Konvexe Analysis *

Martin Brokate †

Inhaltsverzeichnis

*Vorlesungsskript, SS 2009

†Zentrum Mathematik, TU München

Literatur

J.M. Borwein, A.S. Lewis: Convex Analysis and Nonlinear Optimization. Theory and Examples. Springer, New York 2006.

I. Ekeland, R. Temam: Convex Analysis and Variational Problems. North Holland, Amsterdam 1976; Nachdruck und Update: SIAM, Philadelphia 1999.

J.-B. Hiriart-Urruty, C. Lemaréchal: Fundamentals of Convex Analysis. Springer-Verlag, Berlin 2001.

R.B. Holmes: Geometric Functional Analysis and its Applications. Springer-Verlag, New York 1975.

J. Marti: Konvexe Analysis. Birkhäuser, Basel 1977.

R.T. Rockafellar: Convex Analysis. Princeton University Press, Princeton 1970.

H. Tuy: Convex Analysis and Global Optimization. Kluwer, Dordrecht 1998.

F.A. Valentine: Konvexe Mengen. Bibliographisches Institut, Mannheim 1968.

1 Affine Mengen

Definition 1.1 (Affine Menge) *Sei X Vektorraum, $M \subset X$. M heißt affin, wenn*

$$\lambda x + (1-\lambda)y \in M \tag{1.1}$$

für alle $x, y \in M$ und alle $\lambda \in \mathbb{R}$. □

Eine affine Menge enthält also zu je zwei Punkten die Gerade, die durch die beiden Punkte verläuft.

Satz 1.2 *Sei X Vektorraum, $M \subset X$. Dann ist M Unterraum genau dann, wenn M eine affine Menge mit $0 \in M$ ist.*

Beweis: "$\Rightarrow$": klar.
"$\Leftarrow$": Ist $0 \in M$, so ist $\lambda x = \lambda x + (1-\lambda)0 \in M$ für alle $x \in M$ und alle $\lambda \in \mathbb{R}$. Sind $x, y \in M$, so ist auch

$$x + y = 2 \cdot \left(\frac{1}{2}x + \frac{1}{2}y \right) \in M \,. \tag{1.2}$$

□

Affine Mengen entstehen durch Translation von Unterräumen.

Satz 1.3 (Charakterisierung, affine Dimension)
Sei X Vektorraum, $M \subset X$, $M \neq \emptyset$. Dann ist M affin genau dann, wenn es ein $a \in M$ und einen Unterraum U gibt mit

$$M = a + U \,. \tag{1.3}$$

In diesem Fall gilt

$$U = M - M = \{x - y : x, y \in M\} \,. \tag{1.4}$$

Wir definieren die ***affine Dimension*** *von M durch*

$$\dim(M) = \dim(U) \,. \tag{1.5}$$

Beweis: "$\Leftarrow$": Sind $a + x, a + y \in a + U$, so ist auch

$$\lambda(a + x) + (1-\lambda)(a + y) = a + (\lambda x + (1-\lambda)y) \in a + U \,.$$

"$\Rightarrow$": Wähle $a \in M$, setze $U = -a + M$. Dann ist $0 \in U$ und U affin, also U Unterraum. Wir beweisen (1.4).
"$\subset$": $U = M - a \subset M - M$.
"$\supset$": Seien $x, y \in M$. Dann ist $x - y = (x - a) - (y - a) \in U - U = U$. □

Vorsicht: Falls M nicht selbst Unterraum ist, so ist

$$\dim(M) = \dim(\mathrm{span}\,(M)) - 1 \,, \tag{1.6}$$

d.h. die Dimension des von M aufgespannten Unterraums ist in diesem Fall um 1 größer als die affine Dimension von M.

Ist $M \subset X$ affin, $M = a + U$, U Unterraum, und ist X/U der Quotientenraum von X nach U, so gilt

$$M = \{x : p(x) = [a]\}\,, \tag{1.7}$$

wobei $p : X \to X/U$ die kanonische lineare Abbildung ist, welche jedes x auf die zugehörige Äquivalenzklasse abbildet. Jede affine Menge läßt sich also als Urbild eines Punktes unter einer linearen Abbildung darstellen. Im $\mathbb{R}^n$ kann man das auch mit Matrizen und Vektoren hinschreiben.

Satz 1.4 *Sei $M \subset \mathbb{R}^n$ affin mit $\dim(M) = d$. Dann gibt es eine Matrix $A \in \mathbb{R}^{(m,n)}$ und einen Vektor $b \in \mathbb{R}^m$ mit*

$$x \in M \qquad \Leftrightarrow \qquad Ax = b\,, \tag{1.8}$$

wobei $m = n - d$.

Beweis: Sei $M = a + U$, U Unterraum, sei $\{v_1, \ldots, v_m\}$ Basis von $U^\perp$. Sei A diejenige Matrix, deren i-te Zeile gerade der Vektor v_i ist, sei $b = Aa$. Für $y \in \mathbb{R}^n$ gilt

$$Ay = 0 \quad \Leftrightarrow \quad y \in \ker(A) \quad \Leftrightarrow \quad y \in U\,, \tag{1.9}$$

da $U \subset \ker(A)$ nach Konstruktion ($\langle v_i, y\rangle = 0$ für alle $y \in U$) und $\dim(U) = n - m = \dim\ker(A)$. Also gilt mit $x = a + y$

$$b = Ax = A(a + y) \quad \Leftrightarrow \quad Ay = 0 \quad \Leftrightarrow \quad x = a + y \in a + U = M\,. \tag{1.10}$$

□

Es ist klar, daß beliebige Durchschnitte

$$\bigcap_{i \in I} M_i \tag{1.11}$$

von affinen Mengen M_i wieder affin sind.

Definition 1.5 (Affine Hülle) *Sei X Vektorraum, $S \subset X$. Dann heißt*

$$\mathrm{aff}\,(S) = \bigcap_{\substack{S \subset M \subset X \\ M \text{ affin}}} M \tag{1.12}$$

die affine Hülle von S in X. □

Falls $0 \in S$, so gilt

$$\mathrm{aff}\,(S) = \mathrm{span}\,(S)\,, \tag{1.13}$$

da in diesem Fall die Mengen M in (1.12) Unterräume sind.

Lemma 1.6 *Sei X Vektorraum, $S \subset X$, $a \in X$. Dann gilt*

$$\mathrm{aff}\,(S) = a + \mathrm{aff}\,(S - a)\,. \tag{1.14}$$

Falls $a \in S$, so gilt

$$\mathrm{aff}\,(S) = a + \mathrm{aff}\,(S - a) = a + \mathrm{span}\,(S - a)\,. \tag{1.15}$$

Beweis: Wegen (1.13) folgt (1.15) aus (1.14), falls $a \in S$. Wir beweisen (1.14).
"$\subset$": Sei N affin, $S - a \subset N$. Dann ist $S \subset a + N$ und $a + N$ affin, also auch $\mathrm{aff}\,(S) \subset a + N$ und $\mathrm{aff}\,(S) - a \subset N$. Da N beliebig war, folgt $\mathrm{aff}\,(S) - a \subset \mathrm{aff}\,(S - a)$.
"$\supset$": Nach dem eben Bewiesenen gilt $\mathrm{aff}\,(S - a) \subset -a + \mathrm{aff}\,(S - a + a)$. $\square$

Im Spezialfall $S = \{x_0, \ldots, x_n\}$ erhalten wir

$$\mathrm{aff}\,(\{x_0, \ldots, x_n\}) = x_0 + \mathrm{span}\,(\{x_1 - x_0, \ldots, x_n - x_0\})\,. \tag{1.16}$$

Definition 1.7 (Affine Unabhängigkeit) *Sei X Vektorraum. Eine Teilmenge $\{x_0, \ldots, x_n\}$ von X heißt affin unabhängig, wenn $\dim \mathrm{aff}\,(\{x_0, \ldots, x_n\})) = n$.* $\square$

Nach (1.16) ist das genau dann der Fall, wenn $\{x_1 - x_0, \ldots, x_n - x_0\}$ linear unabhängig ist.

Satz 1.8 (Baryzentrische Koordinaten)
Sei X Vektorraum, $S = \{x_0, \ldots, x_n\}$ affin unabhängig. Dann läßt sich jedes $x \in \mathrm{aff}\,(S)$ eindeutig darstellen in der Form

$$x = \sum_{i=0}^{n} \lambda_i x_i\,, \quad \sum_{i=0}^{n} \lambda_i = 1\,. \tag{1.17}$$

Wir nennen $\lambda_0, \ldots, \lambda_n$ die baryzentrischen Koordinaten von x bezüglich $x_0, \ldots, x_n$.

Beweis: Aus (1.16) folgt die Existenz und Eindeutigkeit der Darstellung

$$x - x_0 = \sum_{i=1}^{n} \lambda_i (x_i - x_0)\,, \tag{1.18}$$

also folgt (1.17) mit $\lambda_0 = 1 - \sum_{i=1}^{n} \lambda_i$. Ist

$$x = \sum_{i=0}^{n} \mu_i x_i\,, \quad \sum_{i=0}^{n} \mu_i = 1\,, \tag{1.19}$$

so ist

$$x - x_0 = (\mu_0 - 1)x_0 + \sum_{i=1}^{n} \mu_i x_i = \sum_{i=1}^{n} \mu_i (x_i - x_0)\,, \tag{1.20}$$

also folgt $\mu_i = \lambda_i$ für $i = 1, \ldots, n$ und damit auch $\mu_0 = \lambda_0$. $\square$

Definition 1.9 (Affine Abbildung) *Seien X, Y Vektorräume. Eine Abbildung $f : X \to Y$ heißt affin, wenn*

$$f(\lambda x + (1 - \lambda)y) = \lambda f(x) + (1 - \lambda) f(y) \tag{1.21}$$

gilt für alle $x, y \in X$ und alle $\lambda \in \mathbb{R}$. $\square$

Die Gleichung (1.21) bedeutet, daß die Gerade durch x und y abgebildet wird auf die Gerade durch $f(x)$ und $f(y)$.

Satz 1.10 *Seien X, Y Vektorräume. Eine Abbildung $f : X \to Y$ ist affin genau dann, wenn es eine lineare Abbildung $g : X \to Y$ und ein $a \in Y$ gibt mit*

$$f(x) = g(x) + a\,. \tag{1.22}$$

Beweis: Man rechnet unmittelbar nach, dass (1.21) aus (1.22) folgt. Umgekehrt: Setze $a = f(0)$, $g(x) = f(x) - a$. □

Satz 1.11 *Seien X, Y Vektorräume, $f : X \to Y$ affin. Ist $M \subset X$ affin, so ist auch $f(M)$ affin; ist $N \subset Y$ affin, so ist auch $f^{-1}(N)$ affin. Ist f darüber hinaus invertierbar, so ist auch f^{-1} affin.*

Beweis: Die ersten beiden Behauptungen folgen direkt aus der Definition. Sei nun f invertierbar. Die durch $g(x) = f(x) - f(0)$ definierte Abbildung ist linear und invertierbar, und es gilt

$$f^{-1}(y) = g^{-1}(y) - g^{-1}(f(0))\,, \tag{1.23}$$

was man nachprüft, indem man $y = f(x)$ in die rechte Seite einsetzt. □

2 Konvexe Mengen

Definition 2.1 (Konvexe Menge) *Sei X Vektorraum, $K \subset X$. K heißt konvex, wenn*

$$\lambda x + (1-\lambda) y \in K \tag{2.1}$$

für alle $x, y \in K$ und alle $\lambda \in [0,1]$. Wir schreiben

$$[x,y] = \{\lambda x + (1-\lambda) y : \lambda \in [0,1]\} \tag{2.2}$$

für die Verbindungsstrecke von x nach y, sowie auch $(x,y]$, $[x,y)$, (x,y) für die Strecke ohne die entsprechenden Endpunkte. Wir setzen $[x,x] = \{x\}$. □

Eine konvexe Menge enthält also zu je zwei Punkten deren Verbindungsstrecke.

Eine endliche Summe der Form

$$\sum_{i=1}^{n} \lambda_i x_i\,, \quad x_i \in X\,,\ \lambda_i \geq 0 \text{ mit } \sum_{i=1}^{n} \lambda_i = 1\,, \tag{2.3}$$

heißt Konvexkombination der Vektoren $x_1, \ldots, x_n$.

Satz 2.2 *Sei X Vektorraum. Eine Teilmenge $K \subset X$ ist konvex genau dann, wenn*

$$\sum_{i=1}^{n} \lambda_i x_i \in K \tag{2.4}$$

gilt für alle $n \in \mathbb{N}$ und alle Konvexkombinationen von Elementen $x_1, \ldots, x_n \in K$.

Beweis: Für "$\Leftarrow$" ist nichts zu zeigen. "$\Rightarrow$": wird mit Induktion über n bewiesen. $n = 2$ entspricht der Definition der Konvexität. Induktionsschritt $n-1 \to n$: Sei

$$x = \sum_{i=1}^{n} \lambda_i x_i\,, \quad x_i \in K\,,\ 1 \leq i \leq n\,, \tag{2.5}$$

eine Konvexkombination. Wähle λ_j mit $\lambda_j < 1$, dann gilt

$$x = \lambda_j x_j + (1-\lambda_j) y\,, \quad y = \sum_{i \neq j} \frac{\lambda_i}{1-\lambda_j} x_i\,, \quad \sum_{i \neq j} \frac{\lambda_i}{1-\lambda_j} = 1\,, \tag{2.6}$$

also $y \in K$ nach Induktionsvoraussetzung und damit $x \in K$. □

Beliebige Durchschnitte

$$\bigcap_{i \in I} K_i \tag{2.7}$$

von konvexen Mengen $K_i \subset X$ sind offensichtlich ebenfalls konvex.

