Proseminar Konvexe Mengen
Separierende Hyperebenen

Arne Grenzebach

2. November 2004

Als Literaturgrundlage dient diesem Proseminar das Buch „Convex sets and their applications“ von Steven R. Lay (New York etc.: John Wiley & Sons, 1982). In diesem Vortrag wird das Problem behandelt, zwei konvexe Mengen durch eine Hyperebene zu separieren (☞ Lay, Abschnitt 4, S. 33 bis 40). Zunächst erfolgt aber eine kurze Wiederholung der Begriffe und Zusammenhänge von linearer Abbildung, Linearform und Hyperebene, was aus der Linearen Algebra bekannt sein sollte (☞ Lay, Abschnitt 3, S. 27 bis 32).

Hyperebenen

Im E^n ($\mathbb{R}^n$ mit euklidischem Skalarprodukt) wird ein affiner Unterraum der Dimension $n-1$ als *Hyperebene* bezeichnet, was der natürlichen Verallgemeinerung der Begriffe Linie im E^2 und Ebene im E^3 entspricht. Genau wie eine Gerade im E^2 durch eine Gleichung der Form $\alpha_1 x_1 + \alpha_2 x_2 = \delta$ beschrieben werden kann, können auch Ebenen im E^3 und Hyperebenen im E^n durch je eine lineare Gleichung dargestellt werden.

Definition (3.1): *Eine Abbildung $f\colon \mathbb{R}^n \to \mathbb{R}^m$ heißt **linear**, falls für alle $x, y \in \mathbb{R}^n$ und für alle $\lambda, \mu \in \mathbb{R}$ gilt:*

$$f(\lambda x + \mu y) = \lambda f(x) + \mu f(y).$$

*Eine lineare Abbildung $f\colon \mathbb{R}^n \to \mathbb{R}$ heißt **Linearform**.*

Über Linearformen lassen sich nun auch Hyperebenen definieren; es gilt nämlich der folgende Satz:

Satz (3.2): *$H \subset \mathrm{E}^n$ ist genau dann eine Hyperebene, wenn es eine Linearform $f \neq 0$ gibt, so daß für ein $\delta \in \mathbb{R}$ gilt:*

$$H = \{x \in \mathrm{E}^n \mid f(x) = \delta\}.$$

Da wir dies im folgenden häufig ausnutzen, führen wir für Hyperebenen nachfolgende abkürzende Schreibweise ein:

$$\boxed{H = \{x \in \mathrm{E}^n \mid f(x) = \delta\} =: [f : \delta]}.$$

Nun können Linearformen ihrerseits wieder durch das im E^n gegebene Skalarprodukt dargestellt werden, was das folgende Korollar ausdrückt:

Korollar (3.5): *Wenn f eine Linearform auf dem E^n ist, dann gibt es ein $u \in \mathrm{E}^n$, so daß für alle $x \in \mathrm{E}^n$ gilt: $f(x) = \langle x|u\rangle$.*

Bemerkung: Kombiniert man nun die beiden Ideen aus Satz (3.2) und Korollar (3.5), so ergibt sich: $H \subset \mathrm{E}^n$ ist genau dann eine Hyperebene, wenn es ein $u \in \mathrm{E}^n \setminus \{0\}$ gibt, so daß für ein $\delta \in \mathbb{R}$ gilt:

$$H = \big\{x \in \mathrm{E}^n \;\big|\; \langle x|u\rangle = \delta\big\}.$$

Dies ist die bereits oben erwähnte Gleichungsdarstellung einer Hyperebene.

Weil ein derartiger Vektor u zum Richtungsraum der Hyperebene H orthogonal ist†, berechtigt dies zu folgender Definition:

Definition (3.6): *$u \in \mathrm{E}^n$ heißt **Normalenvektor** zur Hyperebene H, wenn gilt:*

$$H = \big\{x \in \mathrm{E}^n \;\big|\; \langle x|u\rangle = \delta\big\}.$$

Die folgenden drei Sätze beschreiben Eigenschaften von Linearformen und Hyperebenen:

Satz (3.3): *Sind f, g Linearformen auf dem E^n mit $[f : \alpha] = [g : \beta]$, so gibt es ein $\lambda \in \mathbb{R} \setminus \{0\}$, so daß gilt: $f = \lambda \cdot g$ und $\alpha = \lambda \cdot \beta$ (d. h., die Linearformen bestimmen sich gebenseitig bis auf skalare Vielfache).*

Satz (3.4): *Sei $\{e_1, \ldots, e_n\}$ die Standardbasis des E^n. Dann gilt für eine Hyperebene $H := [f : \delta] \subset \mathrm{E}^n$:*

① *Es existieren $\lambda_i \in \mathbb{R}$, so daß für alle $x = \sum\limits_{i=1}^{n} \alpha_i e_i \in \mathrm{E}^n$ gilt: $f(x) = \sum\limits_{i=1}^{n} \alpha_i \lambda_i$,*
② *f stetig,*
③ *H abgeschlossen.*

Satz (3.8): *Lassen sich im E^n zwei Hyperebenen $H_1 = [f : \alpha]$ und $H_2 = [f : \beta]$ durch dieselbe Linearform f beschreiben, gilt: $H_1 \parallel H_2$.*

† Für $x_1, x_2 \in H = \{x \in \mathrm{E}^n \mid \langle x|u\rangle = \delta\}$ gilt: $\langle x_1 - x_2|u\rangle = \langle x_1|u\rangle - \langle x_2|u\rangle = \delta - \delta = 0$.

