Konvexe Analysis *

Martin Brokate †

Inhaltsverzeichnis

---

* Vorlesungsskript, SS 2008

† 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; 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.

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

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

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 $\{v_1, \dots, v_m\}$ 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 $x = a + y$

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

□

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*

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

die affine Hülle von S in X. □

Falls $0 \in S$, so gilt

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

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*

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

Falls $a \in S$, so gilt

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

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 $\text{aff}\,(S) \subset a+N$ und $\text{aff}\,(S) - a \subset N$. Da N beliebig war, folgt $\text{aff}\,(S) - a \subset \text{aff}\,(S-a)$.
"$\supset$": Nach dem eben Bewiesenen gilt $\text{aff}\,(S-a) \subset -a + \text{aff}\,(S-a+a)$. □

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

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

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

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 $x \in \text{aff}\,(S)$ 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 $\mu_0 = \lambda_0$. □

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 $\lambda \in \mathbb{R}$. □

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 $y = f(x)$ 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*

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

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 $\nu \in [0, 1]$

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

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

Definition 2.5 (Simplex) *Sei $\{x_0, \ldots, x_n\} \subset X$ affin unabhängig. Dann heißt*

$$\operatorname{co}(\{x_0, \ldots, x_n\}) \tag{2.12}$$

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

$$x = \sum_{i=0}^{n} \lambda_i x_i \tag{2.13}$$

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

Definition 2.6 (Affine Dimension einer konvexen Menge)
Sei X Vektorraum, $K \subset X$ konvex. Wir definieren die affine Dimension von K als

$$\dim(K) = \dim(\operatorname{aff}(K))\,. \tag{2.14}$$

□

Satz 2.7 *Sei $K \subset \mathbb{R}^n$ konvex. Dann gilt*

$$\dim K = \max_{\substack{S \subset K \\ S \text{ Simplex}}} \dim S\,, \tag{2.15}$$

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

$$x - x_0 \in (K - x_0) \setminus \mathrm{span}\,(S - x_0)\,, \tag{2.16}$$

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

Lemma 2.8 *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

$$\lambda x + (1 - \lambda) y = [\lambda x_1 + (1 - \lambda) y_1] + [\lambda x_2 + (1 - \lambda) y_2] \in K_1 + K_2\,. \tag{2.17}$$

□

Satz 2.9 *Sei X Vektorraum, $K \subset X$ konvex. Dann gilt*

$$(\lambda + \mu) K = \lambda K + \mu K \tag{2.18}$$

für alle $\lambda, \mu \geq 0$.

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

$$\frac{\lambda}{\lambda + \mu} K + \frac{\mu}{\lambda + \mu} K \subset K\,, \tag{2.19}$$

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

Lemma 2.10 *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: Klar. □

Definition 2.11 (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 $\mathrm{aint}\,(S)$. S heißt algebraisch offen, wenn $S = \mathrm{aint}\,(S)$.* □

Satz 2.12 *Sei X Vektorraum, $K \subset X$ konvex. Ist $x \in \mathrm{aint}\,(K)$ und $y \in K$, so gilt $[x, y) \subset \mathrm{aint}\,(K)$.*

Beweis: Sei $a \in [x, y)$, $a = \lambda x + (1 - \lambda) y$, $0 < \lambda \leq 1$. Sei $h \in X$ beliebig. Sei $[a, a + \delta h] \subset K$. 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 $a \in \mathrm{aint}\,(K)$. □

Aus Satz 2.12 folgt unmittelbar, dass $\mathrm{aint}\,(K)$ 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

$$p(tx) = tp(x)\,, \quad \textit{für alle } x \in X,\ t \geq 0, \tag{3.1}$$

$$p(x+y) \leq p(x) + p(y)\,, \quad \textit{für alle } x, y \in X. \tag{3.2}$$

□

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

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 Eichfunktonal 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 2.12, 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 2.8, da $\operatorname{aint}(K_1)$ konvex ist. Weiter ist $\operatorname{aint}(K)$ nichtleer, da $\operatorname{aint}(K) \supset \operatorname{aint}(K_1) - x_2$ für jedes $x_2 \in K_2$, und es gilt $0 \notin K$. Wir wählen gemäß Satz 3.7 ein lineares $\ell : X \to \mathbb{R}$ mit $\ell(x) \le \ell(0) = 0$ für alle $x \in K$ und $\ell(x) < 0$ für $x \in \operatorname{aint}(K)$. Es folgt