Definition 2.3 (Konvexe Hülle) *Sei X Vektorraum, $S \subset X$. Dann heißt*

$$\operatorname{co}(S) = \bigcap_{\substack{S \subset K \subset X \\ K \text{ konvex}}} K \tag{2.8}$$

die konvexe Hülle von S in X. □

Satz 2.4 *Sei X Vektorraum, $S \subset X$. Dann gilt*

$$\operatorname{co}(S) = \{x : x \in X,\ x \text{ ist Konvexkombination von Elementen in } S\}\,. \tag{2.9}$$

Beweis: Wir bezeichnen mit C die durch die rechte Seite von (2.9) definierte Menge, zu zeigen $\operatorname{co}(S) = C$.
"$\supset$": Ist $K \supset S$ konvex, so gilt $K \supset C$ nach Satz 2.2, also $\operatorname{co}(S) \supset C$.
"$\subset$": Da $S \subset C$, genügt es zu zeigen, daß C konvex ist. Seien also $x, y \in C$,

$$x = \sum_{i=1}^{n} \lambda_i x_i\,, \quad y = \sum_{j=1}^{m} \mu_j y_j\,, \tag{2.10}$$

dann gilt für $\alpha \in [0,1]$

$$\alpha x + (1-\alpha) y = \sum_{i=1}^{n} \alpha \lambda_i x_i + \sum_{j=1}^{m} (1-\alpha) \mu_j y_j \in C\,, \tag{2.11}$$

da die Koeffizienten die an eine Konvexkombination gestellten Bedingungen erfüllen. □

Satz 2.5 (Carathéodory)
Sei $S \subset \mathbb{R}^n$. Dann lässt sich jedes $x \in \operatorname{co}(S)$ als Konvexkombination mit $n+1$ (oder weniger) Elementen aus S darstellen, das heißt, es gibt $x_i \in S$ und $\lambda_i \in [0,1]$ mit

$$x = \sum_{i=1}^{n+1} \lambda_i x_i\,. \tag{2.12}$$

Beweis: Nach Satz 2.4 gibt es für $x \in \operatorname{co}(S)$ eine Darstellung

$$x = \sum_{i=1}^{N} \lambda_i x_i\,, \quad x_i \in S\,, \lambda_i \in [0,1]\,.$$

Es genügt zu zeigen: Ist $N > n+1$, so gibt es eine weitere Darstellung

$$x = \sum_{i=1}^{N} \mu_i x_i\,, \quad x_i \in S\,, \mu_i \in [0,1]\,, \tag{2.13}$$

in der mindestens ein μ_j gleich Null ist. Gilt bereits $\lambda_j = 0$ für irgendein j, so ist nichts zu tun. Andernfalls betrachten wir die Vektoren

$$y_i = \lambda_i \begin{pmatrix} x_i \\ 1 \end{pmatrix} \in \mathbb{R}^{n+1}\,, \quad 1 \le i \le N\,.$$

Da $N > n+1$, ist $\{y_1, \dots, y_N\}$ linear abhängig, also gibt es eine Linearkombination

$$\sum_{i=1}^{N} \alpha_i y_i = 0 \tag{2.14}$$

mit $\alpha_i \in \mathbb{R}$, die nicht alle gleich Null sind. Die letzte Komponentengleichung in (2.14) ist

$$\sum_{i=1}^{N} \alpha_i \lambda_i = 0\,. \tag{2.15}$$

Für

$$\alpha_j = \min_{1 \le i \le N} \alpha_i\,,$$

gilt $\alpha_j < 0$ wegen (2.15), da $\lambda_i > 0$ für alle i. Die gesuchte Konvexkombination ergibt sich nun für

$$\mu_i = \left(1 - \frac{\alpha_i}{\alpha_j}\right) \lambda_i\,, \quad 1 \le i \le N\,.$$

Es ist nämlich $\mu_j = 0$ sowie $\mu_i \ge 0$ für $i \ne j$, da $\alpha_i/\alpha_j \le 1$. und es folgt mit (2.15)

$$\sum_{i=1}^{N} \mu_i = \sum_{i=1}^{N} \lambda_i - \frac{1}{\alpha_j} \sum_{i=1}^{N} \alpha_i \lambda_i = \sum_{i=1}^{N} \lambda_i = 1\,,$$

sowie

$$\sum_{i=1}^{N} \mu_i x_i = \sum_{i=1}^{N} \lambda_i x_i - \frac{1}{\alpha_j} \sum_{i=1}^{N} \alpha_i \lambda_i x_i = \sum_{i=1}^{N} \lambda_i x_i = x\,.$$

□

Definition 2.6 (Simplex) *Sei $\{x_0, \ldots, x_n\} \subset X$ affin unabhängig. Dann heißt*

$$\operatorname{co}(\{x_0, \ldots, x_n\}) \tag{2.16}$$

ein (n-dimensionales) Simplex, die x_i heißen die Ecken des Simplex. □

Jeder Punkt x eines Simplex $\operatorname{co}(\{x_0, \ldots, x_n\})$ läßt sich eindeutig als Konvexkombination

$$x = \sum_{i=0}^{n} \lambda_i x_i \tag{2.17}$$

seiner Ecken mit den baryzentrischen Koordinaten $\lambda_0, \ldots, \lambda_n$ schreiben.

Definition 2.7 (Affine Dimension einer konvexen Menge)
Sei X Vektorraum, $K \subset X$ konvex. Wir definieren die affine Dimension von K als

$$\dim(K) = \dim(\operatorname{aff}(K))\,. \tag{2.18}$$

□

Satz 2.8 *Sei $K \subset \mathbb{R}^n$ konvex. Dann gilt*

$$\dim K = \max_{\substack{S \subset K \\ S \text{ Simplex}}} \dim S\,, \tag{2.19}$$

das heißt, eine konvexe Menge der affinen Dimension m enthält ein m-dimensionales Simplex. Für jedes solche Simplex gilt $K \subset \operatorname{aff}(S)$.

Beweis: "$\geq$" ist klar. Sei nun $S = \operatorname{co} E$, $E = \{x_0, \ldots, x_m\}$ ein Simplex maximaler Dimension mit $S \subset K$. Wir zeigen, dass $K \subset \operatorname{aff}(S)$, daraus folgt auch "$\leq$". Wir nehmen im Gegenteil an, daß es ein $x \in K \setminus \operatorname{aff}(S)$ gibt. Dann ist

$$x - x_0 \in (K - x_0) \setminus \operatorname{span}(S - x_0)\,, \tag{2.20}$$

also ist $\{x_1 - x_0, \ldots, x_m - x_0, x - x_0\}$ linear unabhängig und $S' = \operatorname{co}(E \cup \{x\})$ ein Simplex mit $S' \subset K$ und $\dim(S') = m + 1$ im Widerspruch zur Maximalität von S. □

Lemma 2.9 *Sei X Vektorraum, seien $K_i \subset X$ konvex und $\lambda_i \in \mathbb{R}$ für $1 \leq i \leq n$. Dann ist $\sum_{i=1}^n \lambda_i K_i$ konvex.*

Beweis: Ist K konvex und $\lambda \in \mathbb{R}$, so ist auch λK konvex (folgt direkt aus der Definition). Sind K_1 und K_2 konvex, so ist auch $K_1 + K_2$ konvex: Seien $x, y \in K_1 + K_2$, $x = x_1 + x_2$, $y = y_1 + y_2$, wobei $x_1, y_1 \in K_1$ und $x_2, y_2 \in K_2$ sind. Dann gilt

$$\lambda x + (1-\lambda)y = [\lambda x_1 + (1-\lambda)y_1] + [\lambda x_2 + (1-\lambda)y_2] \in K_1 + K_2\,. \tag{2.21}$$

□

Satz 2.10 *Sei X Vektorraum, $K \subset X$ konvex. Dann gilt*

$$(\lambda + \mu)K = \lambda K + \mu K \tag{2.22}$$

für alle $\lambda, \mu \geq 0$.

Beweis: "$\subset$": klar.
"$\supset$": Trivial falls $\lambda = \mu = 0$. Andernfalls gilt nach Definition der Konvexität

$$\frac{\lambda}{\lambda + \mu} K + \frac{\mu}{\lambda + \mu} K \subset K\,, \tag{2.23}$$

woraus nach Multiplikation mit $\lambda + \mu$ die Behauptung folgt. □

Lemma 2.11 *Seien X, Y Vektorräume, $f : X \to Y$ affin. Ist $K \subset X$ konvex, so ist auch $f(K)$ konvex. Ist $L \subset Y$ konvex, so ist auch $f^{-1}(L)$ konvex.*

Beweis: Direkt aus der Definition. □

Definition 2.12 (Algebraisches Inneres) *Sei X Vektorraum, $S \subset X$. Ein $x \in S$ heißt algebraisch innerer Punkt von S, falls es zu jedem $h \in X$ ein $\delta > 0$ gibt mit $[x, x + \delta h] \subset S$. Die Menge aller algebraisch inneren Punkte von S bezeichnen wir mit $\operatorname{aint}(S)$. S heißt algebraisch offen, wenn $S = \operatorname{aint}(S)$.* □

Satz 2.13 *Sei X Vektorraum, $K \subset X$ konvex. Ist $x \in \operatorname{aint}(K)$ und $y \in K$, so gilt $[x, y) \subset \operatorname{aint}(K)$.*

Beweis: Sei $a \in [x, y)$, $a = \lambda x + (1-\lambda)y$, $0 < \lambda \leq 1$. Sei $h \in X$ beliebig. Sei $[x, x + \delta h] \subset K$. Für $t \in [0, \delta]$ gilt dann

$$a + \lambda t h = \lambda(x + th) + (1-\lambda)y \in K\,,$$

also $[a, a + \lambda\delta h] \subset K$. Da h beliebig war, folgt $a \in \operatorname{aint}(K)$. □

Aus Satz 2.13 folgt unmittelbar, dass $\operatorname{aint}(K)$ konvex ist, falls K konvex ist.

3 Algebraische Trennung

Definition 3.1 (Sublineares Funktional)
Sei X Vektorraum. Eine Abbildung $p : X \to \mathbb{R}$ heißt sublinear, falls gilt

$$p(tx) = tp(x)\,, \quad \text{für alle } x \in X,\ t \geq 0, \tag{3.1}$$
$$p(x+y) \leq p(x) + p(y)\,, \quad \text{für alle } x, y \in X. \tag{3.2}$$

□

Jede Halbnorm und jede lineare Abbildung $p : X \to \mathbb{R}$ ist sublinear. Jede sublineare Abbildung ist konvex.

Definition 3.2 (Minkowski-Funktional)
Sei X Vektorraum, sei $S \subset X$. Dann wird durch

$$p_S(x) = \inf\{\alpha : \alpha > 0,\ \frac{1}{\alpha}x \in S\} \tag{3.3}$$

eine Abbildung $p_S : X \to [0, \infty]$ definiert, sie heißt das Minkowski-Funktional oder Eichfunktional von S. Falls $p_S(x) < \infty$ gilt für alle $x \in X$, so heißt S absorbierend. □

Ist K die Einheitskugel in einem normierten Raum X, so ist $p_K(x) = \|x\|$.

Lemma 3.3 *Sei X Vektorraum, $K \subset X$ konvex und absorbierend. Dann gilt $0 \in K$, und das Minkowski-Funktional p_K ist sublinear.*

Beweis: Da K absorbierend ist, ist für $x \in X$ auch $tx \in K$ und $-sx$ in K für geeignete $t, s > 0$, also $0 \in K$ und damit $p(0) = 0$. Für $t > 0$ folgt die Eigenschaft (3.1) unmittelbar aus (3.3). Seien nun $x, y \in X$, sei $\varepsilon > 0$. Wir wählen $\alpha, \beta > 0$ mit

$$\alpha \leq p_K(x) + \varepsilon\,, \quad \frac{1}{\alpha}x \in K\,, \quad \beta \leq p_K(y) + \varepsilon\,, \quad \frac{1}{\beta}y \in K\,.$$

Dann ist, da K konvex ist,

$$\frac{1}{\alpha+\beta}(x+y) = \frac{\alpha}{\alpha+\beta} \cdot \frac{1}{\alpha}x + \frac{\beta}{\alpha+\beta} \cdot \frac{1}{\beta}y \in K\,,$$

also $p_K(x+y) \leq \alpha+\beta \leq p_K(x)+p_K(y)+2\varepsilon$. Da $\varepsilon > 0$ beliebig war, folgt die Behauptung.
□

Lemma 3.4 *Sei X Vektorraum, $K \subset X$ konvex. Dann ist K absorbierend genau dann, wenn $0 \in \operatorname{aint}(K)$.*

Beweis: Folgt direkt aus den Definitionen und Lemma 3.3. □

Satz 3.5 *Sei X Vektorraum, $K \subset X$ konvex, $0 \in \operatorname{aint}(K)$. Dann gilt*

$$\{x : p_K(x) < 1\} = \operatorname{aint}(K) \subset K \subset \{x : p_K(x) \leq 1\}\,. \tag{3.4}$$

Beweis: Ist $x \in K$, so ist $p_K(x) \le 1$ nach Definition von p_K. Ist $p_K(x) < 1$, so ist $tx \in K$ für ein $t > 1$ und damit $x \in \operatorname{aint}(K)$ nach Lemma 2.13, da $x \in [0, tx)$. Ist $x \in \operatorname{aint}(K)$, so ist $tx \in K$ für ein $t > 1$ und damit $p_K(x) < 1$ nach Definition. □

Umgekehrt lässt sich zeigen, dass jedes sublineare Funktional p eine konvexe Menge $K = \{x : p(x) < 1\}$ definiert mit $p_K = p$.

Satz 3.6 *Sei X Vektorraum, $p : X \to \mathbb{R}$ sublinear. Sei U ein Unterraum von X und $\ell : U \to \mathbb{R}$ linear mit $\ell(x) \le p(x)$ für alle $x \in U$. Dann gibt es eine lineare Fortsetzung $L : X \to \mathbb{R}$ von ℓ auf X mit $L(x) \le p(x)$ für alle $x \in X$.*

Beweis: Wir betrachten zunächst den Spezialfall

$$X = \operatorname{span}(U \cup \{y\}), \quad y \in X \setminus U. \tag{3.5}$$

Jedes $x \in X$ läßt sich eindeutig zerlegen in

$$x = z + \alpha y, \quad z \in U, \alpha \in \mathbb{R}. \tag{3.6}$$

Wir definieren $L : X \to \mathbb{R}$ durch

$$L(x) = \ell(z) + \alpha r, \quad \text{falls } x = z + \alpha y, \tag{3.7}$$

wobei $r \in \mathbb{R}$ später festgelegt wird. L ist linear, $L(y) = r$ und $L|U = \ell$. Die verlangte Ungleichung

$$L(z) + \alpha r \le p(z + \alpha y), \quad \text{für alle } z \in U,\ \alpha \in \mathbb{R}, \tag{3.8}$$

ist für $\alpha = 0$ nach Voraussetzung erfüllt, für $\alpha > 0$ gleichbedeutend mit

$$r \le \frac{p(z + \alpha y) - L(z)}{\alpha} = p\left(\frac{z}{\alpha} + y\right) - L\left(\frac{z}{\alpha}\right), \tag{3.9}$$

und für $\alpha < 0$ gleichbedeutend mit

$$r \ge \frac{p(z + \alpha y) - L(z)}{\alpha} = -p\left(-\frac{z}{\alpha} - y\right) + L\left(-\frac{z}{\alpha}\right), \tag{3.10}$$

Ein solches r existiert jedenfalls dann, wenn

$$\sup_{z \in U}(L(z) - p(z - y)) \le \inf_{z \in U}(p(z + y) - L(z)) \tag{3.11}$$

gilt. Nun gilt aber für beliebige $z, \tilde{z} \in U$

$$L(z) + L(\tilde{z}) = L(z + \tilde{z}) \le p(z + \tilde{z}) \le p(z - y) + p(\tilde{z} + y),$$

also auch

$$L(z) - p(z - y) \le p(\tilde{z} + y) - L(\tilde{z}), \quad \text{für alle } z, \tilde{z} \in U,$$

woraus (3.11) folgt. Damit ist der Satz im Spezialfall (3.5) bewiesen. Zum Beweis des allgemeinen Falles verwenden wir das Zornsche Lemma. Wir definieren die Menge

$$\mathcal{M} = \{(V, g) : V \text{ Unterraum},\ U \subset V \subset X,\ g : V \to \mathbb{R} \text{ linear},\ g|U = \ell,\ g \le p \text{ auf } V\}, \tag{3.12}$$

und versehen $\mathcal{M}$ mit der Halbordnung

$$(V_1, g_1) \le (V_2, g_2) \qquad \Leftrightarrow \qquad V_1 \subset V_2\,,\; g_2|V_1 = g_1\,.$$

