{N}|$) Sätze (in einer Sprache), und deswegen nur abzählbare viel Punkten beschreiben werden kann.

Wir (eigentlich, Sie, das waren Hausaufgaben) haben das Auswahlaxiom bereits zweimal benutzt

Einmal im Beweis von Lemma 2 in LAAG I:

Lemma 2 $f : A \rightarrow B$ sei eine Abbildung und $A \neq \emptyset$. Dann gilt:

(1) f ist injektiv $\iff f$ hat eine Linksinverse.

(2) f ist surjektiv $\iff f$ hat eine Rechtsinverse.

Wiederholung. Eine Rechtsinverse Abbildung (zu $f : A \rightarrow B$) ist eine Abbildung $g : B \rightarrow A$ mit $f \circ g = Id_B$

Beweis von (1), und von der Richtung \Leftarrow von (2) war sauber.

Beweis: f ist surjektiv $\implies f$ hat eine Rechtsinverse.

Wiederholung. Eine **rechtsinverse** Abbildung (zu $f : A \rightarrow B$) ist eine Abbildung $g : B \rightarrow A$ mit $f \circ g = Id_B$

Beweis: f sei surjektiv vorausgesetzt. Sei $x \in B$. Die gesuchte **rechtsinverse Abbildung** $g : A \rightarrow B$ wird nun definiert durch $g(x) := y$, wobei y irgendwelcher Punkt aus der Urbildmenge $Urbild_f(y)$.

Definition ist nur dann korrekt, wenn wir aus der Menge $Urbild_f(y)$ einen Punkt auswählen können; z.B. falls wir Auswahlaxiom akzeptieren.

Dann gilt für $\forall x \in B: f \circ g(x) = f(g(x)) = f(y) \stackrel{\text{Def. von } y}{=} x$.

Bemerkung. Das Auswahlaxiom ist zur Existenz von der rechtinversen Abbildung zu eine beliebige Abbildung äquivalent.

Wir (eigentlich, Sie, das waren Hausaufgaben) haben das Auswahlaxiom bereits zweimal benutzt

Zweites Mal in der Hausaufgabe 1a, Blatt 1: eine Menge ist genau dann unendlich, wenn sie zu einer ihrer echten Teilmengen gleichmächtig ist.

Lösung der Hausaufgabe in \Rightarrow Richtung. Wir beweisen, dass eine disjunkte Folge $x_1, \dots, x_k, \dots \in M$ existiert, und dann definieren die Abbildung

$$f : M \rightarrow M \setminus \{x_1\}, \quad \begin{cases} f(x) = x & \text{für } x \neq x_i, i \in \mathbb{N} \\ f(x_i) = x_{i+1} & \text{für } x_i \text{ aus der Folge} \end{cases}$$

Dann gilt:

f ist injektiv; die Bildmenge Bild_f ist $M \setminus \{x_1\}$

\Rightarrow

gilt $f : M \rightarrow M \setminus \{x_1\}$ bijektiv und die Mengen M und $M \setminus \{x_1\}$ sind gleichmächtig.

Um eine solche Folge zu konstruieren, braucht man aber die Auswahlaxiome: die folgende Konstruktion war vorgeplant:
wir **nehmen** einen Punkt $x_1 \in M$. Da M unendlich ist, ist $M \setminus \{x_1\} \neq \emptyset$;
nehmen wir $x_2 \in M \setminus \{x_1\}$; dann $x_3 \in M \setminus \{x_1, x_2\}$ u.s.w.

Es ist gefährlich, mit Mengen naiv umzugehen:
Man bekommt sofort Widerspruch, wenn man sich zuviel erlaubt:

RUSSELL–ZERMELOsches Paradoxon (I)

Sei $R := \{X \mid X \text{ ist Menge und } X \notin X\}$ die Menge aller Mengen, die sich nicht selbst enthalten.

Dann gilt für eine beliebige Menge Y :

$$Y \in R \quad \Leftrightarrow \quad Y \notin Y$$

Insbesondere gilt dann für $Y = R$:

$$R \in R \quad \Leftrightarrow \quad R \notin R$$

Widerspruch !!!

Halbordnung/Ordnung auf der Menge

Wiederholung. Eine **Relation** auf einer Menge M ist eine Teilmenge $R \subseteq M \times M$. (Statt $(x, y) \in R$ schreibt man $x R y$.)

Def. 36 Eine Relation \leq heißt **Halbordnung**, falls sie (für alle $a, b \in M$)

transitiv ist: aus $a \leq b$ und $b \leq c$ folgt $a \leq c$, und

reflexiv ist: $a \leq a$.

antisymmetrisch ist: aus $a \leq b$ und $b \leq a$ folgt $a = b$.

Bsp. Die übliche \leq ist eine Halbordnung auf \mathbb{R} :

Bsp. Die übliche \leq ist eine Halbordnung auf \mathbb{C} : sie besteht aus allen Paaren (x, y) wobei $x, y \in \mathbb{R}$. Z.B. ist weder $(1 + i, i)$ noch $(i, 1 + i)$ in der Relation \leq

Exkurs: Kardinalzahlen

Wir betrachten die Äquivalenzrelation „gleichmächtig“ auf (der Menge von) allen Mengen, siehe Vorl. 1, 2. Die Äquivalenzklassen bzgl. diese Relation heißen **Kardinalzahlen**.

Falls die Menge endlich ist, kann man deren Kardinalzahl mit Anzahl von Elemente identifizieren: die endliche Mengen sind genau dann gleichmächtig, falls sie aus gleichen Anzahl von Elementen bestehen.

Die Kardinalzahl \aleph_0 ist die Äquivalenzklasse von \mathbb{N} . In Vorl. 1 haben wir bewiesen, dass $\mathbb{Q} \in \aleph_0$. (Man schreibt auch $|\mathbb{Q}| = \aleph_0$.)

Die Kardinalzahl \aleph_1 ist die Äquivalenzklasse von \mathbb{R} . In Vorl. 2 haben wir bewiesen, dass $2^{\mathbb{Q}} \in \aleph_1$ (oder, man schreibt $|2^{\mathbb{Q}}| = \aleph_1$.)

Bsp. Die Relation „höchstens gleichmächtig“ ist eine Halbordnung auf der Menge von Kardinalzahlen. In der Tat, sie ist

transitiv: gibt es Injektionen

$\phi : A \rightarrow B$ und $\psi : B \rightarrow C$, so ist die Verkettung $\psi \circ \phi : A \rightarrow C$ eine Injektion.

Also, aus $\aleph_A \leq \aleph_B$ und $\aleph_B \leq \aleph_C$ folgt $\aleph_A \leq \aleph_C$.

reflexiv: $|A| = |A|$,

antisymmetrisch: gibt es Injektionen $\phi : A \rightarrow B$ und $\psi : B \rightarrow A$, dann ist $|A| = |B|$ nach Def. 2

Def. 36 – Voraussetzung Eine Halbordnung „ \leq “ heißt eine **Ordnung**, falls zusätzlich für jede $a, b \in M$ gilt $a \leq b$, oder $b \leq a$.

Bsp. Die übliche „ \leq “ ist nicht nur eine Halbordnung, sondern auch eine Ordnung auf \mathbb{R} .

Bsp. Die Halbordnung „ \leq “ ist keine Ordnung auf \mathbb{C} . (weil weder $i \leq i + 1$ noch $i + 1 \geq i$).

Wir sagen, dass eine Teilmenge $T \subseteq M$ eine **total geordnete Menge** bzg. einer Halbordnung „ \leq “ ist, falls der Halbordnung „ \leq “ beschränkt auf T eine Ordnung ist: für jedes Paar $(x, y) \in T \times T$ gilt $x \leq y$ oder $y \leq x$.

Bsp. In der Halbgeordneten Menge (\mathbb{C}, \leq) ist $\mathbb{R} \subseteq \mathbb{C}$ total geordnet.

Def. 36 – Voraussetzung

Wir sagen, dass eine Teilmenge $T \subseteq M$ **eine obere Schranke** hat, falls es ein $m \in M$ gibt, so dass $x \leq m$ für alle $x \in T$.

Wir sagen, dass ein Element $m \in T$ **maximal** ist, wenn es kein $a \in T$ gibt mit $m \leq a$.

Ein Element o von M heißt **kleinstes Element**, wenn für alle $m \in M$ gilt: $o \leq m$.

Bsp. Ein offenes Intervall auf \mathbb{R} hat kein maximales Element, hat aber eine obere Schranke. Ein abgeschlossenes Intervall hat immer ein maximales Element.

Bemerkung. Ein maximales Element muss keine obere Schranke von M sein. Das wäre ein größtes Element. Z.B., falls die Relation R leer ist, ist jedes Element ein maximales Element, weil weder $a R b$ noch $b R a$ gilt. Also, maximales Element und kleinstes Element sind keine Antonymen.

Def. 36 – Vorsetzung Eine obere Schranke m heißt eine **kleinste obere Schranke** von $S \subseteq M$, wenn für alle oberen Schranken \tilde{m} von S gilt: $m \leq \tilde{m}$. Wenn eine kleinste obere Schranke existiert, dann ist sie eindeutig bestimmt.

Wenn für jede total geordnete Teilmenge von M eine obere Schranke existiert, dann heißt M **induktiv geordnet**. Wenn sogar jeweils eine kleinste obere Schranke existiert, dann heißt M **strikt induktiv geordnet**.

Bsp. Sei $M \subseteq 2^A$. Wir betrachten die folgende Relation (**Inklusionsrelation**) auf M :

$X \leq Y \iff X \subseteq Y$. Eine total geordnete Teilmenge ist eine Familie $T \subseteq 2^A$ von Teilmengen von A , sodass für je zwei verschiedene Elemente $B, C \in T$ entweder $B \subseteq C$ oder $C \subseteq B$ gilt. Wenn nun die Vereinigung aller Mengen aus T auch in M liegt (das ist nicht selbstverständlich!), dann hat man eine (und sogar die kleinste) obere Schranke gefunden.

Fixpunktsatz

Satz 42 (Fixpunktsatz) Es sei (M, \leq) eine nichtleere geordnete Menge mit einem kleinsten Element o . In M habe jede total geordnete Menge T eine kleinste obere Schranke. Sei $F : M \rightarrow M$ eine Abbildung mit der Eigenschaft

$\forall m \in M : m \leq F(m)$. Dann gibt es ein $m \in M$ mit $F(m) = m$.

Beweis. Wie nennen eine Teilmenge S von M **zulässig**, wenn die folgenden drei Bedingungen gelten: $o \in S$, $\text{Bild}_F(S) \subseteq S$ und für jede total geordnete Teilmenge $T \subseteq S$ liegt auch die kleinste obere Schranke \bar{T} (von T) in S . Zum Beispiel ist M selbst zulässig.

Nun sei S_0 der Durchschnitt aller zulässigen Teilmengen von M . S_0 ist auch zulässig: $o \in S_0$;

$$\text{Bild}_F(S_0) = \text{Bild}_F \left(\bigcap_{S \text{ zulässig}} S \right) = \bigcap_{S \text{ zulässig}} \text{Bild}_F(S) \subseteq \bigcap_{S \text{ zulässig}} S = S_0,$$

und wenn T in allen S liegt, dann liegt auch \bar{T} in allen S , schließlich auch in S_0 .

Bemerkung. Wir haben dieselbe Idee früher, insbes. in Vorl. 12 mehrmals benutzt, siehe Def. von lineare, Hülle, konvexe Hüllen.

Also, S_0 ist selbst zulässig und damit die kleinste aller zulässigen Teilmengen von M .

Wir zeigen, dass S_0 total geordnet ist; daraus folgt für die kleinste obere Schranke \bar{S}_0 einerseits,

(a) dass \bar{S}_0 das größte Element von S_0 ist.

(b) Andererseits gilt aber wegen der Zulässigkeit $F(\bar{S}_0) \leq \bar{S}_0$.

Wir bekommen insgesamt $\bar{S}_0 \stackrel{(a)}{\leq} F(\bar{S}_0) \stackrel{(b)}{\leq} \bar{S}_0$, und damit die gewünschte Gleichheit.

Noch zu zeigen ist also die folgende Behauptung

Behauptung: S_0 ist total geordnet.

Behauptung: S_0 ist total geordnet.

Beweis der Behauptung. Wir nennen $e \in S_0$ ein **extremales Element**, wenn für alle $s \in S_0$ mit $\underbrace{s \leq e, s \neq e}_{s < e}$ gilt, dass $F(s) \leq e$. Zum Beispiel

ist das kleinste Element o extremal. Für ein extremales e setzen wir

$$S_e := \{s \in S_0 \mid s \leq e \text{ oder } F(e) \leq s\}.$$

Dann ist für jedes extremale e die Menge S_e zulässig:

- o liegt in S_e .
- Für jedes Element $s \in S_e$ folgt aus $s < e$ sofort $F(s) \leq e$ nach Def. von extremalem Element,

aus $s = e$ folgt $F(s) = F(e) \leq e$ (nach Voraussetzungen),

und aus $s \not\leq e$ folgt $F(e) \leq s \leq F(s)$.

Also gilt insgesamt $\text{Bild}_F(S_e) \subseteq S_e$.

- Es sei T eine total geordnete Teilmenge von S_e . Wenn dann für alle $t \in T$ die Ungleichung $t \leq e$ gilt, dann gilt auch die kleinste obere Schranke $\bar{T} = e$. Wenn es aber mindestens ein t gibt, sodass die Ungleichung $t \leq e$ nicht gilt, dann ist $e \leq F(t) \leq F(\bar{T})$. Wir sehen aus diesen beiden Fällen: $\bar{T} \in S_e$.

Da aber S_0 die kleinste zulässige Teilmenge von M ist, muss also für alle extremalen e gelten: $S_e = S_0$.

Nun müssen wir noch zeigen, dass jedes $e \in S_0$ extremal ist. Dann folgt nämlich für $s \in S_0$:

$$s \in S_e \text{ also } s \leq e \text{ oder } e \leq F(e) \leq s.$$

Das sagt dann, dass S_0 total geordnet ist.

Um zu beweisen, dass jedes $e \in S_0$ extremal ist, betrachten wir $E := \{e \in S_0 \mid e \text{ ist extremal}\}$.

Nun weisen wir nach, dass E zulässig ist, und damit gleich S_0 .
 $o \in E$: klar.

Abgeschlossenheit von E unter F :

$$\forall e \in E, \forall s \in S_0 = S_e, s < F(e) \text{ gilt: } F(s) \leq F(e).$$

Diese letzte Ungleichung gilt, da $F(e) \not\leq s$ und damit $s \leq e$. Nun greift die Extremalität von e zu.

Nun sei noch $T \subseteq E$ total geordnet, und seine kleinste obere Schranke sei \bar{T} . Zu zeigen ist $\bar{T} \in E$. Sei dazu $s \in S_0, s < \bar{T}$. Wenn für jedes $t \in T$ die Relation $F(t) \leq s$ gelten würde, dann wäre \bar{T} als obere Schranke von extremalen Elementen selbst $\leq s$ – Widerspruch.