Separierende Hyperebenen

Wie im vorherigen Abschnitt erwähnt wurde, lassen sich Hyperebenen durch eine Linearform charakterisieren, welche wiederum durch ein Skalarprodukt dargestellt werden kann. Nun wollen wir uns dem Problem zuwenden, zwei konvexe Mengen durch eine Hyperebene zu „separieren".

Bemerkung (4.1): Gilt für eine Linearform f: $f(x) \leq \alpha$ für alle $x \in A$, so schreiben wir $f(A) \leq \alpha$. Analoge Schreibweisen werden für die Relationen $\geq, <, >$ verwendet.

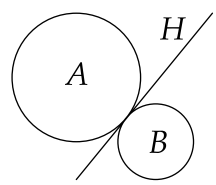
Definition (4.2): *Zwei Mengen A, B werden von der Hyperebene $H := [f : \alpha]$* ***separiert*** *(auch:* ***schwach separiert****), falls gilt:*

$$\begin{array}{rccc} & f(A) \leq \alpha & \text{und} & f(B) \geq \alpha \\ \text{bzw.} & f(A) \geq \alpha & \text{und} & f(B) \leq \alpha \end{array}$$

Definition (4.3): *Zwei Mengen A, B werden von der Hyperebene $H := [f : \alpha]$* ***streng separiert****, falls gilt:*

$$\begin{array}{rccc} & f(A) < \alpha & \text{und} & f(B) > \alpha \\ \text{bzw.} & f(A) > \alpha & \text{und} & f(B) < \alpha \end{array}$$

Aufgrund dieser Definitionen können zwei Mengen überhaupt nur dann streng separiert werden, wenn sie disjunkt sind. Dagegen ist diese Bedingung für schwache Separation nicht notwendig. Betrachtet man nämlich in einer Ebene zwei Kreise A und B, die sich in einem Punkt berühren, so werden diese durch ihre gemeinsame Tangentengerade H (Hyperebene) separiert, jedoch nicht streng separiert.

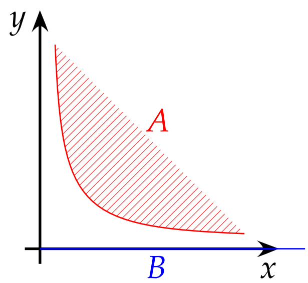


Trotzdem ist die Bedingung, daß zwei Mengen streng separierbar sind, wenn sie disjunkt sind, lediglich notwendig und nicht hinreichend. Dies gilt insbesondere für abgeschlossene konvexe Mengen, wie das folgende Beispiel zeigt:

Die Mengen

$$A := \left\{(x, y) \in \mathbb{R}^2 \;\middle|\; x > 0, y \geq \tfrac{1}{x}\right\}$$
$$B := \left\{(x, y) \in \mathbb{R}^2 \;\middle|\; x \geq 0, y = 0\right\}$$

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.



Bevor wir uns nun dem Problem der Separation zweier Mengen widmen, einige Vorbemerkungen:

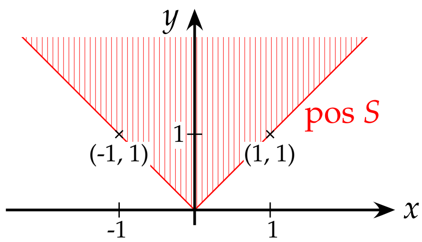
Definition: *Für $i = 1, \dots, m$ seien $\lambda_i \in \mathbb{R}, x_i \in \mathrm{E}^n$. Dann heißt $x = \sum_{i=1}^m \lambda_i x_i$* ***Positivkombination*** *der Punkte $x_1, \dots, x_m$, wenn für alle i gilt: $\lambda_i \geq 0$.*
$\operatorname{pos} S := \{x \in \mathrm{E}^n \mid x = \sum_{i=1}^m \lambda_i x_i$ *mit $\lambda_i \geq 0, x_i \in S$ für alle $i\}$ ist die Menge aller Positivkombinationen der Elemente einer Menge S; sie wird bezeichnet als* ***positive Hülle*** *von S. Insbesondere gilt: $S \subset \operatorname{pos} S$.*

Beispiele:
① Die Menge

$$S := \{(-1, 1), (1, 1)\} \subset \mathbb{R}^2$$

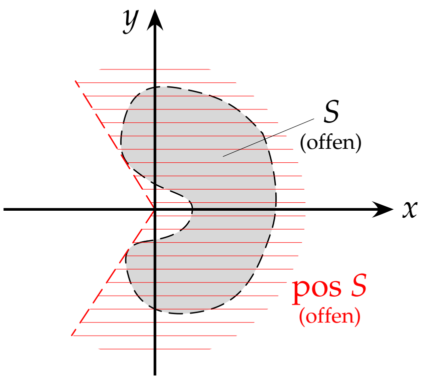
besitzt die positive Hülle

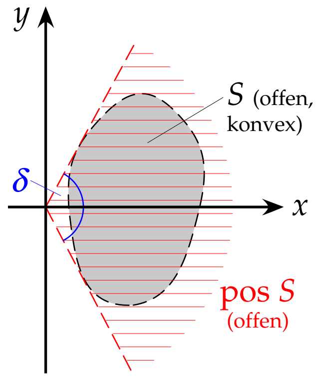
$$\operatorname{pos} S := \{(x, y) \in \mathbb{R}^2 \mid y \geq |x|\}.$$



② Die folgenden Bilder skizzieren die positive Hülle $\operatorname{pos} S$ einer offenen, zusammenhängenden (links) bzw. offenen, konvexen Menge S (rechts) mit $0 \notin S$.