Es ist $(U, \ell) \in \mathcal{M}$, also $\mathcal{M} \neq \emptyset$. Sei $\mathcal{N}$ eine vollständig geordnete Teilmenge von $\mathcal{M}$. Wir definieren

$$V_* = \bigcup_{(V,g)\in\mathcal{N}} V \tag{3.13}$$

und $g_* : V_* \to \mathbb{R}$ durch

$$g_*(x) = g(x)\,, \quad \text{falls } x \in V,\ (V, g) \in \mathcal{N}. \tag{3.14}$$

Aus der Definition von $\mathcal{N}$ folgt nun, dass $g_*(x)$ nicht von der Wahl von (V, g) abhängt, dass V_* ein Unterraum und g_* linear ist (Details hier nicht ausgeführt). Also ist (V_*, g_*) eine obere Schranke von $\mathcal{N}$ in $\mathcal{M}$. Nach dem Zornschen Lemma hat $\mathcal{M}$ ein maximales Element (V, g). Es muss $V = X$ gelten, da wir andernfalls nach dem schon bewiesenen Spezialfall ein $(\tilde{V}, \tilde{g}) \in \mathcal{M}$ konstruieren könnten mit $\tilde{V} = \operatorname{span}(V \cup \{y\})$, $y \in X \setminus V$, im Widerspruch zur Maximalität von (V, g). $\square$

Ist $\ell : X \to \mathbb{R}$ ein lineares Funktional mit $\ell \neq 0$, so heißt eine Niveaumenge der Form

$$H_\alpha = \{x : x \in X,\ \ell(x) = \alpha\}\,, \quad \alpha \in \mathbb{R}\,, \tag{3.15}$$

Hyperebene. Jede solche Hyperebene trennt X in zwei algebraisch offene Halbräume

$$\{x : x \in X,\ \ell(x) < \alpha\} \quad \text{und} \quad \{x : x \in X,\ \ell(x) > \alpha\}\,.$$

Satz 3.7 *Sei X Vektorraum, $K \subset X$ konvex mit* $\operatorname{aint}(K) \neq \emptyset$*, sei $y \in X$ mit $y \notin$* $\operatorname{aint}(K)$*. Dann gibt es ein lineares Funktional $\ell : X \to \mathbb{R}$ mit*

$$\ell(x) \le \ell(y)\,, \quad \textit{für alle } x \in K, \tag{3.16}$$

und $\ell(x) < \ell(y)$, falls $x \in$ $\operatorname{aint}(K)$*.*

Die konvexe Menge K liegt also "auf einer Seite" von der durch ℓ definierten Hyperebene H_α mit $\alpha = \ell(y)$.

Beweis: Sei o.B.d.A. $0 \in \operatorname{aint}(K)$, andernfalls betrachten wir $K - a$ und $y - a$ mit $a \in \operatorname{aint}(K)$. Auf dem eindimensionalen Unterraum $U = \operatorname{span}\{y\}$ definieren wir ℓ durch $\ell(ty) = t$. Dann ist $\ell(y) = 1 \le p_K(y)$ nach Satz 3.5. Nach Satz 3.6 können wir ℓ zu einem linearen Funktional $\ell : X \to \mathbb{R}$ fortsetzen mit $\ell \le p_K$, also $\ell(x) < 1$ für $x \in \operatorname{aint}(K)$ und $\ell(x) \le 1$ für $x \in K$, wieder nach Satz 3.5. $\square$

Satz 3.8 (Trennung zweier konvexer Mengen)
Sei X Vektorraum, seien $K_1, K_2 \subset X$ konvex und nichtleer, sei $\operatorname{aint}(K_1) \neq \emptyset$*, es gelte* $\operatorname{aint}(K_1) \cap K_2 = \emptyset$*. Dann gibt es ein lineares Funktional $\ell : X \to \mathbb{R}$ und ein $\alpha \in \mathbb{R}$ mit*

$$\ell(x_1) \le \alpha \le \ell(x_2)\,, \quad \textit{für alle } x_1 \in K_1,\ x_2 \in K_2, \tag{3.17}$$

und darüber hinaus gilt $\ell(x_1) < \alpha$ für alle $x_1 \in$ $\operatorname{aint}(K_1)$*. Insbesondere ist $\ell \neq 0$.*

Beweis: Wir setzen

$$K = \operatorname{aint}(K_1) - K_2 = \{x_1 - x_2 : x_1 \in \operatorname{aint}(K_1),\ x_2 \in K_2\}\,.$$

Dann ist K konvex nach Lemma 2.9, da $\operatorname{aint}(K_1)$ konvex ist. Wegen

$$K = \bigcup_{z \in K_2} (\operatorname{aint}(K_1) - z)$$

ist K nichtleer und algebraisch offen, und es gilt $0 \notin K$. Wir wählen gemäß Satz 3.7 ein lineares $\ell : X \to \mathbb{R}$ mit $\ell(x) < \ell(0) = 0$ für alle $x \in K = \operatorname{aint}(K)$. Es folgt

$$\ell(x_1) - \ell(x_2) = \ell(x_1 - x_2) < 0\,, \quad \text{für alle } x_1 \in \operatorname{aint}(K_1),\ x_2 \in K_2.$$

Wir setzen $\alpha = \sup\ \ell(\operatorname{aint}(K_1))$. Es folgt $K_2 \subset \{x : \ell(x) \geq \alpha\}$ sowie

$$\operatorname{aint}(K_1) \subset \operatorname{aint}(\{x : \ell(x) \leq \alpha\}) = \{x : \ell(x) < \alpha\}\,.$$

(Siehe Übung für die Gleichung.) Da nach Satz 2.13 jedes $x \in K_1$ Endpunkt einer Strecke $[a, x]$ mit $[a, x) \subset \operatorname{aint}(K_1)$ ist, folgt $\ell(x) \leq \alpha$ für $x \in K_1$ durch Grenzübergang längs dieser Strecke. □

Im Endlichdimensionalen gilt der Trennungssatz auch ohne Voraussetzungen an das algebraische Innere. Sind K_1, K_2 disjunkte nichtleere konvexe Teilmengen des $\mathbb{R}^n$, so gibt es ein lineares Funktional $\ell : \mathbb{R}^n \to \mathbb{R}$ mit $\ell \neq 0$, so dass (3.17) gilt.

In einem unendlichdimensionalen Vektorraum X ist es immer möglich, zu zwei disjunkten konvexen Teilmengen K_1, K_2 von X eine Zerlegung $X = A \cup B$ in zwei disjunkte konvexe Teilmengen A und B von X mit $K_1 \subset A$, $K_2 \subset B$ zu finden (Satz von Stone). Diese müssen aber nicht die Form von Halbräumen haben, wenn K_1 und K_2 beide ein leeres algebraisches Inneres haben. (Siehe Holmes: Geometric Functional Analysis.)

4 Lokalkonvexe Räume, Trennungssatz

Im Unendlichdimensionalen ist der Trennungssatz primär dann nützlich, wenn das trennende lineare Funktional auch stetig ist, da sich die Menge aller linearen stetigen Funktionale (im Gegensatz zur Menge aller linearen Funktionale) in vielen Fällen gut beschreiben lässt. Der Begriff des lokalkonvexen Raumes liefert den angemessenen Rahmen für den Trennungssatz.

Sei $(X, \|\cdot\|)$ ein normierter Raum. Dann gelten bekanntlich die folgenden Aussagen:

(i) Die Addition $+ : X \times X \to X$ und die Skalarmultiplikation $\cdot : \mathbb{R} \times X \to X$ sind stetig.

(ii) Je zwei verschiedene Punkte $x, y \in X$ lassen sich durch disjunkte Umgebungen trennen.

(iii) Jede Nullumgebung V enthält eine konvexe Nullumgebung.

Die Menge $V \subset X$ heißt **Nullumgebung**, wenn V offen ist und $0 \in V$.

Definition 4.1 (Lokalkonvexer Raum) *Sei X ein mit einer Topologie τ versehener Vektorraum, in dem (i) – (iii) gelten. Dann heißt X lokalkonvex.*

Eine Topologie τ auf X ist ein System von Teilmengen von X, welche "offene Mengen" genannt werden. Es wird verlangt, dass $\emptyset$ und X offen sind, und dass endliche Durchschnitte und beliebige Vereinigung von offenen Mengen wieder offen sind. Komplemente von offenen Mengen in X heißen abgeschlossen. Eine Funktion zwischen topologischen Räumen ist definitionsgemäß stetig, wenn das Urbild jeder offenen Menge ebenfalls offen ist. Da in $\mathbb{R}$ jede offene Menge sich als Vereinigung (sogar abzählbare) von offenen Intervallen schreiben lässt, ist eine Funktion $f : X \to \mathbb{R}$ stetig genau dann, wenn $f^{-1}(I)$ offen ist für jedes offene Intervall $I = (a, b)$.

Die Stetigkeit von Addition und Skalarmultiplikation hat zur Folge, dass für alle $a \in X$ und alle $\lambda \in \mathbb{R}$, $\lambda \neq 0$, die Translationen $x \mapsto x + a$ und die Streckungen $x \mapsto \lambda x$ bijektive und in beide Richtungen stetige Abbildungen sind. Die Topologie τ ist damit bereits festgelegt durch die offenen Mengen, welche 0 enthalten, also durch die Nullumgebungen.

Lemma 4.2 *Sei X lokalkonvex. Jede Nullumgebung enthält eine absolutkonvexe Nullumgebung. Eine Teilmenge V von X heißt absolutkonvex, wenn V konvex ist und $tV \subset V$ gilt für alle $|t| \leq 1$.*

Beweis: Ist W eine konvexe Nullumgebung, so ist $V = W \cap (-W)$ eine absolutkonvexe Nullumgebung mit $V \subset W$. $\square$

Ist $S \subset X$ eine beliebige Teilmenge, so ist also $a \in \operatorname{int}(S)$ genau dann, wenn es eine absolutkonvexe Nullumgebung V gibt mit $a + V \subset S$.

Satz 4.3 *Sei X lokalkonvex, $K \subset X$ konvex. Ist $x \in \operatorname{int}(K)$ und $y \in K$, so ist $[x, y) \subset \operatorname{int}(K)$.*

Beweis: Sei $a \in [x, y)$, $a = \lambda x + (1-\lambda)y$, $0 < \lambda \leq 1$. Sei V Nullumgebung von x mit $x + V \subset K$, dann gilt

$$K \supset \lambda(x+V) + (1-\lambda)y = a + \lambda V \ni a\,,$$

also $a \in \mathrm{int}\,(K)$, da V eine Nullumgebung ist. □

Satz 4.4 *Sei X lokalkonvex, $K \subset X$ konvex, es gelte* $\mathrm{int}\,(K) \neq 0$. *Dann gilt*

$$\mathrm{int}\,(K) = \mathrm{aint}\,(K)\,. \tag{4.1}$$

Insbesondere ist jede konvexe Nullumgebung absorbierend.

Beweis: "$\subset$": Sei $x \in \mathrm{int}\,(K)$, $h \in X$. Da $t \mapsto x + th$ stetig und $\mathrm{int}\,(K)$ offen ist, muss $[x, x + \delta h] \subset \mathrm{int}\,(K)$ gelten, falls δ hinreichend klein ist.
"$\supset$": Sei $x \in \mathrm{aint}\,(K)$. Wir wählen $z \in \mathrm{int}\,(K)$ beliebig sowie $\delta > 0$ so, dass $y = x - \delta(z - x) \in K$ gilt. Dann ist $x \in [z, y)$ und damit $x \in \mathrm{int}\,(K)$ nach Satz 4.3. □

Definition 4.5 (Dualraum) *Sei X lokalkonvex. Die Menge*

$$X^* = \{x^* |\; x^* : X \to \mathbb{R} \text{ linear und stetig}\} \tag{4.2}$$

heißt der Dualraum von X. □

Satz 4.6 *Sei X lokalkonvex, $\ell : X \to \mathbb{R}$ linear, $O \subset X$ offen und nichtleer, es gelte $\ell(O) \geq 0$. Dann ist ℓ stetig.*

Beweis: Wir betrachten $A_\alpha = \{x : x \in X,\ \ell(x) > \alpha\}$. Es ist $\mathrm{aint}\,(A_\alpha) = A_\alpha$, da $\ell(x+th) = \ell(x) + t\ell(h) > \alpha$ für $x \in A_\alpha$, falls t hinreichend klein ist. A_α ist konvex, und $x + O \subset A_\alpha$ falls $x \in A_\alpha$. Aus Satz 4.4 folgt $\mathrm{int}\,(A_\alpha) = \mathrm{aint}\,(A_\alpha)$, also ist A_α offen. Da $-\ell(-O) = \ell(O) \geq 0$, sind die Mengen $\{x : x \in X,\ \ell(x) < \beta\} = \{x : x \in X,\ -\ell(x) > -\beta\}$ ebenfalls offen. Also ist $\ell^{-1}(I)$ offen für jedes offene Intervall $I = (\alpha, \beta)$ in $\mathbb{R}$ und damit ℓ stetig. □

Folgerung 4.7 *Sei X lokalkonvex, $\ell : X \to \mathbb{R}$ affin linear, $O \subset X$ offen und nichtleer, es gelte $\ell(O) \geq c$ oder $\ell(O) \leq c$ für ein $c \in \mathbb{R}$. Dann ist ℓ stetig.*

Beweis: Sei ℓ nichtkonstant (konstante Funktionen sind stetig). Im Fall $\ell(O) \geq c$ wählen wir $a \in X$ mit $\ell(a) = c$ und setzen $\tilde{\ell}(x) = \ell(x) - \ell(0)$, $\tilde{O} = O - a$. Dann ist $\tilde{\ell}$ linear, $\tilde{O}$ offen und

$$\tilde{\ell}(\tilde{O}) = \tilde{\ell}(O) - \tilde{\ell}(a) = \ell(O) - \ell(a) \geq 0\,.$$

Aus Satz 4.6 folgt, dass $\tilde{\ell}$ und damit auch ℓ stetig ist. Im Fall $\ell(O) \leq c$ gehen wir zu $-\ell$ über. □

Satz 4.8 (Trennungssatz) *Sei X lokalkonvex, seien $K_1, K_2 \subset X$ konvex und nichtleer, es gelte $\mathrm{int}\,(K_1) \cap K_2 = \emptyset$ und $\mathrm{int}\,(K_1) \neq \emptyset$. Dann gibt es ein $x^* \in X^*$ mit $x^* \neq 0$ und ein $\alpha \in \mathbb{R}$ mit*

$$x^*(K_1) \leq \alpha \leq x^*(K_2)\,, \quad x^*(\mathrm{int}\,K_1) < \alpha\,. \tag{4.3}$$

Beweis: Die Existenz eines linearen Funktionals mit den verlangten Eigenschaften folgt aus Satz 3.8, da $\operatorname{aint}(K_1) = \operatorname{int}(K_1)$ nach Satz 4.4, und dessen Stetigkeit aus Satz 4.7. □

Satz 4.9 (Strikte Trennung) *Sei X lokalkonvex, sei $K \subset X$ konvex, nichtleer und abgeschlossen, sei $y \notin K$. Dann gibt es ein $x^* \in X^*$ mit $x^* \neq 0$ und ein $\alpha \in \mathbb{R}$ mit*

$$x^*(x) \leq \alpha < x^*(y)\,, \quad \textit{für alle } x \in K\,. \tag{4.4}$$

Beweis: Da $X \setminus K$ offen ist, gibt es eine konvexe Nullumgebung V mit $K \cap (y+V) = \emptyset$. Nach Satz 4.8 gibt es ein $x^* \in X^*$ und ein α mit $x^*(K) \leq \alpha < x^*(y+V)$. □

5 Konvexe Funktionen

Wir schreiben $[-\infty, \infty]$ für $\mathbb{R} \cup \{-\infty, +\infty\}$ und entsprechend $(-\infty, \infty]$, $[0, \infty]$ etc.