Also gibt es ein extremales $e \in T$ mit $F(e) \neq s$, und da $S_0 = S_e$ gilt folgt daraus zwangsweise $s \leq e$. Aber $s \neq e$ impliziert dann $F(s) \leq e \leq \bar{T}$, und aus $s = e$ folgt $F(s) = F(e) \in E$, denn E ist unter F abgeschlossen. Also ist auch in diesem Fall $F(s) \leq \bar{T}$. Damit folgt insgesamt, dass \bar{T} extremal ist

Satz 43 (Lemma von Zorn) Jede nichtleere halbgeordnete Menge, in der jede total geordnete Teilmenge eine obere Schranke hat, enthält mindestens ein maximales Element (falls das Auswahlaxiom gilt).

Beweis. (a) Wir behandeln zuerst den Fall einer strikt induktiv geordneten Menge. (**Wiederhol.** – wenn für jede total geordnete Teilmenge von M eine kleinste obere Schranke existiert, dann heißt M strikt induktiv geordnet.)

Da es eben langt, ein maximales Element zu finden, das $\leq x$ für ein willkürlich gewähltes x ist, dürfen wir uns auf den Fall beschränken, dass M ein kleinstes Element enthält. Dann nehmen wir an, es gebe kein maximales Element. Wir finden also für jedes $m \in M$ ein größeres Element $F(m)$ und definieren damit eine Funktion $F : M \rightarrow M$, für die gilt:

$$\forall m \in M : m < F(m).$$

Bemerkung. Wir brauchen dafür das Auswahlaxiom. Die Funktion F ist wie im Satz 42, deswegen muss F einen Fixpunkt haben: Widerspruch.

Nun sei M induktiv geordnet und H die Menge aller total geordneten Teilmengen von M . Dann ist H bezüglich der Inklusion geordnet, und zwar strikt induktiv, denn die Vereinigung einer total geordneten Familie von total geordneten Teilmengen ist wieder eine total geordnete Teilmenge und offensichtlich die kleinste obere Schranke der Familie (bezüglich Inklusion). Insbesondere besitzt H nach (a) ein maximales Element T (das ist eine total geordnete Teilmenge von M). Es sei O eine obere Schranke von T . Dann muss O schon zu T gehören, da $T \cup O$ eine total geordnete Menge ist, die T enthält, aber T ist schon maximal. Dieses Element O ist dann ein maximales Element in M , denn für jedes $m \in M$ folgt aus $O \leq m$, dass m eine obere Schranke von T ist und somit ebenfalls zu T gehören muss, also insbesondere $m \leq O$ und damit $m = O$.

Hamel-Basen

Def. 37 Eine Hamel-Basis von \mathbb{K} -Vektorraum $(V, +, \cdot)$ ist eine Teilmenge $B \subseteq V$, so dass

- sie **linear unabhängig** ist (d.h., das Null-Element $\vec{0}$ nur als triviale **ENDLICHE** Linearkombination der Elemente von B dargestellt werden kann.)
- sie ist **erzeugend** : Man kann jedes Element von V als **ENDLICHE** Linearkombination der Elemente von B darstellen.

Bemerkung. In Satz 25 LAAG I haben wir bewiesen, dass (••) zur folgenden Aussage äquivalent ist: Die Menge B ist **maximal**: wenn wir noch ein Element hinzufügen, wird die Menge linear abhängig.

Falls der Vektorraum V endlich erzeugt ist, ist Hamel-Basis fast dasselbe wie eine Basis: wenn wir die Elemente von B als ein Tupel schreiben, bekommen wir eine Basis, und umgekehrt.

Satz 44. Jeder Vektorraum hat eine Hamel-Basis

Beweis. Sei V ein Vektorraum. Wir betrachten:

$$P := \{X \subseteq V \mid X \text{ linear unabhängig}\} \subseteq 2^V.$$

Die Menge P ist bezüglich der Relation „ \subseteq “ halbgeordnet. Es gilt:

1. Besteht V nicht nur aus dem Nullvektor, dann ist P nicht leer (weil jede Einermenge $\{v\}$ mit $v \in V$ und $v \neq \vec{0}$ ein Element von P ist).

2. Für jede **total geordnete Teilmenge** $T \subseteq P$ ist auch

$$\bigcup T := \bigcup_{X \in T} X = \{v \mid \exists X \in T : v \in X\} \text{ in } P.$$

In der Tat, die Menge $\bigcup T$ ist linear unabhängig: endlich viel Vektoren

$\underbrace{v_1}_{\in X_1}, \dots, \underbrace{v_k}_{\in X_k}$ liegen alle in der grössten von X_1, \dots, X_k ; (**existiert, weil T**

total geordnet ist) dann sind sie nach Konstruktion linear unabhängig.

Also, T hat eine obere Schranke.

Aus dem Lemma von Zorn folgt nun, dass P ein maximales Element hat.

Die maximalen Elemente von P sind nun aber genau die maximalen (in Sinne der Bemerkung oben) linear unabhängigen Teilmengen von V , also die Basen von V . Daher hat V eine Basis und es gilt darüber hinaus,

dass jede linear unabhängige Teilmenge von V in einer Basis von V enthalten ist,

Bsp. Da \mathbb{Q} ein Unterkörper von \mathbb{R} ist, ist \mathbb{R} ein \mathbb{Q} -Vektorraum. Nach Satz 1 ist er unendlichdimensional. Er hat trotzdem eine Hamel-Basis, d.h., die Menge $B \subseteq \mathbb{R}$ sodass jedes $a \in \mathbb{R}$ eine eindeutige Linearkombination von der Elementen von B ist. Diese Menge B kann man nicht konstruieren, wir kennen nur die Existenz.

Die Basis B ist nicht explizit gegeben, und ist mit Hilfe von Auswahlfunktion konstruiert. Wir können sie nicht explizit geben: „Explizit“ bedeutet, dass wir die Basiselemente in Worten beschreiben.

Aber: Kardinalität von allen möglichen Sätze ist $\aleph_0 = |\mathbb{N}|$ (auch wenn wir die unendlichen Sätze erlauben), aber die Kardinalität von B ist $\aleph_1 = |\mathbb{R}|$.

Frage. Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ erfülle

$$f(x + y) = f(x) + f(y). \quad (*)$$

Ist sie linear?

(Umformulieren: folgt $f(\lambda \cdot x) = \lambda f(x)$ aus $(*)$?)

Angenommen $f(1) = \alpha \neq 0$. Dann ist $f(2) = f(1 + 1) \stackrel{(*)}{=} 2\alpha$, induktiv $f(n) = n \cdot \alpha$.

Dann ist $f(0 + 1) \stackrel{(*)}{=} f(0) + f(1) = f(1) \implies f(0) = 0$.

Analog $f(-n)$ ist die Zahl sodass $f(-n) + f(n) = f(0) = 0$; schliesslich $f(-n) = -n\alpha$

Analog $f(p/q)$ ist die Zahl sodass

$$q \cdot f(p/q) = \underbrace{f(p/q) + \dots + f(p/q)}_{q \text{ Stück}} \stackrel{(*)}{=} \underbrace{f(p/q + \dots + p/q)}_{q \text{ Stück}} = f(p) =$$

$$p\alpha \implies f(p/q) = p/q \cdot \alpha.$$

Also, $f(x)$ fällt mit der (linearen Funktion) $\alpha \cdot x$ auf \mathbb{Q} zusammen.

Insbesondere gilt: wenn wir zusätzlich verlangen, dass f stetig ist, dann ist sie linear.

Es gibt aber nichtlineare Funktionen mit der Bedingung (*):

Man betrachte z.B. zwei Zahlen $a, b \in \mathbb{R}$, die über \mathbb{Q} linear unabhängig sind (z.B. $a = 1$ und $b = \sqrt{2}$), und eine Hamel-Basis B in $\underbrace{\mathbb{R}}_{\text{ein } \mathbb{Q}\text{-Vektorraum}}$,

sodass $a, b \in B$ sind.

Wiederholung. Um eine lineare Abbildung zu definieren, brauchen wir nur die Bilder von Basiselementen einzugeben. (Satz 30 LAAG I)

Wir wählen $\alpha, \beta \in \mathbb{R}$ und setzen $f(a) = \alpha$, $f(b) = \beta$ und $\forall x \in B \setminus \{a, b\}$ setzen wir $f(x) := 0$.

Die Abbildung ist linear (über \mathbb{Q}).

Sie erfüllt deswegen $f(x + y) = f(x) + f(y)$, wie wir wollen.

Bemerkung. Man kann selbstverständlich mehr als zwei Basisvektor benutzen, sogar \aleph_0 Vektoren.

Was kennen wir über f

Def von f : $f(a) = \alpha$, $f(b) = \beta$ und $\forall x \in B \setminus \{a, b\} f(x) := 0$.

Also, wir können nur die Werte von Funktion f auf der Zahlen der Form $\frac{p_1}{q_1}a + \frac{p_2}{q_2}b$ berechnen (wobei $\frac{p_i}{q_i} \in \mathbb{Q}$);

$f\left(\frac{p_1}{q_1}a + \frac{p_2}{q_2}b\right) = \frac{p_1}{q_1}\alpha + \frac{p_2}{q_2}\beta$. Dies reicht, um 3. Problem von Hilbert zu lösen.

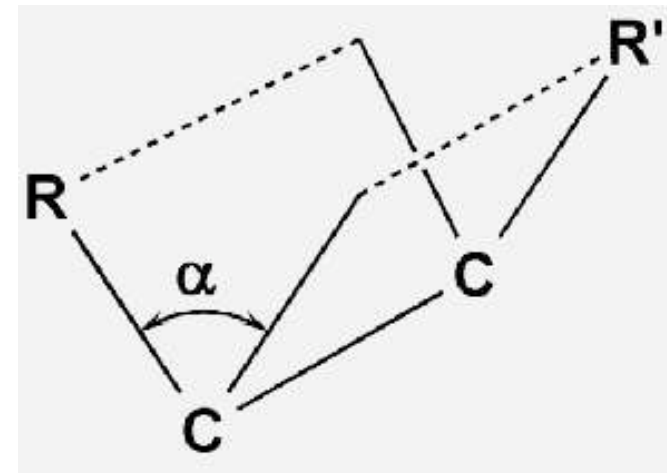
Wiederholung: Das dritte Hilbertsche Problem (1900): Sind je zwei 3-dim Polyeder mit gleichem Volumen **zerlegungsgleich**, das heißt kann man immer ein Polyeder in polyedrische Teile schneiden und das andere Polyeder aus diesen Teilen zusammensetzen?

Antwort: Nein: Sogar ein Würfel kann man nicht in polyedrische Teile schneiden und dann ein regelmäßiges Tetraeder aus diesen Teilen zusammensetzen.

Beweis. Sei \mathcal{P} die Menge aller Polyedra. Wir betrachten die folgende Abbildung $F : \mathcal{P} \rightarrow \mathbb{R}$:

$$F(P) := \sum_{\substack{l \text{ ist} \\ \text{eine Kante} \\ \text{von } P}} |l| \cdot f(\phi(l)),$$

wobei $|l|$ ist die Länge von der Kante l , $\phi(l)$ ist der Diederwinkel um Kante l , s. Bild.



Und f ist die oben konstruierte Abbildung (mit der Eigenschaft $f(x + y) = f(x) + f(y)$); die Daten $a, b, \alpha, \beta \in \mathbb{R}$ sind wie folgt:

a ist der Diederwinkel von Würfel (d.h., $a = \pi/2$)

b ist der Diederwinkel vom regelmäßigen Tetraeder (die Zahlen a und b sind über \mathbb{Q} linearunabhängig; wir werden es nicht beweisen).

$\alpha = 0$ und $\beta = 1$

Die Funktion F vom Würfel und vom regelmäßigen Tetraeder

Wir können die Funktion f nur für einige Einträge ausrechnen, und zwar nur für Linearkombinationen von a und b . Das reicht, um $F(\text{Würfel})$ und $F(\text{Reg. Tetraeder})$ auszurechnen.

$$F(\text{Würfel}) = \sum_{\substack{l \text{ ist} \\ \text{eine Kante}}} |l| \underbrace{\alpha}_{=0} = 0 ;$$

$$F(\text{Reg. Tetraeder}) = \sum_{\substack{l \text{ ist} \\ \text{eine Kante}}} |l| \underbrace{\beta}_{=1} > 0.$$

Wir sehen, dass $F(\text{Würfel}) \neq F(\text{Reg. Tetraeder})$

Das folgende Behauptung löst das 3. Problem von Hilbert:

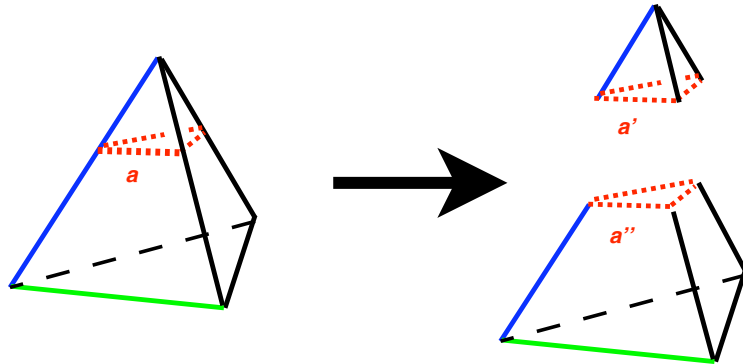
Behauptung Sind zwei Polyeder P_1 und P_2 zerlegungsgleich, so gilt $F(P_1) = F(P_2)$.

Lösung des 3. Problem von Hilbert. Da

$F(\text{Würfel}) \neq F(\text{Reg. Tetraeder})$ ist, sind Würfel und reg. Tetraeder nicht zerlegungsgleich.

Behauptung Sind zwei Polyeder P_1 und P_2 zerlegungsgleich, so gilt $F(P_1) = F(P_2)$.

Beweis der Behauptung: Wenn wir ein Polyeder (T auf dem Bild) mit einer Ebene in zwei Polyedra (T_1, T_2) schneiden, gilt $F(T) = F(T_1) + F(T_2)$.



In der Tat, der Ausdruck $|l|f(\phi(l))$, der grüne Kante entspricht, bleibt unverändert.

Der Ausdruck $|l|f(\phi(l))$, der blaue Kante entspricht, ist die Summe von Ausdrücke der blauen Kanten von T_1 und T_2 .

Die Summe von Ausdrücke für die neuen kanten a' und a'' ist 0, weil $|a'| = |a''| = |a|$, und

$$|a'|f(\phi(a')) + |a''|f(\phi(a'')) \stackrel{(*)}{=} |a|f(\phi(a') + \phi(a'')) = |a|f(\pi) = 2|a|f(\pi/2) = 0.$$

Da man OBdA annehmen kann, dass P_1 mit Hilfe von endlich viel Anwendungen von schneiden in zwei Polyedra zerlegt ist, ist die Behauptung bewiesen.

Neuer (kurzer) Abschnitt: Transformationsgruppen (Einführung)

- ▶ Sei \mathcal{S}_M die Menge aller Bijektionen a von M auf sich (heissen oft **Selbstabbildungen**).
- ▶ \mathcal{S}_M ist eine Gruppe bzgl. Verkettungen, siehe Satz 3 Vorl. 4 LAAG I.
- ▶ **Transformationsgruppe** (informell) ist eine Untergruppe von \mathcal{S}_M , die irgendwelche Eigenschaften von Objekten von M nicht verändert.