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

Sei nun $x_1 \in K_1$ beliebig. Ist $\tilde{x}_1 \in \operatorname{aint}(K_1)$, so ist $x_1 + t(\tilde{x}_1 - x_1) \in \operatorname{aint}(K_1)$ nach Satz 2.12, also gilt für alle $x_2 \in K_2$

$$0 \ge \ell(x_1 + t(\tilde{x}_1 - x_1) - x_2) = \ell(x_1 - x_2) + t\ell(\tilde{x}_1 - x_1)\,,$$

und Grenzübergang $t \downarrow 0$ liefert

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

Es folgt

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

für $\alpha = \sup\ \ell(K_1)$. Ist nun $x_1 \in X$ mit $\ell(x_1) = \alpha$, so ist $\ell(x_1 + th) > \alpha$, falls $t > 0$ und $\ell(h) > 0$, also folgt $x_1 \notin \operatorname{aint}(K_1)$, und daher muss $\ell(x_1) < \alpha$ für $x_1 \in \operatorname{aint}(K_1)$ gelten. □

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 $\ell \ne 0$, 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.

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, diese nennen wir im folgenden **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 $a \in \operatorname{int}(K)$. □

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

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

Insbesondere ist jede konvexe Nullumgebung absorbierend.

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

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

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

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 $\operatorname{aint}(A_\alpha) = A_\alpha$, 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 $\operatorname{int}(A_\alpha) = \operatorname{aint}(A_\alpha)$, 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: 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* $\operatorname{int}(K_1) \cap K_2 = \emptyset$ *und* $\operatorname{int}(K_1) \neq \emptyset$ *Dann gibt es ein $x^* \in X^*$ mit $x^* \neq 0$ und ein $\alpha \in \mathbb{R}$ mit*

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

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

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

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

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 leicht, 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 $\{x : x \in X,\ f(x) \leq \alpha\}$ 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 $f''(x) \geq 0$ 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: f ist konvex genau dann, wenn alle Funktionen

$$g_{x,y}(t) = f(x + t(y - x))\,, \quad x, y \in G\,, \tag{5.13}$$

konvex sind. Nun ist aber

$$g''_{x,y}(t) = (y - x)^T D^2 f(x + t(y - x))(y - x)\,, \tag{5.14}$$

und die Behauptung folgt aus Satz 5.6. □

Satz 5.8 *Sei X Vektorraum, seien $f : X \to (-\infty, \infty]$ und $\varphi : \mathbb{R} \to (-\infty, \infty]$ konvex, sei φ monoton wachsend ($s \leq t \Leftarrow \varphi(s) \leq \varphi(t)$). 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

$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y)\,, \tag{5.15}$$

also

$$(\varphi \circ f)(\lambda x + (1-\lambda)y) \le \varphi(\lambda f(x) + (1-\lambda)f(y)) \le \lambda\varphi(f(x)) + (1-\lambda)\varphi(f(y))\,. \tag{5.16}$$

□

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: Klar.

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

$$f(x) = \inf_{(x,\mu)\in K} \mu \tag{5.17}$$

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*

$$f(x) = \inf_{\substack{x_1, x_2 \in X \\ x_1 + x_2 = x}} (f_1(x_1) + f_2(x_2)) \tag{5.18}$$

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

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

$$K = \{(x_1 + x_2\,,\, \mu_1 + \mu_2) : x_j \in X\,,\ f_j(x_j) \le \mu_j \text{ für } j = 1, 2\}\,. \tag{5.19}$$

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

$$\inf_{(x,\mu)\in K} \mu = \inf_{\substack{x_1 + x_2 = x \\ f_j(x_j) \le \mu_j}} (\mu_1 + \mu_2) = \inf_{x_1 + x_2 = x} (f_1(x_1) + f_2(x_2)) = f(x)\,. \tag{5.20}$$

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*

$$f = f_1 \# f_2\,. \tag{5.21}$$

Es gilt
$$(f_1 \# f_2)(x) = \inf_{y\in X} (f_1(x-y) + f_2(y)) .$$
Falls das Minimum in (5.18) immer dann angenommen wird, wenn das Infimum endlich ist, gilt $\mathrm{epi}\,(f_1 \# f_2) = \mathrm{epi}\,(f_1) + \mathrm{epi}\,(f_2)$.

Definition 5.14 (Indikatorfunktion) *Sei X Vektorraum, $K \subset X$. Die durch*
$$I_K(x) = \begin{cases} 0\,, & x \in K \\ +\infty\,, & \text{sonst} \end{cases} \tag{5.22}$$
definierte Funktion $I_K : X \to [0,\infty)$ heißt Indikatorfunktion von K.