Dabei ist $\operatorname{pos} S$ jeweils ein offener† Sektor mit Eckpunkt 0, da $\operatorname{pos} S$ hier aus allen Ursprungsstrahlen besteht, die ein Element aus S treffen. Für jedes konvexe S ist der Winkel δ des Sektors stets kleiner oder gleich π‡.





Lemma (4.4): *Sei $S \subset \mathrm{E}^2$ offen und konvex, so daß $\mathrm{E}^2 \setminus S \neq \varnothing$. Dann gilt für $x \in \mathrm{E}^2 \setminus S$: Es existiert eine Gerade G mit $x \in G$ und $G \cap S = \varnothing$.*

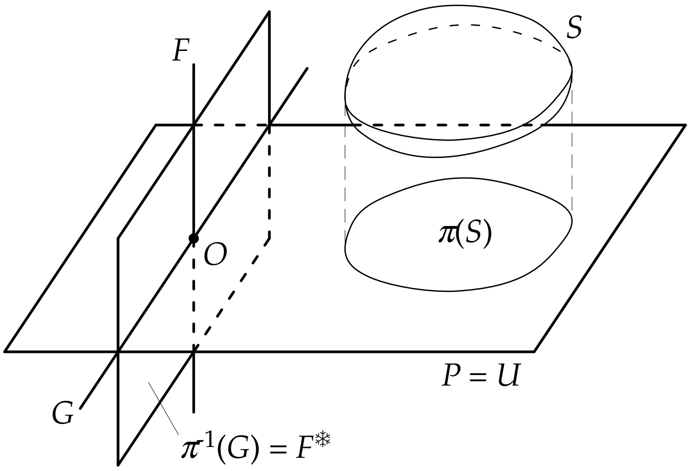
Beweis: O. B. d. A. sei $0 \notin S$, d. h. $x := 0 \in \mathrm{E}^2 \setminus S$. Da die Menge S offen und konvex ist, bildet die positive Hülle $\operatorname{pos} S$ von S einen offenen Sektor mit Eckpunkt $x = 0$ und Winkel $\delta \leq \pi$ (☞ Beispiel ②).

Eine Ursprungsgerade G (d. h. $x = 0 \in G$), die eine Seite des Sektors enthält, erfüllt nun obige Bedingung $G \cap S = \varnothing$, da der Sektor offen ist. ■

† Es existiert zu jedem $s \in \operatorname{pos} S, s \neq 0$, ein $\lambda > 0$ mit $\lambda \cdot s \in S$. Da S offen ist, gibt es ein $\epsilon > 0$, so daß $B(\lambda s, \epsilon) \subset S$; dann ist $B(s, \frac{1}{\lambda}\epsilon) \subset \operatorname{pos} S$ (Strahlensatz!) und somit $\operatorname{pos} S$ offener Sektor.

‡ Bei „$\delta(\operatorname{pos} S) > \pi$" wäre $\operatorname{pos} S$ nicht konvex. Seien $s_1, s_2 \in \operatorname{pos} S$, $s_j = \lambda_j x_j$ mit $\lambda_j > 0$, $x_j \in S$ (konvex). Dann gilt für $\alpha \in [0, 1]$: $\alpha s_1 + (1-\alpha)s_2 = (\alpha\lambda_1)x_1 + ((1-\alpha)\lambda_2)x_2 \in \operatorname{pos} S$ ↯.

Satz (4.5): *Sei $F \subset \mathrm{E}^n$ affiner Unterraum mit* $\dim F = k$. *Sei $S \subset \mathrm{E}^n$ offen und konvex, so daß $F \cap S = \varnothing$.*
Falls $0 \leq k \leq n-2$ ist, existiert ein $(k+1)$-dimensionaler affiner Unterraum $F^{*} \subset \mathrm{E}^n$ mit $F \subset F^{*}$ und $F^{*} \cap S = \varnothing$.



Beweis: Sei $\{e_1, \ldots, e_n\}$ die Standardbasis des E^n.

- Für $0 < k \leq n-2$ sei
  o. B. d. A. $\quad 0 \in F := \mathscr{L}(e_1, \ldots, e_k) \quad (\Rightarrow \dim F = k)$
  und $\quad U := \mathscr{L}(e_{k+1}, \ldots, e_n) \quad (\Rightarrow \dim U = n-k)$.