Definition 5.1 (Epigraph) *Sei X Vektorraum, $f : X \to [-\infty, \infty]$. Die durch*

$$\operatorname{epi} f = \{(x, \mu) : x \in X,\, \mu \in \mathbb{R},\, \mu \geq f(x)\} \tag{5.1}$$

definierte Teilmenge von $X \times \mathbb{R}$ heißt der Epigraph von f, die durch

$$\operatorname{dom} f = \{x : x \in X,\, f(x) < +\infty\} \tag{5.2}$$

definierte Teilmenge von X der effektive Definitionsbereich von f.

Bezeichnet $p_X : X \times \mathbb{R} \to X$ die Projektion auf die erste Komponente, so ist

$$\operatorname{dom} f = p_X(\operatorname{epi} f)\,. \tag{5.3}$$

Definition 5.2 (Konvexe Funktion) *Sei X Vektorraum, $f : X \to [-\infty, \infty]$. f heißt konvex, falls der Epigraph von f konvex ist. f heißt konkav, wenn $-f$ konvex ist.*

Das Addieren und Multiplizieren von Funktionen, deren Werte $-\infty$ oder $+\infty$ annehmen können, erfordert Rechenregeln wie etwa

$$x + \infty = \infty + x = \infty\,, \quad x - \infty = -\infty + x = -\infty\,, \quad 0 \cdot \infty = \infty \cdot 0 = 0\,, \tag{5.4}$$

für $x \in \mathbb{R}$; sie entstehen alle durch Grenzübergang, etwa aus

$$\lim_{r \to \infty} x + r = \infty\,. \tag{5.5}$$

Wir definieren $\infty + \infty = \infty$. Nicht definiert und nicht zugelassen ist

$$\infty - \infty\,. \tag{5.6}$$

Satz 5.3 *Sei X Vektorraum, $f : X \to (-\infty, \infty]$. Dann ist f konvex genau dann, wenn*

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda) f(y) \tag{5.7}$$

für alle $x, y \in X$ und alle $\lambda \in [0, 1]$.

Beweis: "$\Rightarrow$": Falls $f(x) = \infty$ oder $f(y) = \infty$, so ist die Aussage trivial. Seien also $x, y \in \operatorname{dom} f$ und $\lambda \in [0, 1]$. Dann sind $(x, f(x))$ und $(y, f(y))$ Elemente von $\operatorname{epi} f$, also auch

$$(\lambda x + (1 - \lambda)y\,,\, \lambda f(x) + (1 - \lambda) f(y)) \in \operatorname{epi} f\,, \tag{5.8}$$

woraus (5.7) folgt.
"$\Leftarrow$": Seien $(x, \mu), (y, \nu) \in \operatorname{epi} f$, dann gilt

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda) f(y) \leq \lambda \mu + (1 - \lambda)\nu\,, \tag{5.9}$$

also ist $\lambda(x, \mu) + (1 - \lambda)(y, \nu) = (\lambda x + (1 - \lambda)y, \lambda \mu + (1 - \lambda)\nu) \in \operatorname{epi} f$. $\square$

Jedes sublineare Funktional $p : X \to \mathbb{R}$ auf einem Vektorraum ist konvex, insbesondere jede Norm und Halbnorm.

Ist $f : X \to (-\infty, \infty]$ konvex, so gilt

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \le \sum_{i=1}^{n} \lambda_i f(x_i) \tag{5.10}$$

für beliebige Konvexkombinationen $x = \sum_i \lambda_i x_i$.

Definition 5.4 *Sei X Vektorraum, $S \subset X$. Eine Funktion $f : S \to [-\infty, \infty]$ heißt konvex, wenn die durch*

$$\tilde{f}(x) = \begin{cases} f(x)\,, & \text{falls } x \in K\,, \\ +\infty\,, & \text{sonst}\,, \end{cases} \tag{5.11}$$

definierte Funktion $\tilde{f} : X \to [-\infty, \infty]$ konvex ist.

Man sieht unmittelbar, daß eine Funktion $f : K \to (-\infty, \infty]$, welche (5.7) auf einer konvexen Menge $K \subset X$ erfüllt, konvex ist im Sinne von Definition 5.4.

Lemma 5.5 *Sei X Vektorraum, $f : X \to [-\infty, \infty]$ konvex. Dann sind die Mengen $\{x : x \in X,\ f(x) \le \alpha\}$ und $\{x : x \in X,\ f(x) < \alpha\}$ konvex für alle $\alpha \in [-\infty, \infty]$.*

Beweis: Folgt unmittelbar aus der Definition. □

Satz 5.6 *Sei $f : (a, b) \to \mathbb{R}$ zweimal stetig differenzierbar. Dann ist f konvex genau dann, wenn $f''(x) \ge 0$ für alle $x \in (a, b)$.*

Beweis: Siehe Analysis. □

Hieraus folgt beispielsweise, daß

$$f(x) = -\log x \tag{5.12}$$

konvex ist auf $(0, \infty)$.

Satz 5.7 *Sei $G \subset \mathbb{R}^n$ offen und konvex, sei $f : G \to \mathbb{R}$ zweimal stetig differenzierbar. Dann ist f konvex genau dann, wenn $D^2 f(x)$ positiv semidefinit ist für alle $x \in G$.*

Beweis: Übungsaufgabe. □

Satz 5.8 *Sei X Vektorraum, seien $f : X \to (-\infty, \infty]$ und $\varphi : \mathbb{R} \to (-\infty, \infty]$ konvex, sei φ monoton wachsend ($s \le t \Rightarrow \varphi(s) \le \varphi(t)$). Dann ist auch $\varphi \circ f$ konvex. (Dabei wird $\varphi(\infty) = \infty$ gesetzt.)*

Beweis: Für alle $x, y \in X$ und $\lambda \in [0, 1]$ gilt

$$f(\lambda x + (1 - \lambda) y) \le \lambda f(x) + (1 - \lambda) f(y)\,, \tag{5.13}$$

also

$$(\varphi \circ f)(\lambda x + (1 - \lambda) y) \le \varphi(\lambda f(x) + (1 - \lambda) f(y)) \le \lambda \varphi(f(x)) + (1 - \lambda) \varphi(f(y))\,. \tag{5.14}$$

□

Definition 5.9 (Eigentliche konvexe Funktion)
Sei X Vektorraum, $f : X \to [-\infty, \infty]$ konvex. f heißt eigentlich, wenn $\operatorname{epi} f$ *nichtleer ist und keine senkrechten Linien enthält, also wenn $f(x) < \infty$ für mindestens ein $x \in X$ und $f(x) > -\infty$ für alle $x \in X$. Ist f nicht eigentlich, so heißt f uneigentlich.*

Lemma 5.10 *Sei X Vektorraum, seien $f_1, f_2 : X \to (-\infty, \infty]$ konvex. Dann ist $f_1 + f_2$ konvex. $f_1 + f_2$ ist eigentlich genau dann, wenn* $\operatorname{dom} f_1 \cap \operatorname{dom} f_2 \neq \emptyset$.

Beweis: Folgt unmittelbar aus Satz 5.3. □

Lemma 5.11 *Sei X Vektorraum, sei $K \subset X \times \mathbb{R}$ konvex. Dann ist die durch*

$$f(x) = \inf_{(x,\mu) \in K} \mu \tag{5.15}$$

definierte Funktion $f : X \to [-\infty, \infty]$ konvex. (Man beachte die Konvention $\inf_K = +\infty$ *falls $K = \emptyset$.)*

Beweis: Übung.

Satz 5.12 *Sei X Vektorraum, seien $f_1, f_2 : X \to (-\infty, \infty]$ konvex. Dann ist die durch*

$$f(x) = \inf_{\substack{x_1, x_2 \in X \\ x_1 + x_2 = x}} (f_1(x_1) + f_2(x_2)) = \inf_{y \in X} (f_1(y) + f_2(x - y)) \tag{5.16}$$

definierte Funktion $f : X \to [-\infty, \infty]$ konvex.

Beweis: Sei $K = \operatorname{epi} f_1 + \operatorname{epi} f_2$. Dann ist K konvex und

$$K = \{(x_1 + x_2 \,,\, \mu_1 + \mu_2) : x_j \in X \,,\, f_j(x_j) \leq \mu_j \text{ für } j = 1, 2\} \,. \tag{5.17}$$

Sei nun $x \in X$ fest gewählt. Dann gilt

$$\inf_{(x,\mu) \in K} \mu = \inf_{\substack{x_1 + x_2 = x \\ f_j(x_j) \leq \mu_j}} (\mu_1 + \mu_2) = \inf_{x_1 + x_2 = x} (f_1(x_1) + f_2(x_2)) = f(x) \,. \tag{5.18}$$

Die Behauptung folgt nun aus Lemma 5.11. □

Definition 5.13 (Infimale Faltung) *Die in Satz 5.12 definierte Funktion f heißt infimale Faltung von f_1 und f_2, geschrieben*

$$f = f_1 \# f_2 \,. \tag{5.19}$$

Es gilt nach (5.16)

$$(f_1 \# f_2)(x) = \inf_{y \in X} (f_1(x - y) + f_2(y)) \,.$$

Falls das Minimum immer dann angenommen wird, wenn das Infimum endlich ist, gilt $\operatorname{epi}(f_1 \# f_2) = \operatorname{epi}(f_1) + \operatorname{epi}(f_2)$.

Definition 5.14 (Indikatorfunktion) *Sei X Vektorraum, $K \subset X$. Die durch*

$$I_K(x) = \begin{cases} 0\,, & x \in K \\ +\infty\,, & \textit{sonst} \end{cases} \tag{5.20}$$

definierte Funktion $I_K : X \to [0, \infty)$ heißt Indikatorfunktion von K.

Wegen $\operatorname{epi} I_K = K \times \mathbb{R}_+$ ist I_K genau dann konvex, wenn K konvex ist. Für eine einpunktige Menge $K = \{a\}$ gilt

$$(f \# I_{\{a\}})(x) = f(x - a)\,. \tag{5.21}$$

Ist $(X, \|\cdot\|)$ normiert und $K \subset X$ konvex, und wählen wir $f = \|\cdot\|$, so gilt

$$(f \# I_K)(x) = \inf_{y \in K} \|x - y\| =: d(x, K)\,, \tag{5.22}$$

und nach Satz 5.12 ist also $x \mapsto d(x, K)$ konvex. (Letzteres kann man einfacher direkt beweisen.)

Lemma 5.15 *Sei X Vektorraum, seien $f_i : X \to [-\infty, \infty]$ konvex, $i \in I$. Dann ist*

$$f = \sup_{i \in I} f_i \tag{5.23}$$

konvex.

Beweis: Es ist $f(x) \le \mu$ genau dann, wenn $f_i(x) \le \mu$ für alle $i \in I$, also folgt

$$\operatorname{epi} f = \bigcap_{i \in I} \operatorname{epi} f_i\,,$$

also ist $\operatorname{epi} f$ konvex. □

Definition 5.16 (Unterhalbstetigkeit) *Sei X lokalkonvex, $f : X \to [-\infty, \infty]$. f heißt unterhalbstetig, wenn die Subniveaumengen*

$$M_\alpha = \{x : x \in X,\ f(x) \le \alpha\} \tag{5.24}$$

abgeschlossen sind für alle $\alpha \in \mathbb{R}$.

Satz 5.17 *Sei X lokalkonvex, $f : X \to [-\infty, \infty]$. Dann ist f unterhalbstetig genau dann, wenn $\operatorname{epi} f$ abgeschlossen ist in $X \times \mathbb{R}$.*

Beweis: "$\Leftarrow$": Sei $\alpha \in \mathbb{R}$. Dann ist

$$F_\alpha = \{(x, \alpha) : x \in X,\ f(x) \le \alpha\} = \operatorname{epi} f \cap (X \times \{\alpha\}) \tag{5.25}$$

abgeschlossen in $X \times \mathbb{R}$, also auch die Subniveaumenge $M_\alpha = j_\alpha^{-1}(F_\alpha)$, wobei $j_\alpha : X \to X \times \mathbb{R}$ die Einbettung $j_\alpha(x) = (x, \alpha)$ bezeichnet.
"$\Rightarrow$": Wir zeigen, daß das Komplement von $\operatorname{epi} f$ offen ist. Sei $(x, \alpha) \notin \operatorname{epi} f$, also $f(x) > \alpha$. Wähle $\varepsilon > 0$ mit $f(x) > \alpha + \varepsilon$. Nach Voraussetzung ist $U = \{y : f(y) > \alpha + \varepsilon\}$ offen in X und $V = U \times (\alpha - \varepsilon, \alpha + \varepsilon)$ eine offene Umgebung von (x, α) mit $V \cap \operatorname{epi} f = \emptyset$, da $f(y) > \alpha + \varepsilon > \beta$ für alle $(y, \beta) \in V$. □

Folgerung 5.18 *Sei X lokalkonvex, $K \subset X$. Dann ist K abgeschlossen genau dann, wenn I_K unterhalbstetig ist.*

Beweis: "$\Rightarrow$": $\operatorname{epi} I_K = K \times [0, \infty)$.
"$\Leftarrow$": $K = \{x : I_K(x) \leq 0\}$. $\square$

Lemma 5.19 *Sei X lokalkonvex. Sind $f, g : X \to [-\infty, \infty]$ unterhalbstetig, so ist auch $\min\{f, g\}$ unterhalbstetig. Sind $f_i : X \to [-\infty, \infty]$ unterhalbstetig für alle $i \in I$, so ist auch $\sup_{i \in I} f_i$ unterhalbstetig.*

Beweis:

$$\operatorname{epi}(\min\{f, g\}) = \operatorname{epi} f \cup \operatorname{epi} g\,, \quad \operatorname{epi}(\sup_{i \in I} f_i) = \bigcap_{i \in I} \operatorname{epi} f_i\,. \tag{5.26}$$

$\square$

Satz 5.20 *Sei X lokalkonvex, $f : X \to (-\infty, \infty]$ konvex, $x \in X$ mit $f(x) \in \mathbb{R}$. Dann ist f stetig in x genau dann, wenn es eine Umgebung U von x gibt, auf der f nach oben beschränkt ist, d.h. es gibt ein $c \in \mathbb{R}$ mit*

$$f(y) \leq c\,, \quad \textit{für alle } y \in U\,. \tag{5.27}$$

Beweis: "$\Rightarrow$": Folgt direkt aus der Definition der Stetigkeit; wähle eine Umgebung U von x mit $f(U) \subset (f(x) - 1, f(x) + 1)$.
"$\Leftarrow$": Sei o.B.d.A. $x = 0$, $f(0) = 0$, sei $c \in \mathbb{R}$ und U Nullumgebung mit $f(U) \leq c$. Sei $V \subset U$ eine absolutkonvexe Nullumgebung. Es genügt zu zeigen, dass $f(\varepsilon V) \subset [-\varepsilon c, \varepsilon c]$ gilt für jedes $\varepsilon \in (0, 1)$. Sei $y \in \varepsilon V$ beliebig. Dann gilt

$$y = (1 - \varepsilon) \cdot 0 + \varepsilon \frac{y}{\varepsilon}\,, \quad f(y) \leq \varepsilon f\left(\frac{y}{\varepsilon}\right) \leq \varepsilon c\,, \tag{5.28}$$

und außerdem, da $-\varepsilon V = \varepsilon V$,

$$0 = \frac{1}{1 + \varepsilon} y + \frac{\varepsilon}{1 + \varepsilon}\left(-\frac{y}{\varepsilon}\right)\,, \tag{5.29}$$

$$0 \leq \frac{1}{1 + \varepsilon} f(y) + \frac{\varepsilon}{1 + \varepsilon} f\left(-\frac{y}{\varepsilon}\right) \leq \frac{1}{1 + \varepsilon}(f(y) + \varepsilon c)\,, \tag{5.30}$$

also

$$|f(y)| \leq \varepsilon c\,. \tag{5.31}$$

$\square$

Bemerkung: Man kann außerdem zeigen: Ist $f : X \to [-\infty, \infty]$ konvex, $f(x) \in \mathbb{R}$ und f auf einer Umgebung von x nach oben beschränkt, so kann f den Wert $-\infty$ nicht annehmen.