Affine Transformationen

Zwei Definitionen:

Ursprüngliche Def. aus LAAG I; und auch aus LAAG II, siehe

Def. 13 Eine affine Transformation, oder Affinität, (von $M = \mathbb{R}^n$) ist eine Abbildung der Form $f(x) = Ax + b$, wobei $A \in GL(n, \mathbb{R})$.

Folgerung aus Satz 22 (Fundamentalsatz der affinen Geometrie). Eine affine Transformation ist eine Bijektion $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, die Geraden auf Geraden abbildet.

Euklidische Transformationen

Wiederholung. Euklidische Räume sind affine Räume über Vektorräumen mit Skalarprodukt. In dem Fall ist der Abstand zwischen zwei beliebigen Punkten erklärt ist: $d(x, y) = |\overrightarrow{xy}| \stackrel{\text{in } \mathbb{R}^n}{=} |x - y|$.

$a : \mathbb{R}^n \rightarrow \mathbb{R}^n$ heißt eine **Isometrie**, wenn für beliebige $x, y \in \mathbb{R}^n$ gilt:
 $d(x, y) = d(a(x), a(y))$.

Bsp. Wir betrachten \mathbb{R}^n mit Standard-Skalarprodukt $\langle \cdot, \cdot \rangle$. Dann sind die Abbildungen

$F_{O,b} : V \rightarrow V$, $F_{O,b} := Ox + b$, (*)
wobei O eine orthogonale $(n \times n)$ -Matrix und $b \in \mathbb{R}^n$ ist, (bijektive)
Isometrien (bzgl. Standard-Metrik).

Satz 68 LAAG I. *Jede Isometrie eines endlichdimensionalen Euklidischen Raums ist wie in (*)*

Das heißt Im Standard-Raum \mathbb{R}^n mit Standard-Skalarprodukt:
jede Isometrie g kann man als

$F \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = O \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ schreiben,

wobei O eine orthogonale Matrix ist.

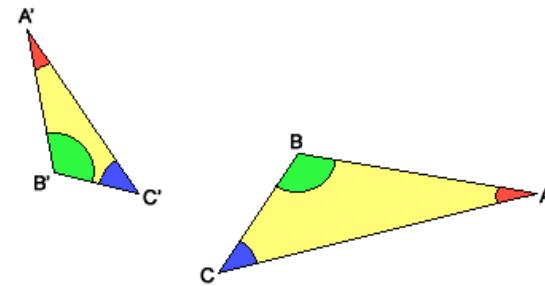
In Worten Jede Isometrie (der Standard-Metrik) ist Verkettung von
Drehung und Verschiebung

Noch ein Bsp.: Ähnlichkeitstransformationen in Elementargeometrie

Schuldefinition: **Ähnlichkeitsabbildungen** sind Abbildungen, die die Gestalt von Figuren erhalten.

Def. 38 Wir betrachten \mathbb{R}^n mit dem Standard-Skalarprodukt $\langle \cdot, \cdot \rangle$.

Eine Bijektion a von \mathbb{R}^n heißt ein **Ähnlichkeitstransformation**, wenn es eine reelle Zahl $k > 0$ gibt, sodass für beliebige $x, y \in \mathbb{R}^n$ gilt:
 $d(a(x), a(y)) = k \cdot d(x, y)$.



Ähnliche Dreiecke

k heißt Ähnlichkeitsfaktor von a .

Bemerkung. Isometrien sind Ähnlichkeitsabbildungen mit $k = 1$.

Folgerung. $a(x) = Cx + c$ mit $C = k \cdot O$, wobei $O \in O(n)$.

Beweis. die Abbildung \tilde{a} , $\tilde{a}(x) = \frac{1}{k}a(x)$ ist eine Isometrie.

Folgerung. (Hausaufgabe) Ähnlichkeitsabbildungen sind die Bijektionen, die gleichzeitig Geradentreu sind (d.h., Bild jeder Gerade ist eine Gerade) und Winkeltreu (d.h., Winkel zwischen Geraden \mathcal{G}_1 und \mathcal{G}_2 ist gleich den Winkel zwischen Geraden $Bild_a(\mathcal{G}_1)$ und $Bild_a(\mathcal{G}_2)$) sind.

Alle oben genannten Transformationsgruppen kann man aus Untergruppen von Matrizingruppen Darstellen:

Wir werden dies in dim 2 erklären, und zuerst nur für Ähnlichkeitsabbildungen obwohl die Dimension nicht wichtig ist: Wir betrachten die erweiterte Koordinaten“

$\begin{pmatrix} x \\ y \end{pmatrix}$ Ersetzt \leftrightarrow $\begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$. In diesen Koordinaten gilt:

- **Drehung um 0-Punkt:**

$$\begin{bmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}$$

- **Skalierung:**

$$\begin{bmatrix} k & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} kx \\ ky \\ 1 \end{bmatrix}$$

- **Translation:**

$$\begin{bmatrix} 1 & 0 & x_v \\ 0 & 1 & y_v \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x + x_v \\ y + y_v \\ 1 \end{bmatrix}$$

- **Spiegelung:**

$$\begin{bmatrix} \cos(\alpha) & \sin(\alpha) & 0 \\ \sin(\alpha) & -\cos(\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}$$

Da jede oben genannte Transformationen (in dim 2) Verkettungen von den Abbildungen aus der Liste links ist, sind die oben genannten Gruppe von Transformationen (isomorph zu) Untergruppen von $GL(n+1, \mathbb{R})$

=3

\implies Jede Transformation wird (in den erweiterten Koordinaten“)
zu einer Matrizenmultiplikation mit der Matrix der Form

$$\begin{pmatrix} \boxed{k \mathbf{O}} & \begin{matrix} a_{13} \\ a_{23} \\ 1 \end{matrix} \\ \begin{matrix} 0 & 0 \end{matrix} & \end{pmatrix}, \text{ wobei } O \in O(2).$$

Wir haben diese Formel bereits benutzt, in Vorl. 25– 26, siehe
Hauptsätze der Theorie der Quadriken (Normalformen).

Analog gilt: Matrizen der Form $\bar{A} = \begin{pmatrix} \boxed{A} & \begin{matrix} a_{13} \\ a_{23} \\ 1 \end{matrix} \\ \begin{matrix} 0 & 0 \end{matrix} & \end{pmatrix}$, wobei

$A \in GL(2, \mathbb{R})$ repräsentieren alle affinen Transformationen von \mathbb{R}^2 .

Der projektive Raum $P(V)$ und projektive Transformationen

$(V, +, \cdot)$ sei ein Vektorraum (über \mathbb{K}).

Def. 39. Der zu V gehörige projektive Raum $P(V)$ ist definiert als die Menge aller 1-dimensionalen Untervektorräume von V . Ist $\dim(V) < \infty$, so setzt man $\dim(P(V)) := \dim(V) - 1$. (Nach Def. ist $P(\{\vec{0}\}) = \emptyset$ und $P(\{\vec{0}\}) = -1$ – wir können damit leben).

Bemerkung. Man hat eine kanonische Abbildung $V \setminus \{\vec{0}\} \rightarrow P(V), v \mapsto \mathcal{G}_v := \{\lambda v \mid \lambda \in \mathbb{K}\}$.

Homogene Koordinaten

Def. 40 Sei $V = \mathbb{K}^{n+1}$. Dann sind die homogenen Koordinaten von $v = (x_0, x_1, \dots, x_n) \in V \setminus \{\vec{0}\}$ definiert durch

$$(x_0 : x_1 : \dots : x_n) := \mathcal{G}_v := \{\lambda v \mid \lambda \in \mathbb{K}\}.$$

Es gilt:

$$\begin{aligned} \mathcal{G}_v = \mathcal{G}_{v'} &\iff \exists \lambda \in \mathbb{K} \setminus 0 \text{ mit } v' = \lambda v \\ &\iff \exists \lambda \in \mathbb{K} \setminus 0 \text{ mit } (x'_0, \dots, x'_n) = (\lambda x_0, \dots, \lambda x_n) \end{aligned}$$

Die homogenen Koordinaten sind also nur bis auf einen gemeinsamen Faktor $\lambda \neq 0$ aus \mathbb{K} festgelegt.

Wiederholung.

Def. 39. Der zum Vektorraum V gehörige projektive Raum $P(V)$ ist definiert als die Menge aller 1-dimensionalen Untervektorräume von V .

Falls $V = \mathbb{K}^{n+1}$ ist, bezeichnet man $P(V)$ oft mit $P^n(\mathbb{K})$ (s.d. $P(\mathbb{K}^{n+1}) = P^n(\mathbb{K})$), oder sogar mit $\mathbb{P}\mathbb{K}^n$.

Bemerkung. Man hat eine kanonische Abbildung $V \setminus \{\vec{0}\} \rightarrow P(V)$, $v \mapsto \mathcal{G}_v := \{\lambda v \mid \lambda \in \mathbb{K}\} := \mathbb{K}v$.

Def. 40 Sei $V = \mathbb{K}^{n+1}$. Dann sind die homogenen Koordinaten von $v = (x_0, x_1, \dots, x_n) \in V \setminus \{\vec{0}\}$ definiert durch

$$(x_0 : x_1 : \dots : x_n) := \mathcal{G}_v := \{\lambda v \mid \lambda \in \mathbb{K}\}.$$

Die homogenen Koordinaten sind also nur bis auf einen gemeinsamen Faktor $\lambda \neq 0$ aus \mathbb{K} festgelegt.

Eine Funktion $f : \mathbb{K}^{n+1} \rightarrow \mathbb{K}$ heißt **homogen** mit Grad $k \in \mathbb{Z}$, falls $f(\lambda x_0, \dots, \lambda x_n) = \lambda^k f(x_0, \dots, x_n)$ für alle $(x_0, \dots, x_n) \in \mathbb{K}^{n+1}$ und $\lambda \in \mathbb{K}$.

Bsp. Eine lineare Funktion $f \in (\mathbb{K}^{n+1})^*$ ist 1-homogen.
 f^2 ist 2-homogen.

Bemerkung. Sei f eine homogene Funktion. Liegt (x_0, \dots, x_n) in der Lösungsmenge der Gleichung $f(x) = 0$, so liegen alle Punkte der Gerade $\mathbb{K}(x_0, \dots, x_n)$ in der Lösungsmenge der Gleichung $f(x) = 0$. Also, die Gleichung $f(x) = 0$ kann man als Gleichung auf $P(\mathbb{K}^{n+1})$ verstehen.

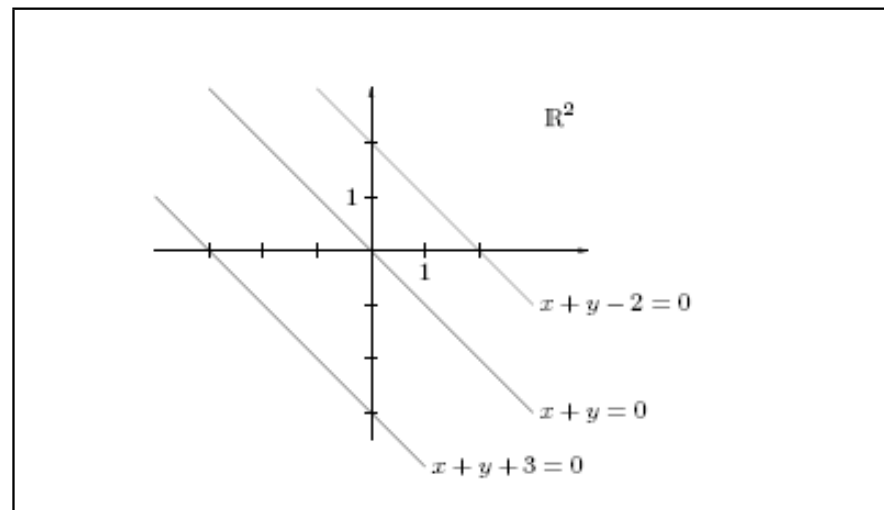
Eine Funktion $f : \mathbb{K}^{n+1} \rightarrow \mathbb{K}$ heißt **homogen** mit Grad $k \in \mathbb{Z}$, falls $f(\lambda x_0, \dots, \lambda x_n) = \lambda^k f(x_0, \dots, x_n)$ für alle $(x_0, \dots, x_n) \in \mathbb{K}^{n+1}$ und $\lambda \in \mathbb{K}$

Bsp. Eine lineare Funktion $f \in (\mathbb{K}^{n+1})^!$ ist 1-homogen.
 f^2 ist 2-homogen.

Bemerkung. Sei f eine homogene funktion. Liegt (x_0, \dots, x_n) in der Lösungsmenge der Gleichung $f(x) = 0$, so liegen alle Punkte der Gerade $\mathbb{K}(x_0, \dots, x_n)$ in der Lösungsmenge der Gleichung $f(x) = 0$.

Beispiel.

Betrachte in \mathbb{R}^2 die parallelen Geraden
 $x + y - 2 = 0, x + y = 0, x + y + 3 = 0.$



Homogenisierung liefert:

$$x + y - 2 = 0 \quad \rightsquigarrow \quad -2x_0 + x_1 + x_2 = 0$$

$$x + y = 0 \quad \rightsquigarrow \quad 0 \cdot x_0 + x_1 + x_2 = 0$$

$$x + y + 3 = 0 \quad \rightsquigarrow \quad 3x_0 + x_1 + x_2 = 0$$

In allen drei Fällen ist $(0 : 1 : -1)$ eine Lösung in $\mathbb{P}^2(\mathbb{R})$. Die Geraden haben also einen Schnittpunkt in $\mathbb{P}^2(\mathbb{R})$.

Projektive Geraden in $P^2(\mathbb{K})$

Eine **projektive Gerade** $G \in P^2(\mathbb{K})$ ist die Nullstellenmenge in $P^2(\mathbb{K})$ einer Gleichung

$$ax_0 + bx_1 + cx_2 = 0 \quad \text{mit} \quad (a, b, c) \neq (0, 0, 0).$$

Es ist also

$$G = \{(x_0 : x_1 : x_2) \in P^2(\mathbb{K}) \mid ax_0 + bx_1 + cx_2 = 0\}.$$

Ist $b = c = 0$, so ist G die **unendlich ferne Gerade**

$$P^1 := \{(x_0 : x_1 : x_2) \in P^2(\mathbb{K}) \mid x_0 = 0\}.$$

Die übrigen Geraden in $P^2(\mathbb{K})$ erhält man, wie in Bsp. oben beschrieben, durch Homogenisierung und durch Hinzufügen des **unendlich fernen Punktes** $(0 : c : -b) \in P^1(\mathbb{K})$.

Ist $c \neq 0$, so ist $U = \{(x_0, x_1, x_2) \in \mathbb{K}^3 \mid ax_0 + bx_1 + cx_2 = 0\}$ ein

Untervektorraum von \mathbb{K}^3 mit Basis $\left(\begin{pmatrix} 1 \\ 0 \\ -a \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -b \end{pmatrix} \right)$ und es ist $G = P(U)$.

Projektive Unterräume in $P(V)$

Def. 41 Eine Teilmenge $X \subseteq P(V)$ heißt projektiver Unterraum, falls $X = P(U)$ mit einem Untervektorraum U von V gilt. Es ist dann

- X eine projektive Gerade in $P(V)$, falls $\dim(X) = 1$
- eine projektive Hyperebene in $P(V)$, falls $\dim(X) = \dim(P(V)) - 1$ und $\dim V < \infty$.