  Sei $\pi\colon \mathrm{E}^n \to U, (\alpha_1, \ldots, \alpha_n) \mapsto \pi(\alpha_1, \ldots, \alpha_n) := (0, \ldots, 0, \alpha_{k+1}, \ldots, \alpha_n)$ die Projektion des E^n auf U. Nach Aufgabe (4.3) (☞ Anhang A) ist $\pi(S)$ eine relativ offene, konvexe Teilmenge von U.
  Ebenso ist $P := \mathscr{L}(e_{n-1}, e_n) \subset U$ (Ebene durch 0) eine relativ offene, konvexe Menge, weshalb auch $P \cap \pi(S) \subset P$ relativ offen, konvex ist.

  Wegen $F \cap S = \mathscr{L}(e_1, \ldots, e_k) \cap S = \varnothing$ gilt: $0 \notin \pi(S)$, also $0 \notin P \cap \pi(S)$. Gemäß Lemma (4.4) existiert nun in P eine Gerade G mit $x := 0 \in G$ und $G \cap \pi(S) = \varnothing$.

  Für $F^{*} := \pi^{-1}(G) = F \oplus G$ gilt dann: $\dim F^{*} = \dim F + \dim G = k+1$. Außerdem ist $F \subset F^{*}$ und[†] $F^{*} \cap S = \varnothing$ (☞ Abbildung oben).

- Ist $k = 0$, gilt: $F = \{a\}$ für ein $a \in \mathrm{E}^n$. Dann gilt für eine beliebige Ebene P mit $F \subset P$: $S \cap P$ ist eine relative offene, konvexe Teilmenge von P. Wegen $F = \{a\} \subset \mathrm{E}^n \setminus (S \cap P)$ existiert nach Lemma (4.4) in P eine Gerade F^{*} ($\dim F^{*} = 1$) mit $F = \{a\} \subset F^{*}$ und $F^{*} \cap S = \varnothing$. ∎

Korollar (4.6): *Sei $F \subset \mathrm{E}^n$ affiner Unterraum mit* $\dim F = k$. *Sei $S \subset \mathrm{E}^n$ offen und konvex, so daß $F \cap S = \varnothing$. Falls $0 \leq k < n$ ist, existiert eine Hyperebene H mit $F \subset H$ und $H \cap S = \varnothing$.*

Beweis: $((n-1)-k)$-maliges Anwenden von Satz (4.5) liefert die Behauptung, da die Dimension des affinen Raumes jeweils um 1 wächst, bis der Raum eine Hyperebene ist. ∎

[†] Gäbe es ein $x \in (F^{*} \cap S) \subset S$, wähle Zerlegung $x = x_F + x_G$ mit $x_F \in F$ und $x_G \in G \subset P$. Da die Projektion linear ist, gilt: $\pi(S) \ni \pi(x) = \pi(x_F) + \pi(x_G) = 0 + \pi(x_G) = x_G \in G \subset U$ ↯, Widerspruch zu $G \cap \pi(S) = \varnothing$!

Aufbauend auf diesen Vorbemerkungen können wir uns nun der Separation zweier konvexer Mengen durch eine Hyperebene zuwenden:

Satz (4.7): *Seien $A, B \subset \mathrm{E}^n$ konvex, so daß $\operatorname{int} A \neq \varnothing$ und $B \cap \operatorname{int} A = \varnothing$ ist. Dann existiert eine Hyperebene H, die die Mengen A und B separiert.*

Beweis:

- Sei A zunächst zusätzlich offen. Nach Aufgabe (1.10) und (2.6) (☞ Anhang A) gilt: $A - B := \{a - b \in \mathrm{E}^n \mid a \in A, b \in B\}$ ist offen und konvex.

  Da A offen ist, ist $\operatorname{int} A = A$, weshalb wegen $B \cap A = B \cap \operatorname{int} A = \varnothing$ gilt: $0 \notin A - B$. Damit gibt es nach Korollar (4.6) (für $S := A - B, F = \{0\}$) eine *Ursprungshyperebene* $H_0 := [f : 0]$ mit $H_0 \cap (A - B) = \varnothing$.

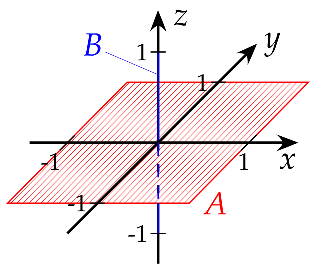
  Da $A - B$ konvex ist, also insbesondere zusammenhängt, können wir o. B. d. A. $f(A - B) > 0$ annehmen. Dann gilt für alle $a \in A$ und $b \in B$: $0 < f(a - b) = f(a) - f(b) \Leftrightarrow f(a) > f(b)$.
  Folglich ist die Menge $f(A) = \{f(a) \mid a \in A\}$ nach unten beschränkt, weshalb $\alpha := \inf\{f(a) \mid a \in A\}$ existiert. Daher werden A und B von der Hyperebene $H := [f : \alpha]$ separiert.