Satz 5.21 *Sei $f : \mathbb{R}^n \to (-\infty, \infty]$ konvex. Dann ist f stetig auf* $\operatorname{int}(\operatorname{dom} f)$.

Beweis: Ist $x = \sum_{i=0}^n \lambda_i x_i$ eine Konvexkombination, so ist

$$f(x) \leq \sum_{i=0}^n \lambda_i f(x_i) \leq \left(\sum_{i=0}^n \lambda_i\right) \max_{0 \leq i \leq n} f(x_i) = \max_{0 \leq i \leq n} f(x_i)\,. \tag{5.32}$$

Sei nun $x \in \operatorname{int}(\operatorname{dom} f)$. Wähle ein n-dimensionales Simplex $S = \operatorname{co}\{x_0, \dots, x_n\}$ mit $x \in \operatorname{int}(S) \subset \operatorname{int}(\operatorname{dom} f)$, dann ist

$$f(S) \le \max_{0 \le i \le n} f(x_i) < \infty\,, \tag{5.33}$$

also ist f stetig nach Lemma 5.20. □

In der Funktionalanalysis zeigt man den folgenden Satz, als Konsequenz des Baireschen Kategoriensatzes.

Satz 5.22 *Sei X Banachraum, $K \subset X$ abgeschlossen, absolutkonvex und absorbierend. Dann ist* $0 \in \operatorname{int}(K)$. □

Satz 5.23 *Sei X Banachraum, $f : X \to (-\infty, \infty]$ konvex und unterhalbstetig. Dann ist f stetig auf* $\operatorname{int}(\operatorname{dom} f)$.

Beweis: Sei $0 \in \operatorname{int}(\operatorname{dom} f)$, sei $f(0) < c$. Wir definieren

$$U = \{x : x \in X,\ f(x) \le c\}\,, \quad V = U \cap (-U)\,. \tag{5.34}$$

Da f konvex und unterhalbstetig ist, ist U konvex und abgeschlossen, und da außerdem $0 \in U$ gilt, ist V absolutkonvex und abgeschlossen. Um zu zeigen, daß V absorbierend ist, wählen wir $x \in X$ beliebig und definieren $g : \mathbb{R} \to (-\infty, \infty]$ durch $g(t) = f(tx)$. Nach Satz 5.21 ist g stetig in 0, also gibt es ein $\varepsilon > 0$ mit $[-\varepsilon x, \varepsilon x] \subset U$, also auch $[-\varepsilon x, \varepsilon x] \subset V$. Damit ist $0 \in \operatorname{aint}(V)$ und V absorbierend. Nach Satz 5.22 ist $0 \in \operatorname{int}(V)$, nach Satz 5.20 ist f stetig in 0. Für beliebiges $y \in \operatorname{int}(\operatorname{dom} f)$ betrachten wir $\tilde{f}(x) = f(x + y)$. □

Satz 5.24 *Sei X lokalkonvex, $f : X \to (-\infty, \infty]$ konvex, unterhalbstetig und eigentlich. Dann gilt*

$$f = \sup\{g|\ g : X \to \mathbb{R} \text{ affin und stetig},\ g \le f\}\,. \tag{5.35}$$

Beweis: "$\ge$": klar.
"$\le$": Es genügt zu zeigen: Ist $(x, a) \in X \times \mathbb{R}$ mit $a < f(x)$, so gibt es eine affine stetige Funktion $g : X \to \mathbb{R}$ mit $a \le g(x)$ und $g \le f$. Sei also $a < f(x)$, dann ist $(x, a) \notin \operatorname{epi} f$, und $\operatorname{epi} f$ ist konvex und abgeschlossen. Aus dem Trennungssatz folgt, daß ein $z^* \in (X \times \mathbb{R})^*$ existiert mit

$$z^*(x, a) < \inf_{(\xi,\alpha) \in \operatorname{epi} f} z^*(\xi, \alpha)\,. \tag{5.36}$$

Wir setzen $\lambda = z^*(0, 1)$ und definieren $x^* \in X^*$ durch $x^*(\xi) = z^*(\xi, 0)$, dann gilt

$$z^*(\xi, \alpha) = x^*(\xi) + \lambda\alpha\,, \quad \forall\, \xi \in X\,,\ \alpha \in \mathbb{R}\,. \tag{5.37}$$

Aus (5.36) erhalten wir also

$$x^*(x) + \lambda a = z^*(x, a) < x^*(\xi) + \lambda\mu\,, \quad \forall\, (\xi, \mu) \in \operatorname{epi} f\,. \tag{5.38}$$

Hieraus folgt $\lambda \ge 0$, da μ beliebig groß gewählt werden kann.
Fall 1: $\lambda > 0$. Wir definieren

$$g(\xi) = \frac{1}{\lambda}(z^*(x, a) - x^*(\xi))\,. \tag{5.39}$$

Aus (5.38) folgt $g(x) = a$ und $g(\xi) < \mu$ für $(\xi, \mu) \in \operatorname{epi} f$, also $g(\xi) < f(\xi)$ für $\xi \in \operatorname{dom} f$ und damit $g \le f$.
Fall 2: $\lambda = 0$. Dann ist $\xi = x$ in (5.38) nicht möglich, also $x \notin \operatorname{dom} f$ und $f(x) = \infty$. Sei $\tilde{g} : X \to \mathbb{R}$ affin und stetig mit $\tilde{g} \le f$; so ein $\tilde{g}$ existiert immer, da $\operatorname{dom} f$ nichtleer ist und bei Elementen von $\operatorname{dom} f$ der Fall 1 zum Tragen kommt. Wir wählen ein $\beta \in \mathbb{R}$ mit

$$x^*(x) = z^*(x, a) < \beta < z^*(\operatorname{epi} f) = x^*(\operatorname{dom} f)\,, \tag{5.40}$$

und setzen

$$g(\xi) = \tilde{g}(\xi) + \delta(\beta - x^*(\xi))\,, \quad \delta > 0\,. \tag{5.41}$$

Aus (5.40) folgt $g(\xi) \le \tilde{g}(\xi) \le f(\xi)$ für $\xi \in \operatorname{dom} f$ und

$$g(x) = \tilde{g}(x) + \delta(\beta - x^*(x)) \ge a\,, \tag{5.42}$$

falls $\delta > 0$ hinreichend groß gewählt wird. □

6 Konjugierte Funktionen

Definition 6.1 (Konjugierte Funktion)
Sei X Banachraum, $f : X \to [-\infty, \infty]$. Die durch

$$f^*(x^*) = \sup_{x \in X} (\langle x^*, x \rangle - f(x)) \tag{6.1}$$

definierte Funktion $f^* : X^* \to [-\infty, \infty]$ heißt die zu f konjugierte Funktion. (Wir schreiben $\langle x^*, x \rangle$ für $x^*(x)$.)

Beispiel 6.2 (i) Sei X Banachraum, $f : X \to \mathbb{R}$ nach oben beschränkt, etwa $f(x) \leq c$ für alle $x \in X$. Dann ist

$$f^*(x^*) = \sup_{x \in X} (\langle x^*, x \rangle - f(x)) \geq \sup_{x \in X} \langle x^*, x \rangle - c\,, \tag{6.2}$$

also

$$f^*(x^*) = \begin{cases} -\inf_{x \in X} f(x)\,, & x^* = 0\,, \\ \infty\,, & \text{sonst.} \end{cases} \tag{6.3}$$

(ii) Sei $K \subset X$, $f = I_K$. Dann ist

$$f^*(x^*) = \sup_{x \in X} (\langle x^*, x \rangle - I_K(x)) = \sup_{x \in K} \langle x^*, x \rangle \ . \tag{6.4}$$

Im Spezialfall $K =$ Einheitskugel gilt also

$$I_K^*(x^*) = \|x^*\|_{X^*}\,. \tag{6.5}$$

(iii) Sei X Banachraum, $f(x) = \|x\|$ für $x \in X$. Dann gilt

$$\begin{aligned} f^*(x^*) &= \sup_{x \in X} (\langle x^*, x \rangle - \|x\|) = \sup_{\lambda \geq 0} \lambda \sup_{\|x\|=1} (\langle x^*, x \rangle - \|x\|) \\ &= \sup_{\lambda \geq 0} \lambda (\|x^*\|_{X^*} - 1) \\ &= \begin{cases} 0\,, & \|x^*\| \leq 1\,, \\ +\infty\,, & \|x^*\| > 1\,, \end{cases} \end{aligned} \tag{6.6}$$

also

$$f^* = I_{K^*}\,, \tag{6.7}$$

wobei K^* die Einheitskugel in X^* ist.

Die Konjugation hat folgende elementare Eigenschaften (Beweis klar bzw. Übung)

$$f^*(0) = \sup_{x \in X} -f(x) = -\inf_{x \in X} f(x)\,, \tag{6.8}$$

$$f \leq g \quad \Rightarrow \quad f^* \geq g^*\,, \tag{6.9}$$

$$(\lambda f)^*(x^*) = \lambda f^* \left(\frac{x^*}{\lambda} \right)\,, \quad \lambda > 0\,, \tag{6.10}$$

$$(f + c)^* = f^* - c\,, \quad c \in \mathbb{R}\,, \tag{6.11}$$

$$\left(\inf_{i \in I} f_i \right)^* = \sup_{i \in I} f_i^*\,. \tag{6.12}$$

Definition 6.3 (Bikonjugierte Funktion)
Sei X Banachraum, $f : X \to [-\infty, \infty]$. Die durch

$$f^{**}(x) = \sup_{x^* \in X^*} (\langle x^*, x\rangle - f^*(x^*)) \tag{6.13}$$

definierte Funktion $f^{**} : X \to [-\infty, \infty]$ heißt die zu f bikonjugierte Funktion.

Gibt es ein $x \in X$ mit $f(x) = -\infty$, so ist $f^* \equiv +\infty$ und $f^{**} \equiv -\infty$.
Gibt es ein $x \in X$ mit $f(x) < \infty$, so ist $f^*(x^*) > -\infty$ für alle $x^* \in X^*$.

Lemma 6.4 *Sei X Banachraum, $f : X \to [-\infty, \infty]$. Dann sind f^* und f^{**} konvex und unterhalbstetig.*

Beweis: Es genügt, den Fall $f : X \to (-\infty, \infty]$ mit $\operatorname{dom} f \neq \emptyset$ zu betrachten, die anderen sind trivial. In diesem Fall gilt

$$f^*(x^*) = \sup_{x \in \operatorname{dom} f} g_x(x^*)\,, \quad g_x(x^*) = \langle x^*, x\rangle - f(x)\,, \tag{6.14}$$

also ist f^* als Supremum einer Familie affiner stetiger Funktionen konvex und unterhalbstetig. Die Aussage für f^{**} wird auf dieselbe Weise bewiesen. □

Lemma 6.5 *Sei X Banachraum, $f : X \to (-\infty, \infty]$ mit $\operatorname{dom} f \neq \emptyset$. Dann gilt $f^* : X^* \to (-\infty, \infty]$ und*

$$\langle x^*, x\rangle \le f(x) + f^*(x^*)\,, \quad \textit{für alle } x \in X\,,\ x^* \in X^*\,. \tag{6.15}$$

Beweis: Folgt direkt aus der Definition von f^*. □

Satz 6.6 *Sei X Banachraum, $f : X \to (-\infty, \infty]$ mit $\operatorname{dom} f \neq \emptyset$. Dann gilt*

$$f^{**} = \sup\{g \mid g : X \to \mathbb{R} \textit{ affin und stetig, } g \le f\}\,, \tag{6.16}$$

sowie $f^{**} \le f$.

Beweis: Aus Lemma 6.5 folgt $f(x) \ge \langle x^*, x\rangle - f^*(x^*)$ für alle x, x^*, also auch

$$f(x) \ge \sup_{x^* \in X^*} (\langle x^*, x\rangle - f^*(x^*)) = f^{**}(x)\,. \tag{6.17}$$

Ist die Menge auf der rechten Seite von (6.16) leer, so gilt $f^* \equiv +\infty$ und damit $f^{**} = -\infty$, also auch (6.16). Sei nun $g : X \to \mathbb{R}$ affin und stetig mit $g \le f$,

$$g(x) = \langle x^*, x\rangle - \alpha\,, \quad x^* \in X^*\,,\ \alpha \in \mathbb{R}\,. \tag{6.18}$$

Dann ist $\langle x^*, x\rangle - \alpha \le f(x)$ für alle $x \in X$, also

$$\alpha \ge \langle x^*, x\rangle - f(x) \tag{6.19}$$

für alle $x \in X$, also $\alpha \ge f^*(x^*)$, und weiter

$$g(x) = \langle x^*, x\rangle - \alpha \le \langle x^*, x\rangle - f^*(x^*) \le f^{**}(x)\,. \tag{6.20}$$

Also ist in diesem Fall f^{**} konvex, unterhalbstetig und eigentlich. Nach Satz 5.24 gilt

$$f^{**} = \sup\{g|\ g : X \to \mathbb{R} \text{ affin und stetig, } g \le f^{**}\}\,. \tag{6.21}$$

Nach dem eben Gezeigten gilt für ein affin lineares $g : X \to \mathbb{R}$

$$g \le f \quad \Leftrightarrow \quad g \le f^{**}\,.$$

Hieraus folgt die Behauptung. □

Folgerung 6.7 *Sei X Banachraum, $f : X \to (-\infty, \infty]$ mit* $\operatorname{dom} f \neq \emptyset$. *Dann gilt*

$$f^{**} = \sup\{g|\ g : X \to (-\infty,\infty] \textit{ konvex und unterhalbstetig, } g \le f\}\,. \tag{6.22}$$

Beweis: "$\le$": Folgt direkt aus Satz 6.6. "$\ge$": Ist $g : X \to (-\infty, \infty]$ konvex und unterhalbstetig mit $g \le f$, so folgt aus Satz 5.24

$$g = \sup\{h|\ h : X \to \mathbb{R} \text{ affin und stetig, } h \le g\}\,, \tag{6.23}$$

also $g \le f^{**}$. □

Folgerung 6.8 *Sei X Banachraum, $f : X \to (-\infty, \infty]$ konvex, unterhalbstetig und eigentlich. Dann ist $f^{**} = f$.*

Beweis: Folgt direkt aus Folgerung 6.7, da $f^{**} \le f$ nach Satz 6.6. □

Folgerung 6.9 *Sei X Banachraum, $f : X \to [-\infty, \infty]$. Dann gilt $f^{***} = f^*$.*

Beweis: Übung. □

Folgerung 6.10 *Sei X Banachraum, $K \subset X$. Dann gilt*

$$I_K^{**} = I_{\mathrm{cl}(\mathrm{co}(K))}\,. \tag{6.24}$$

Beweis: Übung.