Beispiele.

Seien $(X_i = P(U_i))_{i \in I}$ projektive Unterräume von $P(V)$, dann ist

$$\bigcap_{i \in I} X_i = P\left(\bigcap_{i \in I} U_i\right)$$

ein projektiver Unterraum.

Man nennt den kleinsten projektiven Unterraum von $P(V)$, der $\bigcup_{i \in I} X_i$ enthält, den **Verbindungsraum** $\bigvee_{i \in I} X_i$ (das ist ein Analog von Hülle). Es ist

$$\bigvee_{i \in I} X_i = P\left(\sum_{i \in I} U_i\right)$$

Dimensionssatz

Satz 45. Sei $\dim(P(V)) < \infty$. Für projektive Unterräume $X_1 = P(U_1)$ und $X_2 = P(U_2)$ gilt

$$\dim(X_1 \vee X_2) = \dim X_1 + \dim X_2 - \dim(X_1 \cap X_2)$$

Beweis. Es ist

$$\begin{aligned} \dim(X_1 \vee X_2) &= \dim(U_1 + U_2) - 1 && \text{nach Def. 39 und Def. 41.} \\ &= \dim(U_1) + \dim(U_2) - \dim(U_1 \cap U_2) - 1 && \text{nach Aufgabe 3 Blatt 7} \\ &= \dim(X_1) + 1 + \dim(X_2) + 1 - (\dim(X_1 \cap X_2) + 1) - 1 && \text{nach Def. 39 und Def. 41.} \\ &= \dim(X_1) + \dim(X_2) - \dim(X_1 \cap X_2) \end{aligned}$$



Schnittpunktsatz

Satz 46. Sei $\dim(P(V)) < \infty$. Dann gelten:

(i) Ist $\dim(X_1) + \dim(X_2) \geq \dim(P(V))$ für zwei projektive Unterräume X_1, X_2 von $P(V)$, so ist $X_1 \cap X_2 \neq \emptyset$.

(ii) Ist $X_1 = G$ eine projektive Gerade in $P(V)$, $X_2 = H$ eine projektive Hyperebene in $P(V)$, und gilt $G \not\subseteq H$, dann schneiden sich G und H in genau einem Punkt $p \in P(V)$.

(iii) Zwei projektive Geraden in $P^2(\mathbb{K})$ schneiden sich stets.

Beweis (i): $\dim(X_1 \cap X_2) \stackrel{\text{Satz 45}}{=} \dim(X_1) + \dim(X_2) - \dim(X_1 \vee X_2)$
 $\geq \dim(X_1) + \dim(X_2) - \dim(P(V)) \geq 0 \stackrel{\text{Def. 39}}{\implies} X_1 \cap X_2 \neq \emptyset$.

Beweis (ii): Aus $G \not\subseteq H$ folgt $\dim(X_1 \vee X_2) = \dim(P(V))$ und somit $\dim(X_1 \cap X_2) = 0$. Es ist also $X_1 \cap X_2 = \{p\}$ ein Punkt.

(iii) folgt aus (ii), da die Hyperebene H eine projektive Gerade in $P(\mathbb{K}^2)$ ist, □

Projektiver Abschluss von $P(\mathbb{K}^n)$

Sei $V = \mathbb{K}^{n+1}$. Betrachte die Abbildung

$$\psi : \mathbb{K}^n \rightarrow P(V), (x_1, \dots, x_n) \mapsto (1 : x_1 : \dots : x_n) = \mathbb{K}v \text{ mit } v = (1, x_1, \dots, x_n)$$

Es ist $U = \{(x_0, x_1, \dots, x_n) \in \mathbb{K}^{n+1} \mid x_0 = 0\}$ ein n -dimensionaler Untervektorraum von \mathbb{K}^{n+1} , also ist $H := P(U)$ eine Hyperebene in $P(V)$.

Behauptung $\text{Bild}_\psi = P(V) \setminus H$.

Beweis. Nach Definition gilt: $\text{Bild}_\psi \subseteq P(V) \setminus H$.

Sei $(y_0 : y_1 : \dots : y_n) \in P(V) \setminus H$, also $y_0 \neq 0$.

$$\implies (y_0 : y_1 : \dots : y_n) = \left(1 : \frac{y_1}{y_0} : \dots : \frac{y_n}{y_0}\right) = \psi \left(\frac{y_1}{y_0}, \dots, \frac{y_n}{y_0}\right), \quad \square$$

Man nennt H die **unendlich ferne Hyperebene**, heißt kanonische Einbettung von \mathbb{K}^n in $P(\mathbb{K}^{n+1})$, und $P(\mathbb{K}^{n+1})$ wird als **projektiver Abschluss von \mathbb{K}^n bezeichnet**. Man schreibt auch $P(\mathbb{K}^{n+1}) = \mathbb{K}^n \cup A_\infty$, wobei $A_\infty = H$ ist. Für $n = 1$ schreibt man dann speziell $P(\mathbb{K}^2) = \mathbb{K}^1 \cup \infty$, da H dann ein Punkt ist.

Projektivitäten

Def. 42. Seien V, W zwei $(n + 1)$ -dimensionale \mathbb{K} -Vektorräume. Dann heißt eine bijektive Abbildung $! : P(V) \rightarrow P(W)$ eine **Projektivität**, falls es eine bijektive \mathbb{K} -lineare Abbildung

$$\vec{\varphi} : V \longrightarrow W \text{ gibt mit } \varphi(P) = \vec{\varphi}(P) \quad \forall P \in P(V)$$

(Hierbei ist P ein Punkt von $P(V)$ und also ein eindimensionaler Teilraum von V .) Es ist dann $\vec{\varphi}$ *nicht* eindeutig durch φ bestimmt, denn für $\lambda \in K^*$ ist $(\lambda\vec{\varphi})(P) = \vec{\varphi}(\lambda P) = \vec{\varphi}(P)$.

Ferner gilt $\boxed{\varphi(P(U)) = P(\vec{\varphi}(U))}$ für jeden Untervektorraum U von V und jede Projektivität φ .

Bemerkung. Wir haben oben die folgende Bezeichnung benutzt:

$$!(P) = \text{Bild}_{\varphi}(\mathbb{K}P).$$

Kollineationen

Def. 43 Eine bijektive Abbildung $\varphi : P(V) \rightarrow P(W)$ heißt **Kollineation**, wenn das Bild einer projektiven Geraden G durch die Punkte $P, Q \in \subseteq P(V)$ die projektive Gerade durch die Punkte $\varphi(P), \varphi(Q) \in P(W)$ ist.

Satz 47. Es gilt:

- (i) Zu je zwei Punkten $P, Q \in P(V)$ mit $P \neq Q$ gibt es genau eine projektive Gerade G , die P und Q enthält.
- (ii) Ist $\varphi : P(V) \rightarrow P(W)$ eine Projektivität, so ist φ eine Kollineation.

Beweis.

(i) Es ist $P = \mathbb{K}v_1$ und $Q = \mathbb{K}v_2$ mit $v_1, v_2 \in V \setminus \vec{0}$. Da $P \neq Q$ ist folgt, dass v_1 und v_2 linear unabhängig sind. Es ist also $g = P(U)$ mit $U = \mathbb{K}v_1 + \mathbb{K}v_2 :=$ die gesuchte projektive Gerade.

(ii) Es ist $\varphi(G) = P(\vec{\varphi}(U))$. Daraus folgt die zweite Behauptung, □

Weitere Beispiele zur Homogenisierung

Analog konstruiert man zu einer Hyperebene

$$X = \{(x_1, \dots, x_n) \in \mathbb{K}^n \mid a_1 x_1 + \dots + a_n x_n + b = 0\}$$

durch Homogenisierung ihrer definierenden Gleichung den Abschluss

$$\bar{X} = \{(y_0 : y_1 : \dots : y_n) \in P(\mathbb{K}^{n+1}) \mid b y_0 + a_1 y_1 + \dots + a_n y_n = 0\}.$$

(in $P(\mathbb{K}^{n+1})$). Man setze $x_i = \frac{y_i}{y_0}$ für alle $i = 1, \dots, n$.)

Analog erhält man durch Homogenisierung den projektiven Abschluss von Quadriken, Kegelschnitten, etc.

Beispiele.

Betrachte die definierenden Gleichungen **Homogenisierung** Mit $x = \frac{x_1}{x_0}$ und $y = \frac{x_2}{x_0}$ erhält man:

i) $x^2 + y^2 - 1 = 0$ Kreis

i) $-x_0^2 + x_1^2 + x_2^2 = 0$

ii) $x^2 - y^2 - 1 = 0$ Hyperbel

ii) $-x_0^2 + x_1^2 - x_2^2 = 0$

iii) $x^2 - y = 0$ Parabel

iii) $x_1^2 - x_2 x_0 = 0$

Behauptung. Kreis, Hyperbel und Parabel sind projektiv äquivalent, d.h. die zugehörigen projektiven Kurven gehen durch eine Projektivität $\varphi : P(\mathbb{R}^3) \rightarrow P(\mathbb{R}^3)$ ineinander über.

Beispiele.

Betrachte die definierenden Gleichungen Homogenisierung Mit $x = \frac{x_1}{x_0}$ und $y = \frac{x_2}{x_0}$ erhält man:

i) $x^2 + y^2 - 1 = 0$ Kreis

i) $-x_0^2 + x_1^2 + x_2^2 = 0$

ii) $x^2 - y^2 - 1 = 0$ Hyperbel

ii) $-x_0^2 + x_1^2 - x_2^2 = 0$

iii) $x^2 - y = 0$ Parabel

iii) $x_1^2 - x_2x_0 = 0$

Beweis. Durch $(x_0 : x_1 : x_2) \mapsto (x_1 : x_0 : x_2)$ geht die projektive Quadrik mit der Gleichung (ii) in die projektive Quadrik mit der Gleichung (i) über. Durch $(x_0 : x_1 : x_2) \mapsto (x_0 + x_2 : x_1 : x_0 - x_2)$ erhält man aus Gleichung (iii) die Gleichung (i), □

Bemerkung. Ist $H = P(U)$ eine projektive Hyperebene in $P(V)$, so ist $\mathcal{A} := P(V) \setminus H$ ein affiner Raum mit zugehörigem Vektorraum U , und es ist $A_\infty = H$

Frage: Was geschieht dabei mit einer projektiven Quadrik in $P(V)$?

Beispiel. Sei $Q = \{(x_0 : x_1 : x_2) \in P(\mathbb{R}^3) \mid -x_0^2 + x_1^2 + x_2^2 = 0\}$.

(1) Sei $H = \{(x_0 : x_1 : x_2) \in P(\mathbb{R}^3) \mid x_0 = 0\}$ und

$$\psi : \mathbb{R}^2 \rightarrow P(\mathbb{R}^2) \setminus H, \quad (x_1, x_2) \mapsto (1 : x_1 : x_2).$$

Dann ist $Q \cap H = \emptyset$, und $\text{Urbild}_\psi(Q) = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 + x_2^2 - 1 = 0\}$ ist ein Kreis. Ersetzt man \mathbb{R} durch \mathbb{C} , so besteht $Q \cap H$ aus zwei Punkten.

(2) Sei $H = \{(x_0 : x_1 : x_2) \in P(\mathbb{R}^3) \mid x_1 = 0\}$. Dann besteht $Q \cap H$ aus zwei Punkten, und für

$\psi : \mathbb{R}^2 \rightarrow P(\mathbb{R}^3) \setminus H, \quad (x_0, x_2) \mapsto (x_0 : 1 : x_2)$ ist

$\text{Urbild}_\psi(Q) = \{(x_0, x_2) \in \mathbb{R}^2 \mid x_0^2 - x_2^2 - 1 = 0\}$ eine Hyperbel.

(3) Sei $H = \{(x_0 : x_1 : x_2) \in P(\mathbb{R}^3) \mid x_0 + x_1 = 0\}$. Dann ist $Q \cap H$ ein Punkt, und für

$\psi : \mathbb{R}^2 \rightarrow P(\mathbb{R}^3) \setminus H, \quad (x_1, x_2) \mapsto ((1 - x_1) : x_1 : x_2)$ ist

$\text{Urbild}_\psi(Q) = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_2^2 + 2x_1 - 1 = 0\}$ eine Parabel. (Die Substitution $x_1 = -z_1 + \frac{1}{2}$ ergibt $\frac{1}{2}x_2^2 - z_1 = 0$.) (Je nachdem, was man als unendlich ferne Hyperebene auszeichnet, erhält man aus Q die drei affinen Kurven: Kreis, Hyperbel, Parabel)

Explizite Beschreibung von Projektivitäten

Sei $V = \mathbb{K}^{n+1}$ und $\varphi : P(V) \rightarrow P(V)$ eine Projektivität. Bezüglich der Standardbasis wird $\vec{\varphi} : V \rightarrow V$ durch eine Matrix

$$A := \begin{pmatrix} a_{00} & a_{01} & \dots & a_{0n} \\ \vdots & \vdots & & \vdots \\ a_{n0} & a_{n1} & \dots & a_{nn} \end{pmatrix} \in GL_{n+1}(K)$$

beschrieben. Es ist dann $\varphi(x_0 : x_1 : \dots : x_n) = (y_0 : y_1 : \dots : y_n)$ mit $y_i = a_{i0}x_0 + \dots + a_{in}x_n$. Übergang zu inhomogenen Koordinaten ergibt

$$y'_i = \frac{y_i}{y_0} = \frac{a_{i0}x_0 + \dots + a_{in}x_n}{a_{00}x_0 + \dots + a_{0n}x_n} = \frac{a_{i0} + a_{i1}x'_1 + \dots + a_{in}x'_n}{a_{00} + a_{01}x'_1 + \dots + a_{0n}x'_n}$$

mit $x'_i = \frac{x_i}{x_0}$ und $x_0 \neq 0$. Die Abbildung

$$\varphi' : \mathbb{K}^n \rightarrow \mathbb{K}^n, \quad (x'_1, \dots, x'_n) \mapsto (y'_1, \dots, y'_n)$$

ist auf der affinen Hyperebene

$\mathcal{E} = \{(x'_1, \dots, x'_n) \in \mathbb{K}^n \mid a_{00} + a_{01}x'_1 + \dots + a_{0n}x'_n = 0\}$ nicht definiert. Die

Punkte von \mathcal{E} werden auf die unendlich ferne Hyperebene abgebildet.

Ist speziell $n = 1$ und $P(\mathbb{K}^1) = \mathbb{K} \cup \infty$, so ist mit $x'_1 =: x$ und $y'_1 =: y$ eine Projektivität gegeben durch

$$x \mapsto y := \frac{a_{11}x + a_{10}}{a_{01}x + a_{00}}$$

(„Möbiustransformation“). Der Punkt $x = -\frac{a_{00}}{a_{01}}$ wird auf ∞ abgebildet, falls $a_{01} \neq 0$. Definiere $\infty \mapsto -\frac{a_{00}}{a_{01}}$.

Wiederholung.