- Ist A nicht offen, existiert gemäß der vorangegangenen Argumentation für die offene Menge $\operatorname{int} A$ eine Hyperebene $H := [f : \alpha]$ mit z. B. $f(\operatorname{int} A) \geq \alpha$ und $f(\operatorname{int} B) \leq \alpha$. Da f stetig ist, folgt: $f(A) \geq \alpha$, weshalb die Hyperebene $H = [f : \alpha]$ auch A und B separiert. ∎

Die in diesem Satz getroffene Voraussetzung $\operatorname{int} A \neq \varnothing$ kann nicht weggelassen werden, wie wir nun sehen werden:

Für die konvexen Mengen

$$A := \{(x, y, z) \in \mathbb{R}^3 \mid |x| \leq 1, |y| \leq 1, z = 0\}$$
$$B := \{(x, y, z) \in \mathbb{R}^3 \mid x = 0, y = 0, |z| \leq 1\}.$$



ist $\operatorname{int} A = \varnothing$ und folglich $B \cap \operatorname{int} A = \varnothing$. Es kann jedoch keine Hyperebene die Mengen A und B separieren (☞ Skizze rechts).

Diese Voraussetzung läßt sich aber ein wenig abschwächen. Der nächste Satz gibt die beste mögliche Charakterisierung dafür an, ob es möglich ist, zwei konvexe Mengen zu separieren. Diesem sind zwei Definitionen und ein Lemma vorangestellt:

Definition (4.8): *Sei $H := [f : \alpha]$ eine Hyperebene im E^n. Dann heißen die Mengen $\{x \in \mathrm{E}^n \mid f(x) \geq \alpha\}$, $\{x \in \mathrm{E}^n \mid f(x) \leq \alpha\}$ **abgeschlossene Halbräume** bezüglich H und $\{x \in \mathrm{E}^n \mid f(x) > \alpha\}$, $\{x \in \mathrm{E}^n \mid f(x) < \alpha\}$ **offene Halbräume** bezüglich H.*

Definition (4.9): *Sei $H := [f : \alpha]$ eine Hyperebene im E^n, sei $S \subset \mathrm{E}^n$. Dann **beschränkt** die Hyperebene H die Menge S, wenn $f(S) \geq \alpha$ oder $f(S) \leq \alpha$ gilt. Beschränkt die Hyperebene H die Menge S nicht, so **schneidet** sie S (d. h., es gibt $x, y \in S$ mit $f(x) < \alpha$ und $f(y) > \alpha$).*

Lemma (4.10): *Eine Hyperebene $H := [f : \alpha]$ im E^n schneidet eine konvexe Menge $S \subset \mathrm{E}^n$ genau dann, wenn die folgenden beiden Bedingungen gelten:*

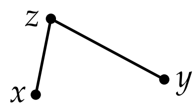
$$\mathrm{I}\colon S \not\subset H, \qquad \mathrm{II}\colon H \cap \operatorname{relint} S \neq \varnothing.$$

Beweis:

„⇒“: Angenommen, H schneidet S. Dann gibt es $x, y \in S$ mit $f(x) < \alpha$ und $f(y) > \alpha$. Da $f(H) = \alpha$ gilt, ist I wahr.

Sei $z \in \operatorname{relint} S$. Da $\operatorname{aff} S$ der kleinste affine Unterraum mit $S \subset \operatorname{aff} S$ ist, entspricht $\operatorname{relint} S$ dem Inneren von S bezüglich $\operatorname{aff} S$, d. h. es gilt: $\operatorname{int}|_{\operatorname{aff} S} S = \operatorname{relint} S$. Damit läßt sich auf $S \subset \operatorname{aff} S$ der Satz (2.9)† anwenden; dies ergibt: $\operatorname{relint} \overline{xz} \subset \operatorname{relint} S$ und $\operatorname{relint} \overline{yz} \subset \operatorname{relint} S$. Wegen $z \in \operatorname{relint} S$ folgt sogar: $(\overline{xz} \setminus \{x\}) \cup (\overline{yz} \setminus \{y\}) = \overline{xz} \cup \overline{yz} \setminus \{x, y\} \subset \operatorname{relint} S$.

Da $(\overline{xz} \cup \overline{yz})$ verbunden, f stetig und $f(x) < \alpha$ und $f(y) > \alpha$ ist, ergibt sich nach dem Zwischenwertsatz:



Falls $f(z) > \alpha$, $\exists\, p \in \overline{xz}$ mit $f(p) = \alpha$ (da $f(x) < \alpha$).
Falls $f(z) < \alpha$, $\exists\, p \in \overline{yz}$ mit $f(p) = \alpha$ (da $f(y) > \alpha$).
Falls $f(z) = \alpha$, wähle $p := z \in \overline{xz} \cup \overline{yz}$.
Offenbar gibt es stets $p \in (\overline{xz} \cup \overline{yz} \setminus \{x, y\}) \subset \operatorname{relint} S$ mit $f(p) = \alpha$. Folglich ist $H \cap \operatorname{relint} S \neq \varnothing$, also gilt II.