7 Das Subdifferential

Definition 7.1 (Subdifferential)
Sei X Banachraum, $f : X \to [-\infty, \infty]$. Ein $x^* \in X^*$ heißt Subgradient für f in x, wenn $f(x) \in \mathbb{R}$ und

$$f(y) \geq f(x) + \langle x^*, y - x \rangle \ , \quad \text{für alle } y \in X \, . \tag{7.1}$$

Die Menge

$$\partial f(x) = \{x^* : x^* \in X^*, \, x^* \text{ Subgradient für } f \text{ in } x\} \tag{7.2}$$

heißt Subdifferential von f in x. Falls $f(x) = \pm\infty$, so setzen wir $\partial f(x) = \emptyset$.

Gibt es ein $y \in X$ mit $f(y) = -\infty$, so ist $\partial f(x) = \emptyset$ für alle $x \in X$.

Beispiel 7.2

(i) Für $f : \mathbb{R} \to \mathbb{R}$, $f(x) = |x|$, gilt $\partial f(x) = \{1\}$ falls $x > 0$, $\partial f(x) = \{-1\}$ falls $x < 0$, sowie $\partial f(0) = [-1, 1]$.

(ii) Für $f : \mathbb{R} \to \mathbb{R}$,

$$f(x) = \begin{cases} 1 \, , & x \geq 0 \, , \\ 0 \, , & x < 0 \, , \end{cases} \tag{7.3}$$

gilt $\partial f(0) = \emptyset$. Definieren wir aber

$$f(x) = \begin{cases} 1 \, , & x > 0 \, , \\ 0 \, , & x \leq 0 \, , \end{cases} \tag{7.4}$$

so gilt $\partial f(0) = \{0\}$.

Definition 7.3 (Stützfunktional) *Sei X Banachraum, $K \subset X$ konvex. $x \in K$. Ein $x^* \in X^*$ heißt Stützfunktional für K in x, falls*

$$\langle x^*, x \rangle \geq \langle x^*, y \rangle \ , \quad \text{für alle } y \in K \, . \tag{7.5}$$

Lemma 7.4 *Sei X Banachraum, $K \subset X$ konvex, $K \neq \emptyset$. Dann gilt*

$$\partial I_K(x) = \{x^* : x^* \text{ Stützfunktional für } K \text{ in } x\} \, , \quad \text{falls } x \in K \, , \tag{7.6}$$

und $\partial I_K(x) = \emptyset$ andernfalls.

Beweis: Direkt aus der Definition. □

Satz 7.5 *Sei X Banachraum, $f : X \to [-\infty, \infty]$, $x \in X$ mit $f(x) < \infty$. Dann gilt*

$$f(x) = \min_{y \in X} f(y) \qquad \Leftrightarrow \qquad 0 \in \partial f(x) \, . \tag{7.7}$$

Beweis: Direkt aus der Definition. □

Satz 7.6 *Sei X Banachraum, $f : X \to (-\infty, \infty]$, $x \in X$. Dann gilt*

$$x^* \in \partial f(x) \qquad \Leftrightarrow \qquad f(x) + f^*(x^*) = \langle x^*, x \rangle \ , \tag{7.8}$$

$$x \in \partial f^*(x^*) \qquad \Leftrightarrow \qquad f^{**}(x) + f^*(x^*) = \langle x^*, x \rangle \ . \tag{7.9}$$

Auf der linken Seite von (7.9) wird x als Element des Bidualraums X^{**} aufgefaßt.

Beweis: Wir zeigen (7.8), der Beweis von (7.9) verläuft analog. Es gilt

$$\begin{aligned} x^* \in \partial f(x) \ &\Leftrightarrow \ f(x) \in \mathbb{R} \text{ und } f(y) \geq f(x) + \langle x^*, y - x \rangle \quad \forall\, y \in X \\ &\Leftrightarrow \ f(x) \in \mathbb{R} \text{ und } \langle x^*, x \rangle - f(x) \geq \langle x^*, y \rangle - f(y) \quad \forall\, y \in X \\ &\Leftrightarrow \ f(x) \in \mathbb{R} \text{ und } \langle x^*, x \rangle - f(x) \geq f^*(x^*) \\ &\Leftrightarrow \ \langle x^*, x \rangle \geq f(x) + f^*(x^*) \, . \end{aligned} \tag{7.10}$$

Die andere Ungleichung gilt immer, nach Lemma 6.5. □

Satz 7.7 *Sei X Banachraum, $f : X \to (-\infty, \infty]$, $x \in X$ mit $\partial f(x) \neq \emptyset$. Dann gilt $f^{**}(x) = f(x)$ und*

$$x^* \in \partial f(x) \quad \Rightarrow \quad x \in \partial f^*(x^*) \, . \tag{7.11}$$

Beweis: Sei $x^* \in \partial f(x)$. Aus $f^{**}(x) \geq \langle x^*, x \rangle - f^*(x^*)$ folgt

$$\langle x^*, x \rangle \leq f^{**}(x) + f^*(x^*) \leq f(x) + f^*(x^*) \leq \langle x^*, x \rangle \ , \tag{7.12}$$

also gilt überall die Gleichheit und aus Satz 7.6 folgt, daß $x \in \partial f^*(x^*)$. □

Satz 7.8 *Sei X Banachraum, $f : X \to (-\infty, \infty]$ konvex, unterhalbstetig und eigentlich. Dann gilt*

$$x^* \in \partial f(x) \quad \Leftrightarrow \quad x \in \partial f^*(x^*) \, . \tag{7.13}$$

Beweis: Nach Folgerung 6.8 ist $f^{**} = f$. Die Behauptung folgt dann aus Satz 7.6. □

Satz 7.9 *Sei X Banachraum, $f : X \to (-\infty, \infty]$ konvex, $x \in X$. Ist f stetig in x, so gilt $\partial f(x) \neq \emptyset$.*

Beweis: Aus der Stetigkeit von f in x folgt, daß $\operatorname{int}(\operatorname{epi} f)$ nichtleer ist (siehe Übung). Es ist $(x, f(x)) \notin \operatorname{int}(\operatorname{epi} f)$, da $(x, f(x) - \delta) \notin \operatorname{epi} f$ für alle $\delta > 0$. Aus dem Trennungssatz folgt, daß es ein $z^* \in (X \times \mathbb{R})^*$ gibt mit

$$z^*(x, f(x)) \leq z^*(\operatorname{epi} f) \, , \quad z^* \neq 0 \, . \tag{7.14}$$

Wir zerlegen z^* in

$$z^*(\xi, \alpha) = x^*(\xi) + \lambda\alpha \, , \quad x^* \in X^* \, , \ \lambda \in \mathbb{R} \, . \tag{7.15}$$

Wäre $\lambda = 0$, so wäre

$$x^*(x) \leq x^*(\xi) \, , \quad \text{für alle } \xi \in \operatorname{dom} f \, , \tag{7.16}$$

woraus wegen $x \in \operatorname{int}(\operatorname{dom} f)$ folgen würde, daß $x^* = 0$ gilt, was im Widerspruch zu $z^* \neq 0$ steht. Also ist $\lambda > 0$. Für alle $\xi \in \operatorname{dom} f$ gilt

$$x^*(x) + \lambda f(x) \leq x^*(\xi) + \lambda f(\xi)\,, \tag{7.17}$$

also folgt

$$f(\xi) \geq f(x) + \left\langle -\frac{1}{\lambda} x^*\,,\, \xi - x \right\rangle \tag{7.18}$$

für alle $\xi \in X$ und damit

$$-\frac{1}{\lambda} x^* \in \partial f(x)\,. \tag{7.19}$$

□

Lemma 7.10 *Sei X Banachraum, $f : X \to (-\infty, \infty]$, $\lambda > 0$. Dann gilt*

$$\partial(\lambda f)(x) = \lambda \partial f(x)\,, \quad \forall\, x \in X\,. \tag{7.20}$$

Beweis: Direkt aus der Definition. □

Ebenso folgt direkt aus der Definition, sofern $f_1(x) + f_2(x)$ definiert ist,

$$\partial(f_1 + f_2)(x) \supset \partial f_1(x) + \partial f_2(x)\,. \tag{7.21}$$

Satz 7.11 *Sei X Banachraum, seien $f_1, \ldots, f_n : X \to (-\infty, \infty]$ konvex. Es gebe ein $\tilde{x} \in X$ mit $\tilde{x} \in \cap_{i=1}^n \operatorname{dom} f_i$, es seien alle f_i mit $1 \leq i \leq n-1$ stetig in $\tilde{x}$. Dann gilt*

$$\partial\left(\sum_{i=1}^n f_i\right)(x) = \sum_{i=1}^n \partial f_i(x)\,, \quad \textit{für alle } x \in X\,. \tag{7.22}$$

Beweis: Wegen (7.21) ist nur "$\subset$": zu zeigen. Wir betrachten den Fall $n = 2$; der allgemeine Fall folgt mit Induktion aus

$$\partial\left(\sum_{i=1}^k f_i\right)(x) = \partial\left(\sum_{i=1}^{k-1} f_i\right)(x) + \partial f_k(x)\,, \quad k = 3, \ldots, n\,. \tag{7.23}$$

Sei also $x^* \in \partial(f_1 + f_2)(x)$, also

$$f_1(y) + f_2(y) \geq f_1(x) + f_2(x) + \langle x^*, y - x\rangle\ , \quad \forall\, y \in X\,, \tag{7.24}$$

und $f_1(x) + f_2(x) \in \mathbb{R}$. Wir definieren $g : X \to (-\infty, \infty]$ durch

$$g(y) = f_1(y) - f_1(x) - \langle x^*, y - x\rangle\ , \tag{7.25}$$

dann ist

$$g(y) \geq f_2(x) - f_2(y)\,, \quad \forall\, y \in X\,. \tag{7.26}$$

Wir definieren weiter

$$C = \operatorname{epi} g\,, \quad D = \{(z, \gamma) : f_2(z) + \gamma \leq f_2(x)\}\,. \tag{7.27}$$

C ist konvex, da g als Summe konvexer Funktionen konvex ist. D ist als Subniveaumenge einer konvexen Funktion konvex. Weiter gilt $\operatorname{int} C \neq \emptyset$, da $\tilde{x} \in \operatorname{dom} g$ und g stetig in $\tilde{x}$ ist. Wir zeigen, daß $\operatorname{int}(C) \cap D$ leer ist: Sei $(y, \beta) \in \operatorname{int}(C)$, dann gibt es ein $\varepsilon > 0$ mit $(y, \beta - \varepsilon) \in C$. Es gilt dann

$$f_2(y) + \beta \geq f_2(x) - g(y) + \beta \geq f_2(x) - (\beta - \varepsilon) + \beta = f_2(x) + \varepsilon\,, \tag{7.28}$$

also $(y, \beta) \notin D$. Aus dem Trennungssatz folgt, daß es ein $z^* \in (X \times \mathbb{R})^*$ und ein $\alpha \in \mathbb{R}$ gibt mit

$$z^*(C) \leq \alpha \leq z^*(D)\,, \quad z^*(\operatorname{int} C) < \alpha\,. \tag{7.29}$$

Wir zerlegen wieder $z^*(y, \beta) = y^*(x) + \lambda\beta$. Aus $(x, 0) \in C \cap D$ folgt $\langle y^*, x\rangle = \alpha$, also

$$\langle y^*, y\rangle + \lambda\beta \leq \langle y^*, x\rangle = \alpha \leq \langle y^*, z\rangle + \lambda\gamma\,, \tag{7.30}$$

für alle $(y, \beta) \in C$, $(z, \gamma) \in D$. Wir zeigen nun $\lambda < 0$: Es gilt $(\tilde{x}, g(\tilde{x}) + 1) \in \operatorname{int} C$ und $(\tilde{x}, f_2(x) - f_2(\tilde{x})) \in D$, also

$$\langle y^*, \tilde{x}\rangle + \lambda(g(\tilde{x}) + 1) < \langle y^*, \tilde{x}\rangle + \lambda(f_2(x) - f_2(\tilde{x}))\,, \tag{7.31}$$

also

$$\lambda(g(\tilde{x}) - f_2(x) + f_2(\tilde{x}) + 1) < 0\,, \tag{7.32}$$

und die Klammer ist positiv wegen (7.26), also $\lambda < 0$. Wir schreiben nun

$$x^* = \left(x^* - \frac{1}{\lambda}y^*\right) + \frac{1}{\lambda}y^*\,. \tag{7.33}$$

Um den Beweis zu vollenden, genügt es zu zeigen, daß

$$x^* - \frac{1}{\lambda}y^* \in \partial f_1(x)\,, \quad \frac{1}{\lambda}y^* \in \partial f_2(x)\,. \tag{7.34}$$

Wir zeigen zunächst die linke Inklusion. Für alle $y \in \operatorname{dom} f_1 = \operatorname{dom} g$ gilt $(y, g(y)) \in C$, also nach (7.30)

$$\langle y^*, y\rangle + \lambda g(y) \leq \langle y^*, x\rangle\ . \tag{7.35}$$

Division durch λ und Einsetzen der Definition von g ergibt

$$f_1(y) \geq f_1(x) + \langle x^*, y - x\rangle + \left\langle -\frac{1}{\lambda}y^*\,,\, y - x\right\rangle\,, \tag{7.36}$$

und damit die linke Inklusion in (7.34). Zum Beweis der rechten Inklusion beachten wir, daß $(z, f_2(x) - f_2(z)) \in D$ für alle $z \in \operatorname{dom} f_2$ gilt. Aus (7.30) folgt nun

$$\langle y^*, x\rangle \leq \langle y^*, z\rangle + \lambda(f_2(x) - f_2(z))\,. \tag{7.37}$$

Division durch λ liefert

$$f_2(z) \geq f_2(x) + \left\langle \frac{1}{\lambda}y^*\,,\, z - x\right\rangle\,, \quad z \in \operatorname{dom} f_2\,, \tag{7.38}$$

also auch die rechte Inklusion in (7.35). □

Wir betrachten nun das Subdifferential einer Komposition $f \circ A$, wobei $A : X \to Y$ linear und stetig ist. Wir erinnern an die Definition der zu A adjungierten Abbildung $A^* : Y^* \to X^*$,

$$\langle A^* y^*, x\rangle = \langle y^*, Ax\rangle\ . \tag{7.39}$$

Satz 7.12 *Seien X, Y Banachräume, $A : X \to Y$ linear und stetig, $f : Y \to (-\infty, \infty]$ konvex, und es gebe ein $\tilde{x} \in X$, so daß f stetig in $A\tilde{x}$ ist. Dann gilt*

$$\partial(f \circ A)(x) = A^* \partial f(Ax) \, , \quad \textit{für alle } x \in X \, . \tag{7.40}$$