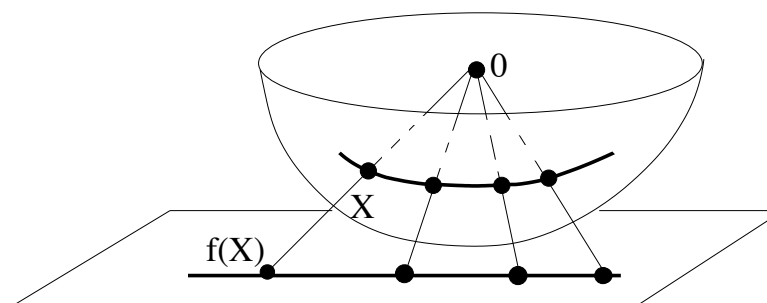
Def. 39. Der zum Vektorraum V gehörige projektive Raum $P(V)$ ist die Menge aller 1-dimensionalen Untervektorräume von V .

Def. 40 Auf $P(V)$ sind **die homogenen Koordinaten** (von dem Punkt $\mathbb{K}v \in P(V)$) definiert durch

$$(x_0 : x_1 : \dots : x_n) := \mathcal{G}_v := \{\lambda v \mid \lambda \in \mathbb{K}\}.$$

In der Nähe des Punktes s.d. $x_0 \neq 0$, kann man die inhomogene Koordinaten (x'_1, \dots, x'_n) definieren: $(x'_1, \dots, x'_n) = \left(\frac{x_1}{x_0}, \dots, \frac{x_n}{x_0}\right)$.

Geometrische Vorstellung der inhomogenen Koordinaten: Man betrachte die Projektion f , die Gerade durch $\vec{0}$ in Schnittpunkt der Gerade mit der Ebene $\{(x_0, \dots, x_n) \mid x_0 = 1\}$ überführt. Dann sind die Koordinaten von $f(x)$ die inhomogene Koordinaten der Geraden.



Projektivitäten

Def. 42. Seien V, W zwei $(n + 1)$ -dimensionale \mathbb{K} -Vektorräume. Dann heißt eine bijektive Abbildung $\varphi : P(V) \rightarrow P(W)$ eine **Projektivität**, falls es eine bijektive \mathbb{K} -lineare Abbildung

$$\vec{\varphi} : V \longrightarrow W \text{ gibt mit } \varphi(P) = \vec{\varphi}(P) \quad \forall P \in P(V)$$

(Hierbei ist P ein Punkt von $P(V)$ und also ein eindimensionaler Teilraum von V .) Es ist dann $\vec{\varphi}$ *nicht* eindeutig durch φ bestimmt, denn für $\lambda \in K^*$ ist $(\lambda\vec{\varphi})(P) = \vec{\varphi}(\lambda P) = \vec{\varphi}(P)$.

Ferner gilt $\boxed{\varphi(P(U)) = P(\vec{\varphi}(U))}$ für jeden Untervektorraum U von V und jede Projektivität φ .

Bemerkung. Wir haben oben die folgende Bezeichnung benutzt:

$$\vec{\varphi}(P) = \text{Bild}_{\vec{\varphi}}(\mathbb{K}P).$$

Def. 42 in koordinaten:

Sei $V = \mathbb{K}^{n+1}$ und $\varphi : P(V) \rightarrow P(V)$ eine Projektivität. Bezüglich der Standardbasis wird $\vec{\varphi} : V \rightarrow V$ durch eine Matrix

$$A := \begin{pmatrix} a_{00} & a_{01} & \dots & a_{0n} \\ \vdots & \vdots & & \vdots \\ a_{n0} & a_{n1} & \dots & a_{nn} \end{pmatrix} \in \text{GL}_{n+1}(K)$$

beschrieben.

In Worten: Eine Projektivität $\varphi : P(\mathbb{K}^{n+1}) \rightarrow P(\mathbb{K}^{n+1})$ ist eine Abbildung, die auf den homogenen Koordinaten durch Multiplikation mit einer invertierbaren $(n+1) \times (n+1)$ -Matrix definiert ist.

Es ist dann $\varphi(x_0 : x_1 : \dots : x_n) = (y_0 : y_1 : \dots : y_n)$ mit $y_i = a_{i0}x_0 + \dots + a_{in}x_n$. Übergang zu inhomogenen Koordinaten ergibt

$$y'_i = \frac{y_i}{y_0} = \frac{a_{i0}x_0 + \dots + a_{in}x_n}{a_{00}x_0 + \dots + a_{0n}x_n} = \frac{a_{i0} + a_{i1}x'_1 + \dots + a_{in}x'_n}{a_{00} + a_{01}x'_1 + \dots + a_{0n}x'_n}$$

mit $x'_i = \frac{x_i}{x_0}$ und $x_0 \neq 0$. Die Abbildung

$$\varphi' : \mathbb{K}^n \rightarrow \mathbb{K}^n, \quad (x'_1, \dots, x'_n) \mapsto (y'_1, \dots, y'_n)$$

ist auf der affinen Hyperebene

$\mathcal{E} = \{(x'_1, \dots, x'_n) \in \mathbb{K}^n \mid a_{00} + a_{01}x'_1 + \dots + a_{0n}x'_n = 0\}$ nicht definiert. Die Punkte von \mathcal{E} werden auf die unendlich ferne Hyperebene abgebildet.

Bsp. Zentralprojektion

Sei V ein $(n + 1)$ -dimensionaler \mathbb{K} -Vektorraum und $X = P(U)$ ein projektiver Unterraum von $P(V)$. Ferner seien $P(U_1)$ und $P(U_2)$ zwei m -dimensionale projektive Unterräume von $P(V)$. Es gelte:

(i) $X \cap P(U_1) = X \cap P(U_2) = \emptyset$

(ii) $X \vee P(U_1) = X \vee P(U_2) = P(V)$

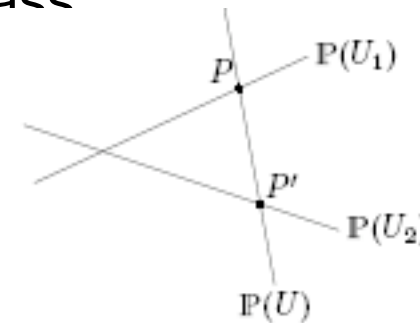
Dann aus dem Dimensionssatz 45 folgt, dass

$$\dim((X \vee P) \cap P(U_2)) = 0,$$

deswegen $\dim((X \vee P) \cap P(U_2)) =: P'$

ein Punkt.

Die Abbildung $\phi : P(U_1) \rightarrow P(U_2)$, $P \mapsto P'$, heißt Zentralprojektion.



Satz 49. Die Zentralprojektion $\phi : P(U_1) \rightarrow P(U_2)$ ist eine Projektivität.

Beweis. Aus (i) folgt $U \cap U_1 = U \cap U_2 = \{\vec{0}\}$ und wegen (ii) ist $U \oplus U_1 = U \oplus U_2 = V$. Nach Definition von Direktprodukt (Def. 49 Vorl. 19 LAAG I, siehe auch Vorl. 5) gibt es zu jedem $u_1 \in U_1$ eindeutig bestimmte Vektoren $u \in U$ und $u_2 \in U_2$ mit $u_1 = \vec{0} + u_1 = u + u_2$. Setze

$$\varphi : U_1 \rightarrow U_2, \quad u_1 \mapsto u_2.$$

Dann ist

$$\varphi(\mathbb{K}u_1) = P(U + \mathbb{K}u_1) \cap P(U_2) = \mathbb{K}u_2 = \vec{\varphi}(\mathbb{K}u_1) \forall u_1 \in U_1 \setminus \{\vec{0}\},$$

und $\vec{\varphi}$ ist linear und injektiv. Da die Dimensionen von U_1 und U_2 gleich sind, ist $\vec{\varphi}$ eine Bijektion nach 1. Dimensionsformel, □

Projektive Fortsetzung affiner Transformationen

Bemerkung. Jede Affinität $F : \mathbb{K}^n \rightarrow \mathbb{K}^n$ lässt sich zu einer projektiven Transformation $\varphi : P(\mathbb{K}^{n+1}) \rightarrow P(\mathbb{K}^{n+1})$ fortsetzen.

Die affine Transformation sei

$$F : (x) = T + Ax$$

mit der invertierbaren Matrix $A \in GL(n, \mathbb{K})$. In Vorl. 19 sahen wir bereits, dass man die Abbildung mit Hilfe von der $(n+1) \times (n+1)$ -Matrix der

Form $\bar{A} = \begin{pmatrix} \boxed{A} & \begin{matrix} T_1 \\ \vdots \\ T_n \end{matrix} \\ 0 & \dots & 0 & 1 \end{pmatrix}$, wobei $A \in GL(n, \mathbb{K})$, angeben kann:

wenn wir \bar{A} mit den Vektor $\begin{pmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{pmatrix}$ multiplizieren, bekommen wir den

Vektor $\begin{pmatrix} F(x) \\ 1 \end{pmatrix}$. Weil A invertierbar ist, ist auch die

$(n+1) \times (n+1)$ -Matrix \bar{A} invertierbar. Multiplikation mit \bar{A} definiert also eine Projektivität $\varphi : P(\mathbb{K}^{n+1}) \rightarrow P(\mathbb{K}^{n+1})$. Auf den ersten n

Koordinaten der Punkten $(x_0 : \dots : x_{n-1} : 1)$ von $P(\mathbb{K}^{n+1})$ ist dies genau die gegebene affine Transformation.

Bemerkung. Man kann selbstverständlich auch die Projektivität φ konstruieren, sodass φ die gegebene affine Transformation auf den letzten n Koordianten der Punkten $(1 : x_1 : \dots : x_n)$, in dem Fall ist

die Matrix von φ $\begin{pmatrix} 1 & 0 & \dots & 0 \\ T_1 & \boxed{A} & & \\ \vdots & & & \\ T_n & & & \end{pmatrix}$,

Def. 43 – Wiederholung. Eine bijektive Abbildung $\varphi : P(V) \rightarrow P(W)$ heißt **Kollineation**, wenn das Bild einer projektiven Geraden $G_{P,Q}$ durch die Punkte $P, Q \in G_{P,Q} \subseteq P(V)$ die projektive Gerade $G_{\varphi(P),\varphi(Q)}$ durch die Punkte $\varphi(P), \varphi(Q) \in P(W)$ ist.

Satz 47 – Wiederholung. Es gilt:

- (i) Zu je zwei Punkten $P, Q \in P(V)$ mit $P \neq Q$ gibt es genau eine projektive Gerade G , die P und Q enthält.
- (ii) Ist $\varphi : P(V) \rightarrow P(W)$ eine Projektivität, so ist φ eine Kollineation.

Hauptsatz der projektiven Geometrie

Satz 50 Seien V, W zwei $(n + 1)$ -dimensionale R -Vektorräume, $n \geq 2$, und sei $\varphi : P(V) \rightarrow P(W)$ eine Kollineation. Dann ist φ ein Projektivität.

Wiederholung: Satz 22 (Fundamentalsatz der affinen Geometrie (über \mathbb{R})) Seien $\mathcal{A}, \mathcal{A}_0$ affine Räume über den \mathbb{R} -Vektorräumen V, V_0 . Sei $\dim(\mathcal{A}) \geq 2$ und $F : \mathcal{A} \rightarrow \mathcal{A}_0$ eine Bijektion, die Geraden auf Geraden abbildet. Dann ist F eine Affinität.

Beweis von Satz 50. Wir zeigen, dass der Satz eine einfache Folgerung aus der Fundamentalsatz der affinen Geometrie ist.

Wir betrachten eine Hyperebene $H \subseteq P(V)$. Setze $H_1 := \text{Bild}_\varphi(H)$. H_1 ist auch eine Hyperebene.

In der Tat, wir betrachten $P_1 = \mathbb{K}v_1, P_2 = \mathbb{K}v_2 \in H_1$. Nach

Voraussetzungen enthält H_1 mit je 2 verschiedenen Punkte

$P_1 = \mathbb{K}v_1, P_2 = \mathbb{K}v_2 \in H_1$ die ganze Gerade durch diese Punkte. Dann

enthält \mathcal{H}_1 mit je 2 linear unabhängige Vektoren $v_1, v_2 \neq \vec{0}$ die ganze

Ebene $\{\lambda_1 v_1 + \lambda_2 v_2 \mid \lambda_i \in \mathbb{R}\}$. Insbesondere enthält \mathcal{H}_1 die Gerade

durch die Punkte v_1, v_2 . Dann ist \mathcal{H}_1 affin abgeschlossen, und deswegen

ein affiner Unterraum nach Satz 21. Da \mathcal{H}_1 den Punkt $\vec{0}$ enthält, ist \mathcal{H}_1

ein Untervektorraum von W . Dimension von \mathcal{H}_1 ist offensichtlich gleich

n . In der Tat, betrachten wir einen Punkt $Q \notin H$. Da

$\dim(Q \vee H) \stackrel{\text{Satz 45}}{=} n$, ist $\dim(Q \vee H) = P(V)$, woraus folgt, dass jedes

$Q' \in P(V)$ auf einer Gerade liegt durch Q und einen Punkt von H . Da φ

Geradentreu Bijektion ist, liegt jedes $Q'_1 \in P(W)$ auf einer Gerade durch

den Punkt $\varphi(Q)$ und einer Punkt von H_1 . Dann ist $\varphi(Q) \vee H_1 = P(W)$,

schließlich $\dim(H_1) = n - 1$.

Wir betrachten ein Koordinatensystem (x_0, \dots, x_n) in V , so dass $H = \{(x_0, \dots, x_n) \mid x_0 = 0\}$, und ein Koordinatensystem (y_0, \dots, y_n) in W , so dass $H_1 = \{(y_0, \dots, y_n) \mid y_0 = 0\}$. Wir betrachten die Abbildung

$$\tilde{\varphi} : \mathbb{K}^n \rightarrow \mathbb{K}^n, \quad \tilde{\varphi}(x_1, \dots, x_n) = (y_1, \dots, y_n)$$

sodass $\varphi(1 : x_1 : \dots : x_n) = (1 : y_1 : \dots : y_n)$. Da $\text{Bild}_\varphi(H) = H_1$ ist, ist die Abbildung wohldefiniert und ist eine Bijektion. Da φ Kollineation ist, ist $\tilde{\varphi}$ auch Geradentreu. Dann ist $\tilde{\varphi}$ eine affine Abbildung nach Satz 22, schliesslich φ eine projektive Abbildung nach Bemerkung oben, \square

Projektive Basen

Seien V, W zwei $(n + 1)$ -dimensionale \mathbb{K} -Vektorräume. Dann sind $n + 2$ Punkte P_0, \dots, P_{n+1} von $P(V)$ in allgemeiner Lage, wenn keine $n + 1$ Punkte davon einen echten projektiven Unterraum von $P(V)$ erzeugen. Man sagt dann, dass P_0, \dots, P_{n+1} eine projektive Basis von $P(V)$ bilden.

Bsp. $\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right)$ ist eine Projektive (Standard-)Basis in $P(\mathbb{R}^3)$

Satz 51. Sind P_0, \dots, P_{n+1} in allgemeiner Lage in $P(V)$ und sind Q_0, \dots, Q_{n+1} in allgemeiner Lage in $P(W)$, dann gibt es genau eine Projektivität $\varphi : P(V) \rightarrow P(W)$ mit $\varphi(P_i) = Q_i$ für alle $i = 0, \dots, n + 1$.