Wegen $\mathrm{epi}\, I_K = K \times \mathbb{R}_+$ ist I_K genau dann konvex, wenn K konvex ist. Für eine einpunktige Menge $K = \{a\}$ gilt
$$(f \# I_{\{a\}})(x) = f(x-a) . \tag{5.23}$$
Ist $(X, \|\cdot\|)$ normiert und wählen wir $f = \|\cdot\|$, $g = I_K$ mit $K \subset X$ konvex, so gilt
$$(f \# g)(x) = \inf_{y\in K} \|x-y\| =: d(x,K) , \tag{5.24}$$
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*
$$f = \sup_{i\in I} f_i \tag{5.25}$$
konvex.

Beweis: Es ist $f(x) \le \mu$ genau dann, wenn $f_i(x) \le \mu$ für alle $i \in I$, also folgt
$$\mathrm{epi}\, f = \bigcap_{i\in I} \mathrm{epi}\, f_i ,$$
also ist $\mathrm{epi}\, f$ konvex. □

Definition 5.16 (Unterhalbstetigkeit) *Sei X lokalkonvex, $f : X \to [-\infty,\infty]$. f heißt unterhalbstetig, wenn die Subniveaumengen*
$$M_\alpha = \{x : x \in X,\ f(x) \le \alpha\} \tag{5.26}$$
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 $\mathrm{epi}\, f$ abgeschlossen ist in $X \times \mathbb{R}$.*

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

$$F_\alpha = \{(x, \alpha) : x \in X,\ f(x) \le \alpha\} = \operatorname{epi} f \cap (X \times \{\alpha\}) \tag{5.27}$$

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.
"⇒": 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: "⇒": $\operatorname{epi} I_K = K \times [0, \infty)$.
"⇐": $K = \{x : I_K(x) \le 0\}$. □

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:

$$\operatorname{epi}(\min\{f, g\}) = \operatorname{epi} f \cup \operatorname{epi} g\,, \quad \operatorname{epi}(\sup_{i \in I} f_i) = \bigcap_{i \in I} \operatorname{epi} f_i\,. \tag{5.28}$$

□

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*

$$f(y) \le c\,, \quad \textit{für alle } y \in U\,. \tag{5.29}$$

Beweis: "⇒": Folgt direkt aus der Definition der Stetigkeit; wähle eine Umgebung U von x mit $f(U) \subset (f(x) - 1, f(x) + 1)$.
"⇐": Sei o.B.d.A. $x = 0$, $f(0) = 0$, sei $c \in \mathbb{R}$ und U Nullumgebung mit $f(U) \le c$. 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

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

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

$$0 = \frac{1}{1 + \varepsilon} y + \frac{\varepsilon}{1 + \varepsilon} \left(-\frac{y}{\varepsilon}\right)\,, \tag{5.31}$$

$$0 \le \frac{1}{1 + \varepsilon} f(y) + \frac{\varepsilon}{1 + \varepsilon} f\left(-\frac{y}{\varepsilon}\right) \le \frac{1}{1 + \varepsilon}(f(y) + \varepsilon c)\,, \tag{5.32}$$

also

$$|f(y)| \le \varepsilon c\,. \tag{5.33}$$

□

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* $\mathrm{int}\,(\mathrm{dom}\, f)$.

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

$$f(x) \le \sum_{i=0}^n \lambda_i f(x_i) \le \left(\sum_{i=0}^n \lambda_i \right) \max_{0 \le i \le n} f(x_i) = \max_{0 \le i \le n} f(x_i) \,. \tag{5.34}$$

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

$$f(S) \le \max_{0 \le i \le n} f(x_i) < \infty \,, \tag{5.35}$$

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* $0 \in \mathrm{int}\,(K)$. □

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

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

$$U = \{x : x \in X,\ f(x) \le c\} \,, \quad V = U \cap (-U) \,. \tag{5.36}$$

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 $0 \in \mathrm{aint}\,(V)$ und V absorbierend. Nach Satz 5.22 ist $0 \in \mathrm{int}\,(V)$, nach Satz 5.20 ist f stetig in 0. Für beliebiges $y \in \mathrm{int}\,(\mathrm{dom}\, f)$ 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*

$$f = \sup\{g |\ g : X \to \mathbb{R} \text{ affin und stetig}, \ g \le f\} \,. \tag{5.37}$$

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 $(x, a) \notin \mathrm{epi}\, f$, und $\mathrm{epi}\, f$ ist konvex und abgeschlossen. Aus dem Trennungssatz folgt, daß ein $z^* \in (X \times \mathbb{R})^*$ existiert mit

$$z^*(x, a) < \inf_{(\xi, \alpha) \in \mathrm{epi}\, f} z^*(\xi, \alpha) \,. \tag{5.38}$$