„⇐“: Andersherum seien I und II erfüllt, wobei wir o. B. d. A. annehmen:

$$0 \in (H \cap \operatorname{relint} S), \qquad \exists\, x \in S \setminus H\colon f(x) > f(0) = 0.$$

Wegen $\operatorname{int}|_{\operatorname{aff} S} S = \operatorname{relint} S$ gibt es in $\operatorname{aff} S$ eine abgeschlossene Umgebung U von 0 mit $U \subset S$. Die Gerade durch 0 und x (beide in S) liegt nun vollständig im Unterraum $\operatorname{aff} S$ und ihr Schnitt mit U ist eine Strecke $\overline{ab}$ mit $0 \in \overline{ab}, 0 \neq a, b$. Daher existiert ein $\delta > 0$ mit $-\delta x \in S$. Dann ist aber $f(-\delta x) = -\delta f(x) < 0$. Folglich schneidet H die Menge S. ■

Satz (4.11): *Seien $A, B \subset \mathrm{E}^n$ konvex mit $\dim(A \cup B) = n$. Dann werden A, B genau dann von einer Hyperebene H separiert, wenn $\operatorname{relint} A \cap \operatorname{relint} B = \varnothing$ ist.*

Beweis:

„⇒“: Sei H eine Hyperebene, die A und B separiert.
Angenommen, es gibt $x \in \operatorname{relint} A \cap \operatorname{relint} B \subset A \cap B$. Dann gilt $x \in H$, also $H \cap \operatorname{relint} A \neq \varnothing \neq H \cap \operatorname{relint} B$ (Bedingung II von Lemma (4.10)).
Wegen $\dim(A \cup B) = n$ und $\dim H = n - 1$ folgt: $A \not\subset H$ oder $B \not\subset H$ (Bedingung I von Lemma (4.10)).
Gemäß Lemma (4.10) schneidet H eine der Mengen A, B. Widerspruch zur Separation der Mengen A und B durch die Hyperebene H, d. h., bei $\operatorname{relint} A \cap \operatorname{relint} B = \varnothing$ kann keine Hyperebene A und B separieren.

„⇐“: Für zwei Mengen A und B sowie eine Hyperebene H gilt: H separiert A und $B \Leftrightarrow H$ separiert $\operatorname{relint} A$ und $\operatorname{relint} B$.
„⇒“: klar, „⇐“: Sei $H = [f : \alpha]$, dann folgt die Behauptung mit dem Folgenkriterium für Stetigkeit von f.

†Satz (2.9): Sei C konvex, sei $a \in \operatorname{int} C$, $b \in C$. Dann gilt: $\operatorname{relint} \overline{ab} \subset \operatorname{int} C$.

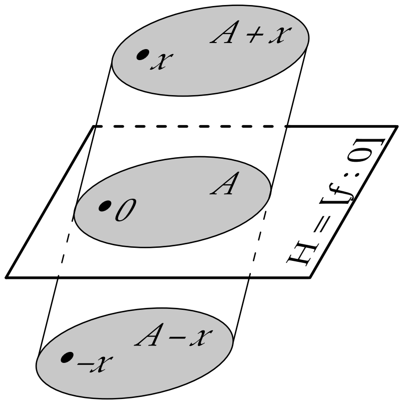
Wir können daher A und B als relativ offene, konvexe Mengen annehmen mit $A \cap B = \varnothing$. Den Schluß des Beweises liefert eine (absteigende) Induktion nach der Dimension von A:

Induktionsanfang: Sei $\dim A = n$. Dann ist $\operatorname{int} A \neq \varnothing$, weshalb nach Satz (4.7) eine A, B separierende Hyperebene existiert.

Induktionsvoraussetzung: Falls die Dimension von einer von zwei relativ offenen, konvexen Mengen größer oder gleich einem k ist mit $1 \leq k \leq n$, gebe es eine die Mengen separierende Hyperebene.

Induktionsschluß: Es sei $\dim A = k - 1$. Dann können wir o. B. d. A. $0 \in A$ annehmen. Sei $H := [f : 0]$ eine Hyperebene mit $A \subset H$.
Sei $x \in \mathrm{E}^n$ mit $f(x) > 0$, also $x \notin A$.
Wir betrachten die beiden konvexen und relativ offenen Mengen

$$C := \operatorname{conv}(A \cup (A + x)),$$
$$D := \operatorname{conv}(A \cup (A - x)).$$



Falls $B \cap C \neq \varnothing$ und $B \cap D \neq \varnothing$, folgt, da B und $C \cup D$ konvex sind: $B \cap A \neq \varnothing$, Widerspruch!

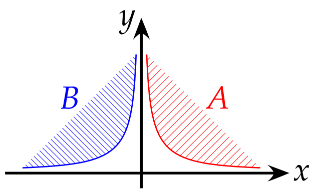
Daher sei etwa $B \cap C = \varnothing$. Weil $\dim C = k$ ist, existiert gemäß der Induktionsannahme eine Hyperebene H, die C und B separiert. Wegen $A \subset C$ separiert diese Hyperebene dann auch A und B. ■

Wesentlich ist bei diesem Satz die Voraussetzung $\dim(A \cup B) = n$. Denn wäre $\dim(A \cup B) < n$, gäbe es eine Hyperebene $H := [f : \alpha]$ mit $(A \cup B) \subset H$. Für diese gilt: $f(A) \leq f(H) = \alpha$ und $f(B) \geq f(H) = \alpha$, d. h. A und B werden durch H stets separiert – unabhängig davon, ob relint A und relint B disjunkt sind. Eine derartige Separation heißt *uneigentliche Separation*; sie hat wenig Wert.

Nachdem wir uns bislang der Separation gewidmet haben, behandeln wir abschließend die strenge Separation zweier konvexer Mengen.