Satz 51. Sind P_0, \dots, P_{n+1} in allgemeiner Lage in $P(V)$ und sind Q_0, \dots, Q_{n+1} in allgemeiner Lage in $P(W)$, dann gibt es genau eine Projektivität $\varphi : P(V) \rightarrow P(W)$ mit $\varphi(P_i) = Q_i$ für alle $i = 0, \dots, n + 1$.

Beweis. Es ist $P_i = \mathbb{K}v_i$ mit $v_i \in V \setminus \vec{0}$ für $i = 0, \dots, n$ und analog $Q_i = \mathbb{K}w_i$ mit $w_i \in W \setminus \vec{0}$. Nach Voraussetzung bilden v_0, \dots, v_n eine Basis von V und w_0, \dots, w_n eine Basis von W . Sei $f : V \rightarrow W$ die durch $f(v_i) = \lambda_i w_i$ mit noch zu bestimmenden $\lambda_i \in \mathbb{K}$ für $i = 0, \dots, n$ definierte \mathbb{K} -lineare Abbildung. Es ist $v_{n+1} = \sum_{i=0}^n \mu_i v_i$ mit $0 \neq \mu_i \in \mathbb{K}$ für alle $i = 0, \dots, n$, da P_0, \dots, P_{n+1} in allgemeiner Lage sind, und analog $w_{n+1} = \sum_{i=0}^n \eta_i w_i$ mit $0 \neq \eta_i \in \mathbb{K}$.

Setze $\lambda_i = \frac{\eta_i}{\mu_i}$, $\eta_i = \mu_i \lambda_i$ für $i = 0, \dots, n$. Dann ist

$$f(v_{n+1}) = \sum_{i=0}^n \mu_i f(v_i) = \sum_{i=0}^n \eta_i w_i = w_{n+1}.$$

Für $\varphi : P(V) \rightarrow P(W)$, $\mathbb{K}v \mapsto \mathbb{K}f(v)$, ist dann $\varphi(P_i) = Q_i$ und $\vec{\varphi} = f$.

Bis auf einen Faktor $\lambda \in \mathbb{K}^*$ ist $\vec{\varphi}$ eindeutig bestimmt, □

Folgerung. Jede Projektivität ist eine Zentralprojektion.

Folgerung. Sei $\varphi : P(\mathbb{K}^{n+1}) \rightarrow P(\mathbb{K}^{n+1})$ eine Projektivität. Es gibt projektive Hyperebenen $H_1, H_2 \in P(\mathbb{K}^{n+2})$, eine Gerade $G \in \mathbb{K}^{n+2}$, und Projektivitäten $\varphi_1 : P(\mathbb{K}^{n+1}) \rightarrow H_1$ und $\varphi_2 : H_2 \rightarrow P(\mathbb{K}^{n+1})$ sodass $\varphi = \varphi_2 \circ \varphi_G \circ \varphi_1$, wobei φ_G die Zentralprojektion bzgl. Gerade G ist.

Beweis. Wir betrachten zwei beliebige $H_1 = P(U_1)$, $H_2 = P(U_2)$ und Gerade $G = P(V)$ mit

(i) $G \cap P(U_1) = G \cap P(U_2) = \emptyset$. Dann ist

(ii) $G \vee P(U_1) = G \vee P(U_2) = P(V)$ automatisch erfüllt.

Die dazugehörige Zentralprojektion bezeichnen wir φ_G . Nach Satz 49 ist φ_G eine Projektivität. Wir nehmen eine Basis (B_0, \dots, B_{n+1}) in $P(\mathbb{K}^{n+1})$. Die Bilder $\varphi_{B_0}, \dots, \varphi_{B_{n+1}}$ sind auch in der allgemeineren Lage, und bilden deswegen eine projektive Basis. Wir nehmen eine projektive Basis (P_0, \dots, P_{n+1}) in H_1 . Dann ist $(Q_0 = \varphi_G(P_0), \dots, Q_{n+1} = \varphi_G(P_{n+1}))$, auch eine projektive Basis in H_2 . Wir betrachten Projektivität φ_1 , die B_0, \dots, B_{n+1} jeweils in P_0, \dots, P_{n+1} Standard-Basispunkten überführt, und eine Projektivität φ_2 , die Q_0, \dots, Q_{n+1} jeweils in $\varphi_{B_0}, \dots, \varphi_{B_{n+1}}$ überführt. Die Verkettung $\varphi_2 \circ \varphi_G \circ \varphi_1 : P(\mathbb{K}^{n+1}) \rightarrow P(\mathbb{K}^{n+1})$ bildet nach Konstruktion die Punkte B_0, \dots, B_{n+1} jeweils auf $\varphi_{B_0}, \dots, \varphi_{B_{n+1}}$ ab; nach Satz 51 ist dann $\varphi = \varphi_2 \circ \varphi_G \circ \varphi_1$,

Das Doppelverhältnis

Sei $n = 1$ und $X = P(V)$ mit $\dim(V) = 2$. Dann sind je drei verschiedene Punkte $P_0, P_1, P_2 \in X$ in allgemeiner Lage. Seien nun P_0, P_1, P_2, P_3 vier paarweise verschiedene Punkte in X . Dann gibt es nach Satz 51 genau eine Projektivität $\varphi_X : X \rightarrow P(\mathbb{K}^2)$ mit

$$\varphi_X(P_0) = (1 : 0), \quad \varphi_X(P_1) = (1 : 1), \quad \text{und} \quad \varphi_X(P_2) = (0 : 1).$$

Dadurch ist $\phi_X(P_3) =: (\lambda : \mu)$ schon eindeutig festgelegt. Man nennt $D(P_0, P_1, P_2, P_3) = (\lambda : \mu) \in P(\mathbb{K}^2) (= \mathbb{K} \cup \infty)$ das **Doppelverhältnis** der vier Punkte $P_0, P_1, P_2, P_3 \in X$.

Satz 52. Sei $\dim(V) = 2 = \dim(W)$. Seien $X = P(V)$ und $Y = P(W)$ zwei projektive Geraden. Ist $\varphi : X \rightarrow Y$ eine Projektivität, so gilt $D(P_0, P_1, P_2, P_3) = D(\varphi(P_0), \varphi(P_1), \varphi(P_2), \varphi(P_3))$ für je vier paarweise verschiedene Punkte $P_0, P_1, P_2, P_3 \in X$. Das Doppelverhältnis bleibt also unter Projektivitäten erhalten.

Satz 52. Sei $\dim(V) = 2 = \dim(W)$. Seien $X = P(V)$ und $Y = P(W)$ zwei projektive Geraden. Ist $\varphi : X \rightarrow Y$ eine Projektivität, so gilt $D(P_0, P_1, P_2, P_3) = D(\varphi(P_0), \varphi(P_1), \varphi(P_2), \varphi(P_3))$ für je vier paarweise verschiedene Punkte $P_0, P_1, P_2, P_3 \in X$. Das Doppelverhältnis bleibt also unter Projektivitäten erhalten.

Beweis. Sei $Q_i = \varphi(P_i)$ für $i = 0, 1, 2, 3$. Dann gelten

$$\begin{aligned} (\varphi_Y \circ \varphi)(P_0) &= \varphi_Y(Q_0) = (1 : 0) = \varphi_X(P_0) \\ (\varphi_Y \circ \varphi)(P_1) &= \varphi_Y(Q_1) = (1 : 1) = \varphi_X(P_1) \\ (\varphi_Y \circ \varphi)(P_2) &= \varphi_Y(Q_2) = (0 : 1) = \varphi_X(P_2), \end{aligned}$$

also ist $\varphi_Y \circ \varphi = \varphi_X$ nach Satz 51. Damit folgt

$$(\varphi_Y \circ \varphi)(P_3) = \varphi_X(P_3) = D(P_0, P_1, P_2, P_3), \quad \square.$$

Beispiel. Seien $P_i = (1 : \lambda_i) \in P(\mathbb{K}^2)$ für $i = 0, 1, 2, 3$ paarweise verschiedene Punkte. Für die Möbiustransformation

$$\varphi : K \cup \{\infty\} \longrightarrow K \cup \{\infty\}, \quad x \longmapsto \frac{\mu_1 - \mu_2}{\mu_1 - \mu_0} \cdot \frac{x - \mu_0}{x - \mu_2}$$

gilt $\varphi(\mu_0) = 0$, $\varphi(\mu_1) = 1$ und $\varphi(\mu_2) = \infty$. Es ist dann

$$D(\mu_0, \mu_1, \mu_2, \mu_3) = \frac{\frac{\mu_1 - \mu_2}{\mu_1 - \mu_0}}{\frac{\mu_3 - \mu_2}{\mu_3 - \mu_0}}$$

das Doppelverhältnis von $\mu_0, \mu_1, \mu_2, \mu_3$.

Das Doppelverhältnis in Determinantenform

$$\mathbf{x} = \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_0 \\ y_1 \end{pmatrix}, \quad \mathbf{u} = \begin{pmatrix} u_0 \\ u_1 \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} v_0 \\ v_1 \end{pmatrix} \in \mathbb{P}_1(\mathbb{K})$$

ist es in Determinantenform

$$DV(\mathbf{x}, \mathbf{y}; \mathbf{u}, \mathbf{v}) = \frac{\det \begin{pmatrix} x_0 & u_0 \\ x_1 & u_1 \end{pmatrix}}{\det \begin{pmatrix} x_0 & v_0 \\ x_1 & v_1 \end{pmatrix}} : \frac{\det \begin{pmatrix} y_0 & u_0 \\ y_1 & u_1 \end{pmatrix}}{\det \begin{pmatrix} y_0 & v_0 \\ y_1 & v_1 \end{pmatrix}}.$$

Organisatorisches

- ▶ Klausur am 1. August, 10-12 Uhr in CZ 3, HS 3
- ▶ Zulassungskriterium: 60% von Hausaufgaben + Bonuspunkten. Bitte bei zweifeln mit Übungsgruppenleitern klären. Keine Anmeldung.
- ▶ 5 Aufgaben; nur eine davon rechnerische davon 3 theoretische; eine davon wird bestehen, einen wichtigen Satz aus der Vorlesung zu beweisen (die Liste von „wichtigen Sätze“ ist auf der nächste Folie)
- ▶ Keine Hilfsmittel zugelassen; Papier wird gegeben.
- ▶ Bitte Ihren Lichtbildausweis (egal welcher) mitbringen
- ▶ Über Einsichtmöglichkeit werden Sie über CAJ und der Homepage der Vorlesung informiert. Vermutlich eine Woche nach der Veröffentlichung der Ergebnisse in CAJ + 2 Wochen irgendwann. Sie können einfach vorbei schauen, aber wegen Sommerpause es lohnt sich, sicherheitshalber zuerst per email/telefon Termin **bei Herr Dr. Schöbel (ich bin nicht da)** vereinbaren.
- ▶ Genauer Termin für Nachklausur wird morgen besprochen; vermutlich die dritte Woche der Oktober-Prüfungszeit

Sätze zur Klausur

- ▶ Mengenlehre
 - ▶ Satz 2 (Wann sind zwei Menge gleichmächtig)
 - ▶ Satz 44 (Hamel-Basis)
 - ▶ Lösung des dritten Hilbertschen Problem
- ▶ Jordan-Form
 - ▶ Satz 5 (Normalform für die Endomorphismen, die über \mathbb{C} diagonalisierbar sind.)
 - ▶ Satz 11 (Zerlegung in Produkt von verallgemeinerten Eigenräumen)
 - ▶ Lemma 4 (Zerlegungslemma)
 - ▶ Satz 13 (Jordan-Normalform)
- ▶ Affine, projektive und konvexe Geometrie
 - ▶ Satz 22 (Fundamentalsatz der affinen Geometrie)
 - ▶ Satz 23 (Konvexe Hülle)
 - ▶ Satz 34 (Extrempunkt = Ecke = ein Element der minimalen Darstellung)
 - ▶ Satz 40 ($K = K^{**}$)
 - ▶ Satz 55 (von Desargues)

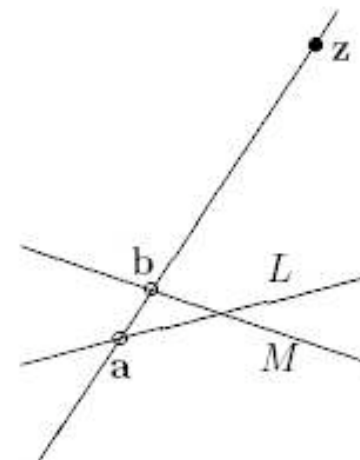
Lineare Geometrie der projektiven Ebene

Def. Eine **Perspektivität** in der Ebene ist folgendes (unten ist $P(\mathbb{K}^3) = \mathbb{P}_2(\mathbb{K})$); wir werden annehmen (obwohl es nicht nötig ist), dass $\#\mathbb{K} = \infty$

Es seien zwei Geraden $L, M \subset \mathbb{P}_2(\mathbb{K})$ und ein Punkt $z \in \mathbb{P}_2(\mathbb{K})$ in der projektiven Ebene gegeben. Wir nennen z das *Perspektivitäts-Zentrum* und die Geraden $S \subset \mathbb{P}_2(\mathbb{K})$ durch z die *Perspektivitäts-Strahlen*. Jeder Strahl S schneidet die Geraden L und M in genau einem Punkt. Umgekehrt gibt es zu jedem Punkt $a \in L$ genau einen Strahl S_a , nämlich die von z und a aufgespannte Gerade. Dieser Strahl schneidet die Gerade M im Punkt $b = S_a \cap M$. Dadurch ist eine Abbildung

$$L \rightarrow M, \quad a \mapsto b = S_a \cap M$$

definiert. Sie heißt die *Perspektivität mit Zentrum z* .



Bemerkung. Diese Definition geht genauso für zwei Hyperebenen $L, M \subseteq P(\mathbb{K}^{n+1})$ im n -dimensionalen projektiven Raum. Man muss nur darauf achten, dass das Perspektivitäts-Zentrum z auf keiner dieser beiden Hyperebenen liegt.

Satz 53. Jede Perspektivität ist eine Projektivität. (Beweis wie von Satz 49)

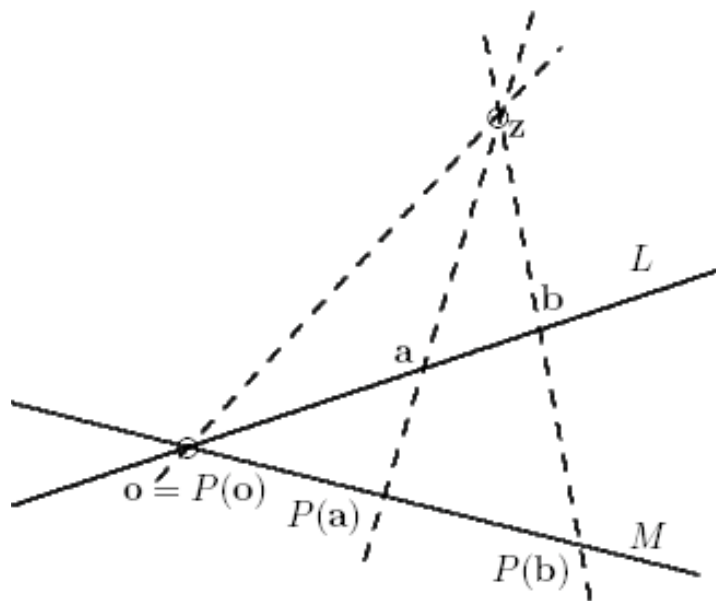
Zunächst wollen wir Projektivitäten $P : L \rightarrow M$ zwischen zwei Geraden $L, M \subseteq P(\mathbb{K}^3)$ betrachten. Wenn nichts anderes gesagt wird, setzen wir $L \neq M$ voraus. Dann nennen wir den Schnittpunkt $L \cap M =: o$.