Beweis: "$\supset$": Sei $y^* \in \partial f(Ax)$. Dann gilt

$$f(y) \geq f(Ax) + \langle y^*, y - Ax \rangle \ , \quad \text{für alle } y \in Y \, , \tag{7.41}$$

also

$$f(A\xi) \geq f(Ax) + \langle y^*, A\xi - Ax \rangle \ , \quad \text{für alle } \xi \in X \, , \tag{7.42}$$

also

$$(f \circ A)(\xi) \geq (f \circ A)(x) + \langle A^* y^*, \xi - x \rangle \ , \quad \text{für alle } \xi \in X \, , \tag{7.43}$$

also $A^* y^* \in \partial(f \circ A)(x)$.
"$\subset$": Sei $x^* \in \partial(f \circ A)(x)$. Wir definieren eine affine Teilmenge U von $Y \times \mathbb{R}$ durch

$$U = \{(A\xi, f(Ax) + \langle x^*, \xi - x \rangle) : \xi \in X\} \, . \tag{7.44}$$

Es ist $\operatorname{int}(\operatorname{epi} f) \neq \emptyset$, da $A\tilde{x} \in \operatorname{int}(\operatorname{epi} f)$. Wir behaupten $U \cap \operatorname{int}(\operatorname{epi} f) = \emptyset$: Ist $(y, \beta) \in \operatorname{int}(\operatorname{epi} f)$, so gilt für hinreichend kleines $\varepsilon > 0$

$$f(y) \leq \beta - \varepsilon \, . \tag{7.45}$$

Wäre $(y, \beta) \in U$, so gäbe es ein $\xi \in X$ mit

$$f(A\xi) \leq f(Ax) + \langle x^*, \xi - x \rangle - \varepsilon \tag{7.46}$$

im Widerspruch zu $x^* \in \partial(f \circ A)(x)$. Nach dem Trennungssatz gibt es also ein $z^* \in (Y \times \mathbb{R})^*$ und ein $\alpha \in \mathbb{R}$ mit

$$z^*(U) \leq z^*(\operatorname{epi} f) \, , \quad \alpha < z^*(\operatorname{int}(\operatorname{epi} f)) \, . \tag{7.47}$$

Wir schreiben

$$z^*(y, \beta) = y^*(y) + \lambda\beta \, , \quad y^* \in Y^* \, , \; \lambda \in \mathbb{R} \, . \tag{7.48}$$

Dann ist

$$\langle y^*, A\xi \rangle + \lambda f(Ax) + \lambda \langle x^*, \xi - x \rangle \leq \alpha \leq \langle y^*, y \rangle + \lambda\mu \, , \tag{7.49}$$

für alle $\xi \in X$, $y \in \operatorname{dom} f$, $\mu \geq f(y)$. Wie früher folgt $\lambda \geq 0$. Wäre $\lambda = 0$, so wäre nach Voraussetzung an $\tilde{x}$

$$\langle y^*, A\tilde{x} \rangle \leq \alpha < \langle y^*, A\tilde{x} \rangle \ , \tag{7.50}$$

also ist $\lambda > 0$. Aus (7.48) folgt

$$\langle y^*, A\xi \rangle + \lambda \langle x^*, \xi \rangle = 0 \, , \quad \forall \, \xi \in X \, , \tag{7.51}$$

also ist

$$x^* = A^* \left(-\frac{1}{\lambda} y^* \right) \, . \tag{7.52}$$

Damit wird (7.49) zu

$$\lambda f(Ax) - \lambda \langle x^*, x \rangle \leq \langle y^*, y \rangle + \lambda f(y) \, , \quad y \in \operatorname{dom} f \, , \tag{7.53}$$

und $-\lambda \langle x^*, x \rangle = \langle A^* y^*, x \rangle = \langle y^*, Ax \rangle$, also

$$f(y) \geq f(Ax) + \left\langle -\frac{1}{\lambda} y^*, y - Ax \right\rangle , \tag{7.54}$$

woraus

$$-\frac{1}{\lambda} y^* \in \partial f(Ax) \tag{7.55}$$

und wegen (7.52) auch $x^* \in A^* \partial f(Ax)$ folgt. $\square$

8 Differenzierbarkeit konvexer Funktionen

Lemma 8.1 *Sei $\varphi : \mathbb{R} \to (-\infty, \infty]$ konvex, $x \in \operatorname{dom} \varphi$. Dann wird durch*

$$d(t) = \frac{\varphi(x+t) - \varphi(x)}{t} \tag{8.1}$$

eine auf $(0, \infty)$ monoton wachsende Funktion d mit Werten in $(-\infty, \infty]$ definiert. Ferner gilt $d(-t) \le d(t)$ für $t > 0$.

Beweis: Für $0 < s < t$ ist

$$x + s = \frac{t-s}{t} x + \frac{s}{t}(x+t)\,, \tag{8.2}$$

also

$$\varphi(x+s) \le \frac{t-s}{t}\varphi(x) + \frac{s}{t}\varphi(x+t)\,. \tag{8.3}$$

Subtraktion von $\varphi(x)$ und Division durch s liefert

$$d(s) = \frac{\varphi(x+s) - \varphi(x)}{s} \le \frac{\varphi(x+t) - \varphi(x)}{t} = d(t)\,. \tag{8.4}$$

Es gilt weiter für $t > 0$

$$\varphi(x) \le \frac{1}{2}\varphi(x-t) + \frac{1}{2}\varphi(x+t)\,, \tag{8.5}$$

also $\varphi(x) - \varphi(x-t) \le \varphi(x+t) - \varphi(x)$ und damit

$$\frac{\varphi(x-t) - \varphi(x)}{-t} \le \frac{\varphi(x+t) - \varphi(x)}{t}\,. \tag{8.6}$$

$\square$

Definition 8.2 (Strikt konvexe Funktion) *Sei X Vektorraum, $K \subset X$ konvex. Eine Funktion $f : K \to \mathbb{R}$ heißt strikt konvex auf K, falls*

$$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda) f(y) \tag{8.7}$$

gilt für alle $x, y \in K$ mit $x \ne y$ und alle $\lambda \in (0,1)$.

Lemma 8.3 *Sei $K \subset \mathbb{R}$ konvex, $\varphi : K \to \mathbb{R}$ strikt konvex, $x \in K$. Dann ist*

$$d(t) = \frac{\varphi(x+t) - \varphi(x)}{t} \tag{8.8}$$

streng monoton wachsend in $(0, \infty) \cap \operatorname{dom} d$.

Beweis: Wie der Beweis von Lemma 8.1, aber mit der strikten Ungleichung. $\square$

Definition 8.4 (Richtungsableitung) *Sei X Vektorraum, $f : X \to (-\infty, \infty]$, $x \in \operatorname{dom} f$ und $h \in X$. Falls der Grenzwert*

$$f'(x; h) = \lim_{t \downarrow 0} \frac{f(x+th) - f(x)}{t}\,. \tag{8.9}$$

in $(-\infty, \infty]$ existiert, so heißt er die Richtungsableitung von f in x in Richtung h.

Unmittelbar aus der Definition folgt, dass die Richtungsableitung positiv homogen ist,

$$f'(x;th) = tf'(x;h)\,, \quad x,h \in X\,,\ t > 0\,. \tag{8.10}$$

Satz 8.5 *Sei X Vektorraum, $f : X \to (-\infty,\infty]$ konvex, $x \in \operatorname{dom} f$. Dann existiert $f'(x;h)$ für jedes $h \in X$, und es gilt*

$$f'(x;h) = \inf_{t>0} \frac{f(x+th)-f(x)}{t}\,. \tag{8.11}$$

Ferner gelten

$$f'(x;h) \le f(x+h) - f(x)\,, \tag{8.12}$$

sowie

$$-f'(x;-h) \le f'(x;h)\,. \tag{8.13}$$

Beweis: Nach Lemma 8.1, angewendet auf $\varphi(t) = f(x+th)$, ist der Differenzenquotient

$$d_h(t) = \frac{f(x+th)-f(x)}{t} \tag{8.14}$$

monoton wachsend auf $(0,\infty)$, also existiert $\lim_{t\downarrow 0} d_h(t)$ und ist gleich $\inf_{t>0} d_h(t)$. Wegen $d_h(1) = f(x+h) - f(x)$ gilt (8.12), und (8.13) folgt aus

$$-\frac{f(x-th)-f(x)}{t} = d_h(-t) \le d_h(t)\,, \tag{8.15}$$

durch Grenzübergang $t \downarrow 0$. □

Definition 8.6 (Gâteaux-Ableitung)
Sei X Vektorraum, $M \subset X$ algebraisch offen, $f : M \to \mathbb{R}$, $x \in M$. Falls die Richtungsableitungen $f'(x;h)$ für alle $h \in X$ existieren und die Abbildung $h \mapsto f'(x;h)$ linear ist, so heißt die durch

$$\langle f'(x), h\rangle = f'(x;h) \tag{8.16}$$

definierte Abbildung $f'(x) : X \to \mathbb{R}$ die Gâteaux-Ableitung von f in x. Hat f in jedem Punkt $x \in M$ eine Gâteaux-Ableitung, so heißt $f' : M \to X^{\#}$ die Gâteaux-Ableitung von f in M.

In Definition 8.6 haben wir die Bezeichnung $X^{\#}$ für den Raum aller linearen (nicht unbedingt stetigen) Funktionale auf X verwendet.

Satz 8.7 *Sei X Banachraum, $f : X \to (-\infty,\infty]$ konvex, $x \in \operatorname{aint}(\operatorname{dom} f)$, f habe in x eine Gâteaux-Ableitung $f'(x) \in X^*$. Dann ist $\partial f(x) = \{f'(x)\}$.*

Beweis: Es ist

$$\langle f'(x), y-x\rangle = f'(x;y-x) \le f(y) - f(x) \tag{8.17}$$

wegen (8.12), also $f'(x) \in \partial f(x)$. Ist $x^* \in \partial f(x)$, so gilt

$$f(x+th) - f(x) \ge t\,\langle x^*, h\rangle \tag{8.18}$$

für alle $h \in X$ und alle $t > 0$, also

$$\langle f'(x), h\rangle = f'(x;h) \ge \langle x^*, h\rangle\,, \quad \text{für alle } h \in X\,, \tag{8.19}$$

also $f'(x) = x^*$. □

Bemerkung 8.8
Sei X normierter Raum, $M \subset X$ offen, $f : M \to \mathbb{R}$ und $x \in M$. Hat f in x eine Gâteaux-Ableitung mit $f'(x) \in X^*$ und gilt außerdem

$$\lim_{\substack{h \to 0 \\ h \neq 0}} \frac{|f(x+h) - f(x) - \langle f'(x), h \rangle\,|}{\|h\|} = 0\,, \tag{8.20}$$

so erhält man den üblichen Differenzierbarkeitsbegriff. In diesem Fall heißt $f'(x)$ die **Fréchet-Ableitung** von f in x.

9 Konvexe Kegel

Definition 9.1 (Kegel)
*Sei X Vektorraum, $K \subset X$ nichtleer. K heißt **Kegel**, falls $\lambda x \in K$ gilt für alle $x \in K$ und alle $\lambda \geq 0$. Ein Kegel heißt **spitz**, falls $K \cap (-K) = \{0\}$.* □

Ist K Kegel, so ist $0 \in K$. Ein Kegel ist spitz genau dann, wenn er keinen eindimensionalen Unterraum (Gerade durch den Nullpunkt) enthält.

Lemma 9.2 *Sei X Vektorraum, $K \subset X$ Kegel. K ist konvex genau dann, wenn $x+y \in K$ gilt für alle $x, y \in K$.*

Beweis: "$\Leftarrow$": Direkt aus der Definition. "$\Rightarrow$": Folgt aus der Identität

$$x + y = 2\left(\frac{1}{2}x + \frac{1}{2}y\right) .$$

□

Lemma 9.3 *Sei X Vektorraum. Eine nichtleere Teilmenge $K \subset X$ ist ein konvexer Kegel genau dann, wenn*

$$\sum_{i=1}^{n} \lambda_i x_i \in K$$

gilt für alle $n \in \mathbb{N}$, $x_i \in S$ und $\lambda_i \geq 0$. Sind K_i konvexe Kegel, $1 \leq i \leq n$, so auch

$$\sum_{i=1}^{n} K_i = \left\{ \sum_{i=1}^{n} x_i : x_i \in K_i \right\} . \tag{9.1}$$

Beweis: Folgt direkt aus Lemma 9.2 und Definition 9.1. □

Definition 9.4 (Kegelhülle)
*Sei X Vektorraum, $S \subset X$ nichtleer. Die **Kegelhülle** von S ist definiert als*

$$\operatorname{cone}(S) = \left\{ \sum_{i=1}^{n} \lambda_i x_i : n \in \mathbb{N},\ x_i \in S,\ \lambda_i \geq 0. \right\} . \tag{9.2}$$

*Eine Linearkombination $\sum_{i=1}^{n} \lambda_i x_i$ mit $\lambda_i \geq 0$ für alle i heißt auch **konische Kombination**.*

Aus Lemma 9.3 folgt unmittelbar, dass $\operatorname{cone}(S)$ gleich ist dem Durchschnitt aller konvexer Kegel, welche S umfassen. Weiter gilt

$$\dim(\operatorname{cone}(S)) \leq |S| \qquad (= \text{Anzahl der Elemente von } S) . \tag{9.3}$$

Ist S endlich, so auch $\dim(\operatorname{cone}(S))$.

Definition 9.5 (Endlich erzeugter Kegel)
*Sei X Vektorraum. Ein Kegel $K \subset K$ heißt **endlich erzeugt**, falls $K = \operatorname{cone}(S)$ für eine endliche Menge $S \subset X$.* □

Beispiel 9.6
(i) Der positive Orthant im $\mathbb{R}^n$,

$$\mathbb{R}^n_+ = \{x : \ x \in \mathbb{R}^n, \, x_i \geq 0 \text{ für alle } i\} \tag{9.4}$$

ist ein spitzer konvexer Kegel, welcher von $S = \{e_i : 1 \leq i \leq n\}$, also den Einheitsvektoren, erzeugt wird.
(ii) Ist Ω eine Menge und X ein Vektorraum reellwertiger Funktionen auf Ω, so ist

$$K = \{x : \ x \in X, \, x(t) \geq 0 \text{ für alle } t \in \Omega\} \tag{9.5}$$

ein spitzer konvexer Kegel. Ist Ω unendlich, so ist K in der Regel nicht endlich erzeugt. Betrachtet man den Raum $L^p(\Omega)$, so ist "für alle" in (9.5) durch "für fast alle" zu ersetzen.

Wir erinnern daran, dass eine Relation auf einer Menge als Ordnungsrelation bezeichnet wird, wenn sie reflexiv, transitiv und antisymmetrisch ist.

Satz 9.7 (Geordneter Vektorraum)
Sei X Vektorraum, $K \subset X$ ein spitzer konvexer Kegel. Dann wird durch

$$x \leq y \qquad \Leftrightarrow \qquad y - x \in K \tag{9.6}$$

eine Ordnungsrelation auf X definiert mit den Eigenschaften

$$x \leq y \qquad \Rightarrow \qquad x + z \leq y + z \,, \quad \lambda x \leq \lambda y \,, \quad -y \leq -x \,, \tag{9.7}$$

für alle $x, y, z \in X$ und alle $\lambda \geq 0$. Wir sagen, dass K die Ordnung "$\leq$" erzeugt.