Es ist einfach, Beispiele für zwei disjunkte nichtkompakte Mengen zu finden, die nicht streng separierbar sind, beispielsweise die folgenden:

$$A := \{(x,y) \in \mathbb{R}^2 \mid y \geq 0, xy \geq 1\}$$
$$B := \{(x,y) \in \mathbb{R}^2 \mid y \geq 0, xy \leq -1\}.$$



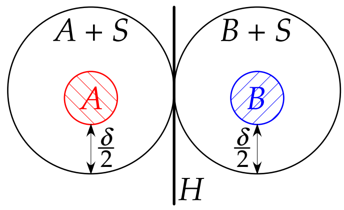
Streng separieren lassen sich stets zwei disjunkte, kompakte, konvexe Mengen; in der Tat gilt nämlich der folgende, etwas strengere Satz:

Satz (4.12): *Seien $A, B \subset \mathrm{E}^n$ nichtleer und konvex. Zudem sei A kompakt und B abgeschlossen. Die Mengen A und B werden genau dann von einer Hyperebene H streng separiert, wenn $A \cap B = \varnothing$ ist.*

Beweis:

„⇐“: Sei $A \cap B = \varnothing$. Unter dem Abstand der Mengen A und B versteht man: $d(A,B) := \inf\{d(a,b) \mid a \in A, b \in B\}$; für diesen gilt, da A kompakt und B abgeschlossen ist: $\delta := d(A,B) > 0$.†

Sei $S := B(0, \frac{\delta}{2})$ die offene Kugel um 0 mit Radius $\frac{\delta}{2}$. Dann sind $A+S, B+S \subset \mathrm{E}^n$ disjunkt, offen und konvex.



Gemäß Satz (4.11) existiert eine Hyperebene $H = [f : \alpha]$, die $A+S$ und $B+S$ separiert, etwa mit $f(A+S) \leq \alpha$ und $f(B+S) \geq \alpha$.

Wegen $f(a+s) \leq f(b+s) \Leftrightarrow f(a) \leq f(b)$ für alle $a \in A, b \in B, s \in S$ separiert H auch A und B. Gäbe es nun $a \in A$, $b \in B$ mit $f(a) = f(b)$, folgt für $s, t \in S$ mit $f(s) < f(t)$: $f(a+t) \leq f(b+s) \Leftrightarrow f(t) \leq f(s)$, Widerspruch! Daher werden A und B durch H streng separiert.

„⇒“: Klar (zu zeigen wäre: H separiert A und B streng $\Rightarrow A \cap B = \varnothing$). ■

Läßt man die Forderung „A, B konvex“ weg, erhält man:

Satz (4.13): *Seien $A, B \subset \mathrm{E}^n$ nichtleer und kompakt. Die Mengen A und B werden genau dann von einer Hyperebene H streng separiert, wenn* $\operatorname{conv} A \cap \operatorname{conv} B = \varnothing$ *ist.*

Beweis:

„⇐“: Sei zunächst $\operatorname{conv} A \cap \operatorname{conv} B = \varnothing$. Da die konvexe Hülle einer kompakten Menge kompakt ist, was in Satz (2.30) bewiesen wurde, existiert nach Satz (4.12) eine Hyperebene H, die $\operatorname{conv} A$ und $\operatorname{conv} B$ streng separiert. Wegen $A \subset \operatorname{conv} A, B \subset \operatorname{conv} B$ separiert diese Hyperebene H auch A und B streng.

„⇒“: Sei jetzt $H := [f : \alpha]$ eine Hyperebene, die A und B streng separiert, etwa mit $f(A) < \alpha$ und $f(B) > \alpha$. Sei $x = \sum_{i=1}^k \lambda_i x_i$ eine Konvexkombination von Elementen aus A. Dann gilt wegen $\sum_{i=1}^k \lambda_i = 1$:

$$f(x) = f\left(\sum_{i=1}^k \lambda_i x_i\right) = \sum_{i=1}^k \lambda_i f(x_i) < \sum_{i=1}^k \lambda_i \alpha = \alpha$$

Also ist $f(\operatorname{conv} A) < \alpha$ und entsprechend $f(\operatorname{conv} B) > \alpha$. Daher separiert die Hyperebene $H = [f : \alpha]$ $\operatorname{conv} A$ und $\operatorname{conv} B$ streng, weshalb $\operatorname{conv} A$ und $\operatorname{conv} B$ nach Satz (4.12) disjunkt sind. ■

†Wäre $\delta = 0$, gäbe es Folgen $(a_n)_{n\in\mathbb{N}}$ in A und $(b_n)_{n\in\mathbb{N}}$ in B mit $0 \leq d(a_n, b_n) \leq \frac{1}{n}$. Da A kompakt ist, können wir $\lim\limits_{n\to\infty} a_n =: a \in A$ annehmen (Bolzano-Weierstraß). Weil d stetig ist, folgt: $0 = \lim\limits_{n\to\infty} \frac{1}{n} \geq \lim\limits_{n\to\infty} d(a_n, b_n) = d(a, \lim\limits_{n\to\infty} b_n) \geq 0$, also: $a = \lim\limits_{n\to\infty} b_n$. Wegen der Abgeschlossenheit von B müßte gelten: $a \in B$, Widerspruch zu $A \cap B = \varnothing$.