Satz 54 (Projektivitäten und Perspektivitäten)

- (a) Eine Projektivität $P : L \rightarrow M$ ist eine Perspektivität genau dann, wenn $P(o) = o$.
- (b) Jede Projektivität $P : L \rightarrow M$ ist ein Produkt von höchstens zwei (falls $L \neq M$) oder drei (falls $L = M$) Perspektivitäten.

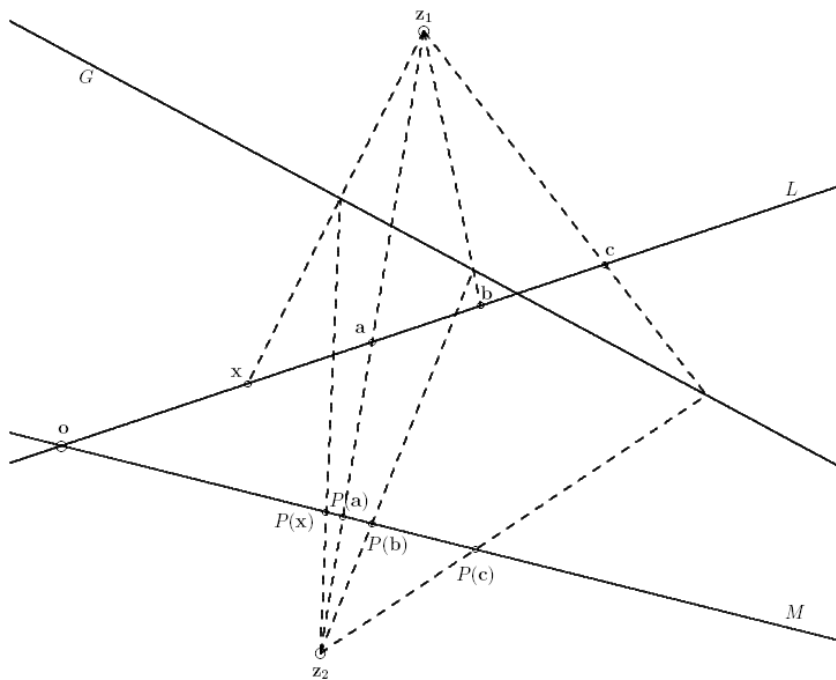
Beweis. (a) Sei $P(o) = o$. Wir wählen zwei Punkte $o \neq a \neq b \neq o$ auf der Geraden L . Dann sind ihre Bilder $P(a), P(b) \in M$ voneinander und von o verschieden.

Als Projektionszentrum wählen wir den Schnittpunkt z der Geraden $oP(a)$ und $oP(b)$. Die Perspektivität mit Zentrum z bildet, ebenso wie die Projektivität P , die drei (Basis)Punkte o, a, b auf $o, P(a), P(b)$ ab.



Aus Satz 51 folgt, dass P die Perspektivität mit Zentrum z ist.

(b) Jede Projektivität $P : L \rightarrow M$ ist ein Produkt von höchstens zwei (falls $L \neq M$) oder drei (falls $L = M$) Perspektivitäten.

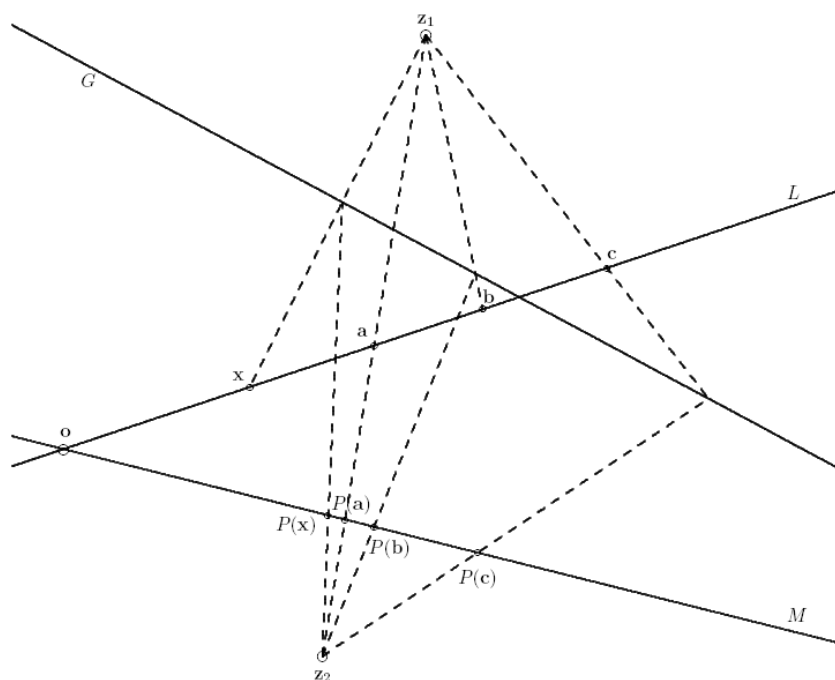


Beweis. Auf L wählen wir drei Punkte a, b, c , die untereinander und von o , sowie $P^{-1}(o)$ verschieden seien. Dann sind ihre Bildpunkte $P(a), P(b), P(c) \in M$ voneinander und von o verschieden. Auf der Verbindungsgeraden von a und $P(a)$ wählen wir zwei beliebige Punkte $z_1 \neq z_2$, die von a und $P(a)$ verschieden sein sollen.

Weil die Punkte $b, c, P(b)$ und $P(c)$ nicht auf der Verbindungsgeraden von a und $P(a)$ liegen, sind die Geraden z_1b und $z_2P(b)$, sowie z_1c und $z_2P(c)$ verschieden, und schneiden sich (Satz 46 (c)) jeweils in einem Punkt.

Die Gerade durch diese beiden Schnittpunkte nennen wir G . Sei $P_1 : L \rightarrow G$ durch die Perspektivität mit Zentrum z_1 gegeben und $P_2 : G \rightarrow M$ durch die Perspektivität mit Zentrum z_2 . Wir verfolgen die Punkte a, b und c unter $P_2 \circ P_1$:

Wir verfolgen die Punkte a, b und c unter $P_2 \circ P_1$:



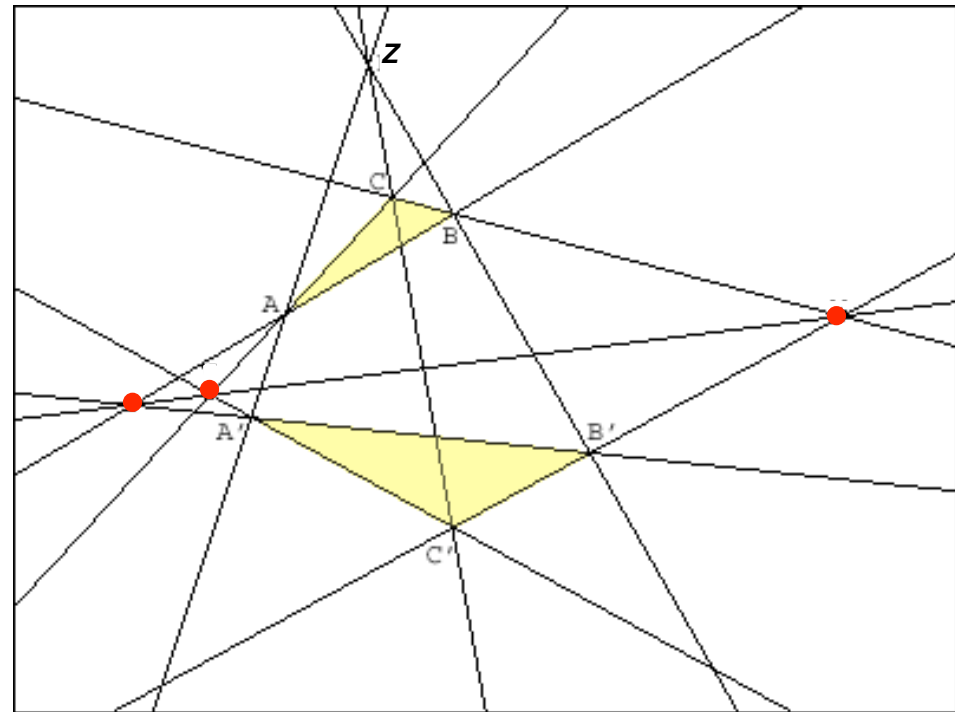
$$P_1 : \begin{cases} a \mapsto G \cap aP(a) \\ b \mapsto G \cap z_1 b = G \cap z_2 P(b) \\ c \mapsto G \cap z_1 c = G \cap z_2 P(c) \end{cases} \quad P_2 : \begin{cases} G \cap aP(a) \mapsto P(a) \\ G \cap z_2 P(b) \mapsto P(b) \\ G \cap z_2 P(c) \mapsto P(c) \end{cases}$$

Wir sehen: Die Projektivität $P_2 \circ P_1$ bildet die drei Punkte a, b, c genauso ab, wie P .
 Nach Satz 49 muß die Projektivität P mit $P_2 \circ P_1$ übereinstimmen. Falls $L = M$ gewesen sein sollte, bilden wir zuerst L auf eine Gerade $L' \neq L$ durch eine Perspektivität ab, und dann $L' \mapsto M$ durch ein Produkt von zwei Perspektivitäten,

Satz 55 (von Desargues (1591-1661))



Die Dreiecke $a_1a_2a_3$ und $b_1b_2b_3$ seien **perspektiv**, d.h., es gebe eine Perspektivität mit Zentrum $z \notin \{a_1, \dots, b_3\}$, welche für $i = 1, 2, 3$ die Ecken a_i auf die Ecken b_i abbildet.



Dann sind die Schnittpunkte $t_1 := a_2a_3 \cap b_2b_3$, $t_2 := a_3a_1 \cap b_3b_1$, $t_3 := a_1a_2 \cap b_1b_2$ kollinear.

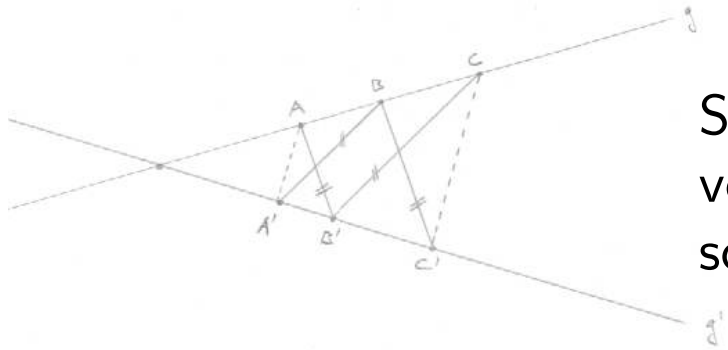
Beweis. Falls zwei der Ecken a_i und b_i zusammenfallen, wird die Aussage trivial: Sei etwa $a_1 = b_1$. Dann wird $t_2 = t_3 = a_1 = b_1$.

Deswegen können wir oBdA $a_i \neq b_i$ für $i = 1, 2, 3$ annehmen. Wir bezeichnen mit $\hat{o} \neq \hat{a}_1, \dots, \hat{b}_3, \hat{z} \in \mathbb{K}^3$ Vektoren, die zu den Punkten $a_1, \dots, b_3, z \in P(\mathbb{K}^3)$ gehören. Weil a_i, b_i und z kollinear sind, sind die Vektoren \hat{a}_i, \hat{b}_i und \hat{z} linear abhängig. Es gibt deswegen Zahlen $\lambda_i, \mu_i \in \mathbb{K}$ mit $\hat{z} = \lambda_i \hat{a}_i + \mu_i \hat{b}_i$ für $i = 1, 2, 3$. Dann gilt:

$$\hat{z} = \lambda_1 \hat{a}_1 + \mu_1 \hat{b}_1 = \lambda_2 \hat{a}_2 + \mu_2 \hat{b}_2 = \lambda_3 \hat{a}_3 + \mu_3 \hat{b}_3$$

Satz 56 von Pappos (IV. Jhd. n. Chr.)

Einfache (affine) Version des Satzes



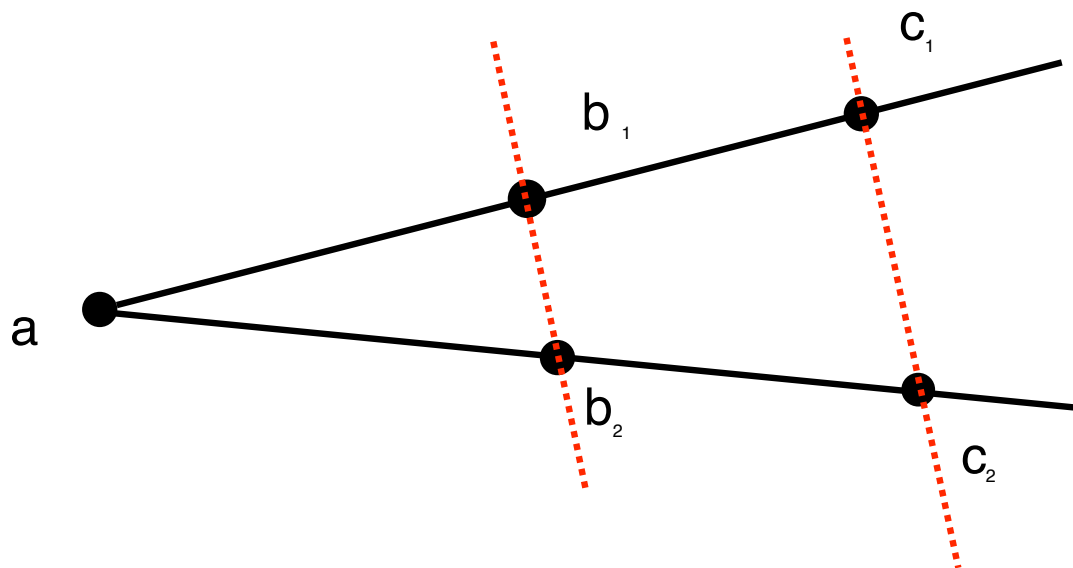
Seien G, G' Geraden in der Ebene A, B, C
verschiedene Punkte auf G A', B', C' ver-
schieden Punkte auf G'

$A, B, C, A', B', C' \notin G \cap G'$.

Dann gilt: Wenn $AB' \parallel BC'$ und $A'B \parallel B'C$, dann $AA' \parallel CC'$

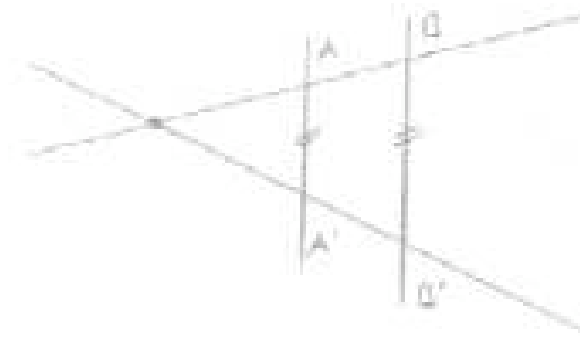
Beweis mit Hilfe von affiner Geometrie

Wiederholung — Strahlensatz (Satz 20) Seien (a, b_1, c_1) und (a, b_2, c_2) kollineare Punkttripel auf verschiedenen Geraden durch $a \notin \{b_1, b_2, c_1, c_2\}$. Dann sind die Geraden \mathcal{G}_{b_1, b_2} und \mathcal{G}_{c_1, c_2} genau dann parallel, wenn $TV(a, b_1, c_1) = TV(a, b_2, c_2)$.