Beweis: Folgt unmittelbar aus der Definition und Lemma 9.2. □

Die in Beispiel 9.6 erzeugten Ordnungen nennt man komponentenweise oder punktweise Ordnungen.

Lemma 9.8 *Seien X, Y Vektorräume, $L : X \to Y$ linear. Ist $K \subset X$ ein (konvexer) Kegel, so auch $L(K)$. Ist $C \subset Y$ ein (konvexer) Kegel, so auch $L^{-1}(C)$.*

Beweis: Folgt unmittelbar aus Lemma 9.3, da L konische Kombinationen auf ebensolche abbildet. □

Aus Lemma 9.8 folgt beispielsweise, dass die Menge

$$\{x : x \in \mathbb{R}^n \,, \, Ax \geq 0\} \,, \quad A \in \mathbb{R}^{(m,n)} \text{ gegeben,}$$

ein konvexer Kegel im $\mathbb{R}^m$ ist.

Definition 9.9 (Polarkegel)
*Sei X Banachraum, $S \subset X$ nichtleer. Wir definieren den **Polarkegel** von S als*

$$S^* = \{x^* : \ x^* \in X^*, \, \langle x^*, x\rangle \leq 0 \textit{ für alle } x \in S\} \,. \tag{9.8}$$

Zu nichtleerem $T \subset X^*$ definieren wir $T^* \subset X$ durch

$$T^* = \{x : \ x \in X, \, \langle x^*, x\rangle \leq 0 \textit{ für alle } x^* \in T\} \,. \tag{9.9}$$

*Im Falle $T = S^*$ erhalten wir den **Bipolarkegel** von S,*

$$S^{**} = \{x : \ x \in X, \, \langle x^*, x\rangle \leq 0 \textit{ für alle } x^* \in S^*\} \,. \tag{9.10}$$

□

Wir können S^* und S^{**} darstellen als

$$S^* = \bigcap_{x \in S} \{x^* : x^* \in X^* \,,\, \langle x^*, x\rangle \le 0\} \,, \quad S^{**} = \bigcap_{x^* \in S^*} \{x : x \in X \,,\, \langle x^*, x\rangle \le 0\} \,. \tag{9.11}$$

Es gilt offenbar

$$S_1 \subset S_2 \qquad \Rightarrow \qquad S_1^* \supset S_2^* \,, \quad S_1^{**} \subset S_2^{**} \,. \tag{9.12}$$

Beispiel 9.10
Sei X Banachraum, K Unterraum von X. Dann gilt

$$K^* = K^\perp = \{x^* : \langle x^*, x\rangle = 0 \text{ für alle } x \in K\} \,. \tag{9.13}$$

Als Spezialfall ergibt sich für Hyperebenen durch den Nullpunkt, $K = \ker(x^*)$ mit $x^* \in X^*$, dass

$$K^* = \operatorname{span}(\{x^*\}) = \{tx^* : t \in \mathbb{R}\} \,. \tag{9.14}$$

Betrachten wir den Halbraum statt der Hyperebene, also $K = \{x : \langle x^*, x\rangle \le 0\}$ für gegebenes $x^* \in X^*$, so gilt

$$K^* = \{tx^* : t \ge 0\} \,. \tag{9.15}$$

Lemma 9.11 *Sei X Banachraum, $S \subset X$ nichtleer. Dann sind S^* und S^{**} abgeschlossene konvexe Kegel, und es gilt*

$$S^* = (\overline{S})^* = (\overline{\operatorname{cone}(S)})^* \,. \tag{9.16}$$

Ist S konvexer Kegel, so gilt

$$S^{**} = \overline{S} \,. \tag{9.17}$$

Beweis: Die Darstellungsformel (9.11) zeigt, dass S^* Durchschnitt abgeschlossener konvexer Kegel ist, da aus $\langle x_i^*, x\rangle \le 0$ folgt, dass $\sum_{i=1}^n \langle \lambda_i x_i^*, x\rangle \le 0$ gilt für jede konische Kombination. Analog wird bewiesen, dass S^{**} abgeschlossener konvexer Kegel ist. Die Inklusionen $S^* \supset (\overline{S})^* \supset (\overline{\operatorname{cone}(S)})^*$ folgen aus (9.12).
"$S^* \subset (\overline{S})^*$": Sei $x^* \in S^*$. Ist $x \in \overline{S}$ und $\{x_k\}$ Folge in S mit $x_k \to x \in X$, so gilt $\langle x^*, x\rangle = \lim_k \langle x^*, x_k\rangle \le 0$. Damit folgt auch $(\operatorname{cone}(S))^* = (\overline{\operatorname{cone}(S)})^*$.
"$S^* \subset (\operatorname{cone}(S))^*$": Ist $x^* \in S^*$, so gilt $\langle x^*, x\rangle = \sum_{i=1}^n \lambda_i \langle x^*, x_i\rangle \le 0$ für jede konische Kombination $\sum_{i=1}^n \lambda_i x_i$ von Elementen $x_i \in S$.
"$S^{**} \supset \overline{S}$": Ist $x \in \overline{S}$, so gilt $\langle x^*, x\rangle \le 0$ für alle $x^* \in S^*$ wegen (9.16), und damit $x \in S^{**}$.
"$S^{**} \subset \overline{S}$": Ist $x \notin \overline{S}$, so gibt es nach Trennungssatz ein $x^* \in X^*$ mit

$$x^*(S) \le \sup x^*(S) =: \alpha < x^*(x) \,.$$

Da S konvexer Kegel ist, muss $\alpha = 0$ gelten. Es folgt $x^* \in S^*$ und daher $x \notin S^{**}$. □

Lemma 9.12 *Sei X Banachraum, $T \subset X^*$ nichtleer. Dann gilt*

$$T^* = (\overline{T})^* = (\overline{\operatorname{cone}(T)})^* \,. \tag{9.18}$$

Ist X außerdem reflexiv und T konvexer Kegel, so gilt

$$T^{**} = \overline{T} \,. \tag{9.19}$$

Beweis: Analog zum Beweis von Lemma 9.11. Die Reflexivität von X ist erforderlich, damit Funktionale in X^{**}, die bei der Anwendung des Trennungssatzes in X^* auftreten, auf Elemente in X zurückgeführt werden können. $\square$

Lemma 9.13 *Sei X Banachraum, seien K_i konvexe Kegel in X, $i \in I$. Dann gilt*

$$\left(\bigcup_{i \in I} K_i \right)^* = \bigcap_{i \in I} K_i^* . \tag{9.20}$$

Sind T_i konvexe Kegel in X^*, $i \in I$, so gilt

$$\left(\bigcup_{i \in I} T_i \right)^* = \bigcap_{i \in I} T_i^* . \tag{9.21}$$

Beweis: Beide Mengen in (9.20) enthalten genau diejenigen $x^* \in X^*$, für die gilt $\langle x^*, x \rangle \leq 0$ für alle $i \in I$ und alle $x \in K_i$. Analog wird (9.21) gezeigt. $\square$

Satz 9.14 *Sei X reflexiver Banachraum, seien K_i abgeschlossene konvexe Kegel in X, $i \in I$. Dann gilt*

$$\left(\bigcap_{i \in I} K_i \right)^* = \overline{\operatorname{cone} \bigcup_{i \in I} K_i^*} . \tag{9.22}$$

Beweis: Wir definieren $T \subset X^*$ durch

$$T = \bigcup_{i \in I} K_i^* .$$

Dann gilt wegen (9.21) und (9.17)

$$T^* = \left(\bigcup_{i \in I} K_i^* \right)^* = \bigcap_{i \in I} K_i^{**} = \bigcap_{i \in I} K_i ,$$

und weiter wegen Lemma 9.12

$$\left(\bigcap_{i \in I} K_i \right)^* = T^{**} = \overline{\operatorname{cone}(T)}^{**} = \overline{\operatorname{cone}(T)} .$$

$\square$

Handelt es sich in Satz 9.14 um eine endliche Indexmenge, so wird (9.22) zu

$$\left(\bigcap_{i \in I} K_i \right)^* = \overline{\sum_{i \in I} K_i^*} , \tag{9.23}$$

da

$$\operatorname{cone} \bigcup_{i=1}^n S_i = \sum_{i=1}^n S_i , \quad \operatorname{cone} \bigcup_{i=1}^n T_i = \sum_{i=1}^n T_i$$

gelten für beliebige $S_i \subset X$ und $T_i \subset X^*$. Es ist daher von Interesse festzustellen, wann die Summe abgeschlossener konvexer Kegel abgeschlossen ist.

Satz 9.15 (Endlich erzeugte Kegel sind abgeschlossen)
Sei X Banachraum, $K = \text{cone}\,(x_1, \ldots, x_n)$ ein endlich erzeugter Kegel in X. Dann gilt:
(i) K ist abgeschlossen.
(ii) Es gibt ein $c > 0$, so dass für jedes $x \in K$ ein $\lambda \in \mathbb{R}^n$ existiert mit $\lambda \geq 0$ und

$$\|\lambda\| \leq c\|x\|\,, \quad x = \sum_{i=1}^{n} \lambda_i x_i\,. \tag{9.24}$$

Beweis: "(ii)$\Rightarrow$(i)": Sei $y \in \overline{K}$, sei $\{y^k\}$ Folge in K mit $y^k \to y$. Zu y^k wählen wir $\lambda^k \in \mathbb{R}^n$ gemäß (ii). Da $\{y^k\}$ beschränkt ist, ist nach (9.24) auch die Folge $\{\lambda^k\}$ beschränkt in $\mathbb{R}^n$, also gilt $\lambda^{k_m} \to \lambda \in \mathbb{R}^n$ für eine Teilfolge. Grenzübergang liefert $\lambda \geq 0$ und $y = \sum_{i=1}^{n} \lambda_i x_i$, also $y \in K$.
Wir zeigen (ii). Zu $\lambda \in \mathbb{R}^n_+$ definieren wir

$$I(\lambda) = \{i : \lambda_i > 0\}\,.$$

Wir nennen $\lambda \in \mathbb{R}^n_+$ minimal, falls es kein $\mu \in \mathbb{R}^n_+$ gibt mit

$$\sum_{i=1}^{n} \lambda_i x_i = \sum_{i=1}^{n} \mu_i x_i\,, \quad I(\mu) \subset I(\lambda)\,, \quad I(\mu) \neq I(\lambda)\,.$$

Da es nur endlich viele Teilmengen J von $\{1, \ldots, n\}$ gibt, und da jedes $x \in K$ sich mit einem minimalen λ konisch darstellen lässt, genügt es zu zeigen: Zu jeder Teilmenge J von $\{1, \ldots, n\}$ gibt es eine Konstante c_J, so dass

$$\|\lambda\| \leq c_J\|x\|\,, \quad x = \sum_{i=1}^{n} \lambda_i x_i \tag{9.25}$$

gilt für alle minimalen λ mit $J = I(\lambda)$.
Zum Beweis nehmen wir das Gegenteil an. Dann gibt es eine Indexmenge J und eine Folge minimaler $\lambda^k \in \mathbb{R}^n_+$ mit $J = I(\lambda^k)$ und

$$\left\| \sum_{i \in J} \lambda_i^k x_i \right\| \leq \frac{1}{k}\|\lambda^k\|\,.$$

Es folgt, dass $\{\lambda^k\|\lambda^k\|^{-1}\}$ einen Häufungspunkt $\mu \in \mathbb{R}^n$ hat. Für diesen gilt $\|\mu\| = 1$, $\mu \geq 0$ und

$$\sum_{i \in J} \mu_i x_i = 0\,.$$

Sei nun $\lambda \geq 0$ minimal mit $J = I(\lambda)$. Da $\lambda_i > 0$ für alle $i \in J$ und $\mu \neq 0$, gibt es ein $j \in J$ und ein $t > 0$ mit

$$0 = \lambda_j - t\mu_j = \min_{i \in J}(\lambda_i - t\mu_i)\,.$$

Die Indexmenge $I(\lambda - t\mu)$ ist daher echt in J enthalten, und

$$\sum_{i=1}^{n} (\lambda_i - t\mu_i) x_i = \sum_{i=1}^{n} \lambda_i x_i$$

im Widerspruch zur Minimalität von λ. □

Satz 9.16 (Lemma von Farkas)
Sei $A \in \mathbb{R}^{(m,n)}$ *mit den Zeilen* $a_1, \ldots, a_m$*, sei*

$$K = \{x : x \in \mathbb{R}^n \,,\, Ax \leq 0\} \,. \tag{9.26}$$

Dann gilt

$$K^* = \mathrm{cone}\,(a_1, \ldots, a_m) \,. \tag{9.27}$$

Beweis: Wir setzen

$$K_i = \{x : x \in \mathbb{R}^n \,,\, a_i^T x \leq 0\} \,.$$

Nach Beispiel 9.10 gilt

$$K_i^* = \{ta_i : t \geq 0\} \,.$$

Es folgt

$$\sum_{i=1}^m K_i^* = \mathrm{cone}\,(a_1, \ldots, a_m) \,.$$

Nach Satz 9.15 ist dieser konvexe Kegel abgeschlossen. Aus (9.23) folgt daher

$$K^* = \left(\bigcap_{i=1}^m K_i \right)^* = \sum_{i=1}^m K_i^* = \mathrm{cone}\,(a_1, \ldots, a_m) \,.$$

□

Man kann das Lemma von Farkas auch in Form einer Alternative formulieren.

Satz 9.17 (Lemma von Farkas, andere Form)
Sei $A \in \mathbb{R}^{(m,n)}$ *mit den Zeilen* $a_1, \ldots, a_m$*, sei* $b \in \mathbb{R}^n$*. Dann ist trifft genau eine der folgenden beiden Aussagen zu:*
(i) Es gibt ein $x \in \mathbb{R}^n$ *mit* $Ax \leq 0$ *und* $b^T x > 0$*.*
(ii) Es gibt ein $y \in \mathbb{R}^m$ *mit* $A^T y = b$ *und* $y \geq 0$*.*

Beweis: Sei $K = \{x : x \in \mathbb{R}^n \,,\, Ax \leq 0\}$. Wir setzen

$$\begin{aligned} M_1 &= \{x : x \in \mathbb{R}^n \,,\, Ax \leq 0 \,,\, b^T x > 0\} \,, \\ M_2 &= \{y : y \in \mathbb{R}^m \,,\, A^T y = b \,,\, y \geq 0\} \,. \end{aligned}$$

Dann gilt (die mittlere Äquivalenz folgt aus Satz 9.16)

$$M_1 = \emptyset \quad \Leftrightarrow \quad b \in K^* \quad \Leftrightarrow \quad b \in \mathrm{cone}\,(a_1, \ldots, a_m) \quad \Leftrightarrow \quad M_2 \neq \emptyset \,.$$

□

Man kann die Argumentation im Beweis von Satz 9.17 umdrehen und auf diese Weise Satz 9.16 aus Satz 9.17 folgern. Setzen wir die Gültigkeit von Satz 9.16 voraus (etwa, weil wir ihn auf andere Weise bewiesen haben), so folgt

$$b \in K^* \quad \Leftrightarrow \quad M_1 = \emptyset \quad \Leftrightarrow \quad M_2 \neq \emptyset \quad \Leftrightarrow \quad b \in \mathrm{cone}\,(\{a_1, \ldots, a_m\} \,.$$