Ein Spezialfall des Satzes (4.13) ist folgender:

Ein Punkt x kann genau dann von einer kompakten Menge S durch eine Hyperebene H streng separiert werden, wenn $S \notin \operatorname{conv} S$ ist.

Dies kann mit Hilfe des Satzes von Carathéodory (☞ Anhang B) umformuliert werden zu:

Ein Punkt x kann genau dann von einer kompakten Menge $S \in \mathrm{E}^n$ durch eine Hyperebene H streng separiert werden, wenn zu jeder Teilmenge T aus $n+1$ oder weniger Elementen aus S eine Hyperebene existiert, welche x und T streng separiert.

Allgemeiner gilt für zwei kompakte Mengen:

Satz (4.14): *Seien $A, B \subset \mathrm{E}^n$ nichtleer und kompakt. Die Mengen A und B werden genau dann von einer Hyperebene H streng separiert, wenn es zu jedem $T \subset B$ mit $^\# T \leq n+1$ eine Hyperebene gibt, die A und T streng separiert.*

Beweis:

„⇒“: Ist H eine Hyperebene, die A und B streng separiert, so separiert H insbesondere A von jedem $T \subset B$ streng.

„⇐“: Für jedes $T \subset B$ mit $^\# T \leq n+1$ existiere eine Hyperebene, die A und T streng separiert. Gilt $\operatorname{conv} A \cap \operatorname{conv} B = \varnothing$, folgt die Behauptung nach Satz (4.13). Angenommen, es gebe ein $x \in \operatorname{conv} A \cap \operatorname{conv} B$, dann folgt gemäß dem Satz von Carathéodory:

$$\exists\, a_1, \ldots, a_{n+1} \in A : x = \sum_{i=1}^{n+1} \lambda_i a_i \quad \text{mit} \quad \sum_{i=1}^{n+1} \lambda_i = 1,\, \lambda_i \geq 0 \;\forall\, i,$$

$$\exists\, b_1, \ldots, b_{n+1} \in B : \; x = \sum_{i=1}^{n+1} \mu_i b_i \quad \text{mit} \quad \sum_{i=1}^{n+1} \mu_i = 1,\, \mu_i \geq 0 \;\forall\, i.$$

Zu $T := \{b_1, \ldots, b_{n+1}\} \subset B$ mit $^\# T \leq n+1$ existiert dann eine Hyperebene $H := [f : \alpha]$, die A und T streng separiert. Gilt o. B. d. A. $f(A) > \alpha$ und $f(B) < \alpha$, folgt:

$$f(x) = f\left(\sum_{i=1}^{n+1} \lambda_i a_i\right) = \sum_{i=1}^{n+1} \lambda_i f(a_i) > \sum_{i=1}^{n+1} \lambda_i \alpha = \alpha,$$

$$f(x) = f\left(\sum_{i=1}^{n+1} \mu_i b_i\right) = \sum_{i=1}^{n+1} \mu_i f(b_i) \overset{\lightning}{<} \sum_{i=1}^{n+1} \mu_i \alpha = \alpha. \quad \text{Widerspruch!} \quad \blacksquare$$

Anhang

A Aufgaben

Definition: *Seien $A, B \subset \mathrm{E}^n$. Dann verstehen wir unter der **Summe** von A und B die Menge $A + B := \{a + b \in \mathrm{E}^n \mid a \in A, b \in B\}$.*

Aufgabe (1.10): *Sei $A \subset \mathrm{E}^n$ offen. Zu zeigen: Für jedes $B \subset \mathrm{E}^n$ ist die Summe von A und B offen.*

Aufgabe (2.6): *Zu zeigen: Sind $A, B \subset \mathrm{E}^n$ konvex, so ist die Summe von A und B konvex.*

Definition: *Sei $A \subset X \subset \mathrm{E}^n$. Eine Teilmenge $B \subset A$ heißt **relativ offen** bezüglich A, wenn es eine offene Menge U in X gibt, so daß $B = U \cap A$ ist.*

Aufgabe (4.3): *Sei $\{e_1, \dots, e_n\}$ Standardbasis des E^n. Sei $U := \mathscr{L}(e_{k+1}, \dots, e_n)$, d. h. $\dim U = n - k$ mit $0 < k \leq n - 2$.*
Sei $\pi\colon \mathrm{E}^n \to U, (\alpha_1, \dots, \alpha_n) \mapsto \pi(\alpha_1, \dots, \alpha_n) := (0, \dots, 0, \alpha_{k+1}, \dots, \alpha_n)$ die Projektion des E^n auf U.
Zu zeigen: Falls $S \subset \mathrm{E}^n$ offen und konvex ist, ist $\pi(S) \subset U$ relativ offen, konvex.

B Satz von Carathéodory

Satz (2.23 – Carathéodory): *Sei $S \subset \mathrm{E}^n$, $S \neq \varnothing$. Dann gilt: Jedes $x \in \operatorname{conv} S$ ist darstellbar als Konvexkombination von höchstens $(n+1)$ Elementen aus S:*

$$\operatorname{conv} S = \left\{ \sum_{j=1}^{n+1} \lambda_j x_j \;\middle|\; x_j \in S, \lambda_j \geq 0, \sum_{j=1}^{n+1} \lambda_j = 1 \right\}.$$