Wiederholung. Ist (a, b, c) ein kollineares Punkttripel und $a \neq b$, so heißt der durch die Gleichung $\vec{ac} = \lambda \vec{ab}$ eindeutig bestimmte Skalar λ das **Teilverhältnis** des kollinearen Punkttripels (a, b, c) , bezeichnet durch $TV(a, b, c)$.

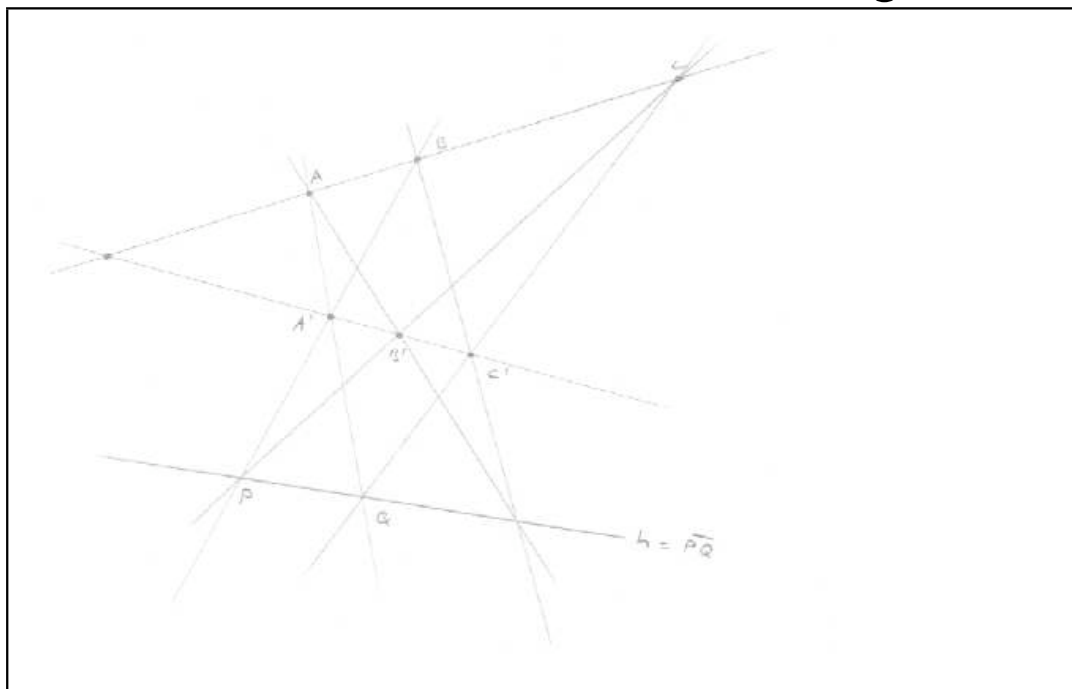
Einfache Formulierung des Strahlensatzes



$$AA' \parallel BB' \Leftrightarrow \exists \lambda : \overrightarrow{OB} = \lambda \overrightarrow{OA} \text{ und } \overrightarrow{OB'} = \lambda \overrightarrow{OA'}$$

Projektive Version des Satzes von Pappus:

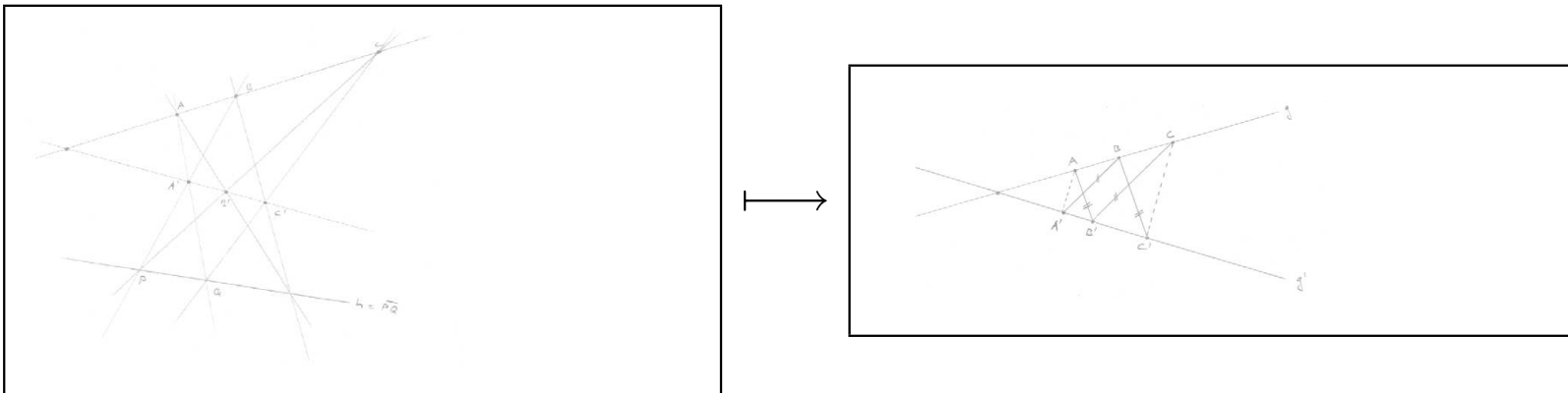
Satz 55: G, G' projektive Geraden in $P(\mathbb{R}^3)$. $G \cap G' =: \{0\}$. O, A, B, C verschiedene Punkte auf G O, A', B', C' verschieden Punkte auf G' . Die Durchschnittspunkte der drei Paare von Geraden $AB', BC', A'B, B'C; AA', CC'$ liegen auf der selben Gerade.



Poetische (Coxeter):	Formulierung
Liegen die Eckpunkte eines 6-Ecks $AB'CC'BA'$ alternierend auf zwei Geraden, so treffen sich die drei Paare gegenüberliegender Seiten in kollinearen Punkten.	

Beweis.

Man betrachte eine Projektivität $\phi : P(\mathbb{R}^3) \rightarrow P(\mathbb{R}^3)$, die die Gerade $H = PQ$ auf unendlich ferne Gerade $H_\infty = \{x_0 : x_1 : x_2 \mid x_0 = 0\}$ abbildet. (Existiert nach Satz 51: es genügt die Punkte P, Q und noch einen Punkt von H auf unendlich ferne Punkte, z.B. $(0 : 0 : 1)$, $(0 : 1 : 0)$, $(0 : 1 : 1)$, abbilden.) Dann ist $\phi(A)\phi(B') \parallel \phi(B)\phi(C')$ und $\phi(A')\phi(B) \parallel \phi(B')\phi(C)$ als Geraden auf der affinen $\mathcal{E}_2 = P(\mathbb{R}^3) \setminus H_\infty (= \mathbb{R}^2$ mit inhomogenen Koordinaten $(\frac{x_1}{x_0}, \frac{x_2}{x_0})$).)
(Das affine Bild wie folgt verändert:)



Wir sehen dass wir genau in den Voraussetzungen des einfachen (affinen) Version des Satzes von Pappos sind, also $\phi(A)\phi(A') \parallel \phi(C)\phi(C')$ (als Geraden auf \mathcal{E}_2). Dann liegt der Schnittpunkt von $\phi(A)\phi(A')$ $\phi(C)\phi(C')$ (als Geraden auf $P(\mathbb{R}^3)$) auf $Bild_\phi(H)$, deswegen ist Schnittpunkt von $AA' CC'$ auf H ,



Dualität in der projektiven Ebene

Ausgangspunkt ist folgende Bemerkung: Genauso, wie ein Punkt $a \in P(\mathbb{K}^3)$ durch seine drei homogenen Koordinaten $(a_0 : a_1 : a_2)$ festgelegt ist, so ist auch eine Gerade $L : \alpha_0 x_0 + \alpha_1 x_1 + \alpha_2 x_2 = 0$ durch drei Zahlen $\alpha_0, \alpha_1, \alpha_2 \in \mathbb{K}$ festgelegt. Und ändert man diese drei Zahlen um einen gemeinsamen Faktor $\lambda \in \mathbb{K}^*$ ab, so ändert die Gerade nicht. Man kann also auch die Gerade L durch ein Tripel $(\alpha_0 : \alpha_1 : \alpha_2)$ homogener Koordinaten charakterisieren. Die **Inzidenz** (=ein Punkt ist mit einer Geraden inzident, falls er auf der Gerade liegt) eines Punktes a mit der Geraden L kann man den homogenen Koordinaten von a und L ansehen: $a \in L \iff \alpha_0 a_0 + \alpha_1 a_1 + \alpha_2 a_2 = 0$.

In dieser Inzidenz-Gleichung kann man die Rollen des Punktes a und der Geraden L vertauschen: Nimmt man die Koordinaten des Punktes a , um eine Gerade

$$a^* : a_0 x_0 + a_1 x_1 + a_2 x_2 = 0 \subseteq P(\mathbb{K}^3)$$

zu definieren, und die Koordinaten von L , um einen Punkt

$$L^* := (\alpha_0 : \alpha_1 : \alpha_2) \in P(\mathbb{K}^3) \text{ zu definieren, so gilt } a \in L \iff a^* \ni L^*.$$

Man nennt $a^* \subseteq P(\mathbb{K}^3)$ die **duale Gerade** zum Punkt a und $L^* \in P(\mathbb{K}^3)$ den dualen Punkt zur Geraden L . Diese Dualitätsbeziehung ist symmetrisch:

$$L = a^* \iff L^* = a.$$

Weil man die Rollen von Punkt und Gerade bei der Inzidenz vertauschen kann, bleibt eine Inzidenz zwischen einem Punkt und einer Geraden beim Übergang zur dualen Geraden und dem dualen Punkt erhalten. Dies ist das

Dualitätsprinzip: Gilt eine Aussage über Punkte und Geraden, die nur mit Hilfe von Inzidenzen formuliert ist (z.B. der Satz von Pappos, der Satz von Desargues), so gilt auch die duale Aussage, die man erhält, wenn man alle









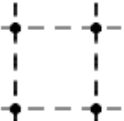
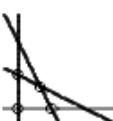


Punkte durch die dualen Geraden

Geraden durch die dualen Punkte

Inzidenzen durch die Inzidenzen zwischen den dualen Objekten ersetzt.

Duale Objekte

Als Beispiele stellen wir eine kleine Liste der einfachsten Figuren in der projektiven Ebene und ihrer dualen Figuren zusammen:

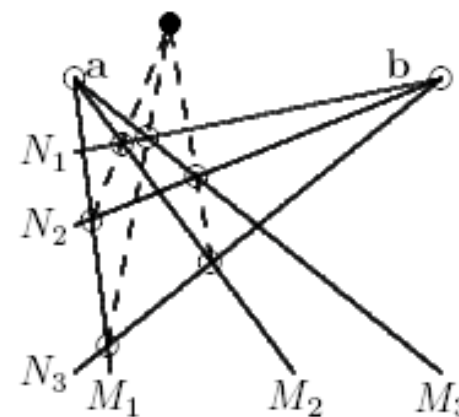
Figur	Dual
 Punkt	 Gerade
 Punkt auf einer Geraden	 Gerade durch einen Punkt
 zwei Punkte auf einer Geraden	 zwei Geraden durch einen Punkt
 Dreieck	 Dreieck
 Viereck	 Viereck
 vollständiges Viereck	 vollständiges Viereck

Als Beispiel einer nicht ganz trivialen Aussage dualisieren wir den Satz von Pappos:

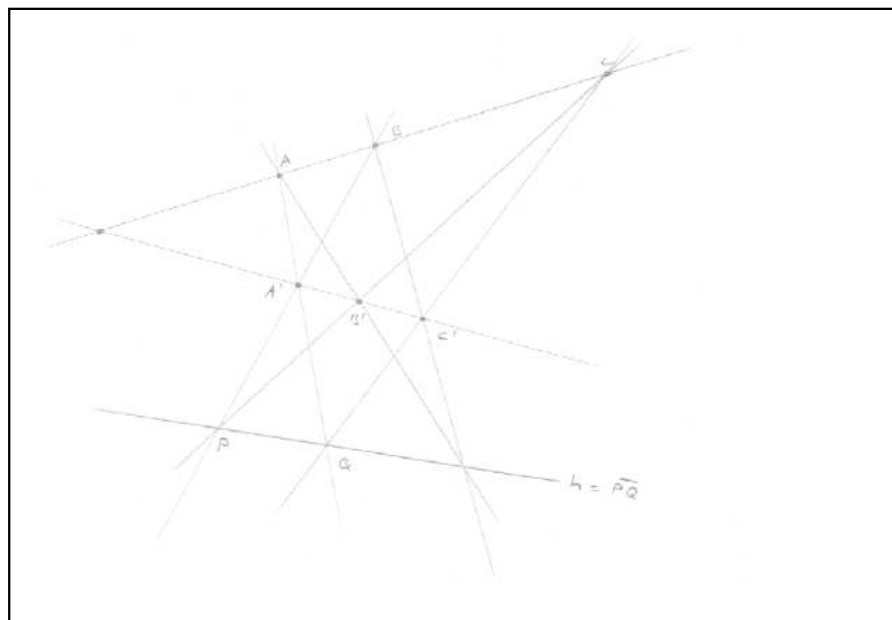
Satz (Dual zum Satz von Pappos)
Es seien a und $b \in \mathbb{P}_2$ zwei verschiedene Punkte. Durch jeden der Punkte seien drei Geraden $M_1, M_2, M_3 \ni a$ und $N_1, N_2, N_3 \ni b$ gewählt, die voneinander und von der Verbindungsgeraden ab verschieden sein mögen. Dann treffen sich die drei Verbindungsgeraden der Punktepaare

$$M_i \cap N_j \text{ und } N_i \cap M_j, \quad i \neq j = 1, 2, 3$$

in einem Punkt.



Satz 55 – Wiederholung: G, G' projektive Geraden in $P(\mathbb{R}^3)$. $G \cap G' =: \{0\}$. O, A, B, C verschiedene Punkte auf G O, A', B', C' verschiedenen Punkten auf G' Die Durchschnittspunkte der drei Paare von Geraden $AB', BC'; A'B, B'C; AA', CC'$ liegen auf der selben Gerade.



Abschied:

- ▶ Die Vorlesung Morgen wird Herr Dr. Schöbel halten – er wird über ein Paar mathematischen Paradoxen sprechen. Ich werde mithören.
- ▶ Die Übungsgruppe am Do fällt aus: stattdessen wird (vermutlich, am Dienstag 29. Juli vor der Klausur) eine Fragestunde angeboten: Wird nichts vorgerechnet/vorbereitet, sondern die Fragen werden beantwortet. Ort und Zeit in CAJ
- ▶ Viel Erfolg!