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

$$z^*(\xi, \alpha) = x^*(\xi) + \lambda\alpha\,, \quad \forall\, \xi \in X\,,\ \alpha \in \mathbb{R}\,. \tag{5.39}$$

Aus (5.38) erhalten wir also

$$x^*(x) + \lambda a = z^*(x,a) < x^*(\xi) + \lambda\mu\,, \quad \forall\, (\xi, \mu) \in \operatorname{epi} f\,. \tag{5.40}$$

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

$$g(\xi) = \frac{1}{\lambda}(z^*(x,a) - x^*(\xi))\,. \tag{5.41}$$

Aus (5.40) 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 $g \leq f$.
Fall 2: $\lambda = 0$. Dann ist $\xi = x$ in (5.40) nicht möglich, also $x \notin \operatorname{dom} f$ und $f(x) = \infty$. Sei $\tilde{g} : X \to \mathbb{R}$ affin und stetig mit $\tilde{g} \leq f$; 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

$$x^*(x) = z^*(x,a) < \beta < z^*(\operatorname{epi} f) = x^*(\operatorname{dom} f)\,, \tag{5.42}$$

und setzen

$$g(\xi) = \tilde{g}(\xi) + \delta(\beta - x^*(\xi))\,, \quad \delta > 0\,. \tag{5.43}$$

Aus (5.42) folgt $g(\xi) \leq \tilde{g}(\xi) \leq f(\xi)$ für $\xi \in \operatorname{dom} f$ und

$$g(x) = \tilde{g}(x) + \delta(\beta - x^*(x)) \geq a\,, \tag{5.44}$$

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 $f(x) \le c$ für alle $x \in X$. Dann ist

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

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

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

also

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

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

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

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

$$f \le g \quad \Rightarrow \quad f^* \ge g^*\,, \tag{6.9}$$

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

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

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

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*

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

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

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

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

$$f^{**} = \sup\{g | \ g : X \to \mathbb{R} \text{ affin und stetig, } g \le f^{**}\} . \tag{6.18}$$

Sei nun $g : X \to \mathbb{R}$ affin und stetig mit $g \le f$,

$$g(x) = \langle x^*, x\rangle - \alpha , \quad x^* \in X^* , \ \alpha \in \mathbb{R} . \tag{6.19}$$

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

$$\alpha \ge \langle x^*, x\rangle - f(x) \tag{6.20}$$

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

$$g(x) = \langle x^*, x\rangle - \alpha \le \langle x^*, x\rangle - f^*(x^*) \le f^{**}(x) . \tag{6.21}$$

Hieraus folgt die Behauptung. □

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

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

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.

Gemäß Definition 7.1 ist $\partial f(x) = \emptyset$ falls $f(x) = +\infty$ gilt. 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. $\square$

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

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

$\square$

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 $x \in \partial f^*(x^*)$. $\square$

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. $\square$

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*

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

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 $f(x) \in \mathbb{R}$. 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 monoton wachsende Funktion $d : (0, \infty) \to (-\infty, \infty]$ definiert. Ferner gilt $d(-t) \leq d(t)$ 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

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

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

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

Es gilt weiter für $t > 0$

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

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

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

□

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 $x \neq y$ 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. □

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.

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 Optimierungsprobleme

Problem 9.1
Sei X Vektorraum. Wir betrachten

$$\min_{x \in K} f(x)\,, \tag{9.1}$$

wobei $K \subset X$ konvex und $f : K \to \mathbb{R}$ konvex sind. Ein $x \in K$ heißt Lösung von (9.1), wenn $f(x) \leq f(y)$ für alle $y \in K$.

Satz 9.2 *Sei X Vektorraum, $K \subset X$ konvex, $f : K \to \mathbb{R}$ konvex. Dann ist die Menge aller Lösungen von Problem 9.1 konvex. Ist f strikt konvex, so gibt es höchstens eine Lösung. Ein $x \in K$ ist genau dann Lösung, wenn*

$$f'(x; y - x) \geq 0\,, \quad \textit{für alle } y \in K. \tag{9.2}$$

Beweis: Die Menge aller Lösungen ist gegeben durch die Menge

$$S = \{x : x \in K,\, f(x) \leq \inf_{y \in K} f(y)\}\,, \tag{9.3}$$

welche konvex ist, da f und K konvex sind. Ist f strikt konvex, so ist

$$f\left(\frac{x+y}{2}\right) < \frac{1}{2}f(x) + \frac{1}{2}f(y)\,, \tag{9.4}$$

falls $x, y \in K$ und $x \neq y$, also ist S höchstens einelementig. Die Charakterisierung durch (9.2) folgt direkt aus Satz 8.5. □

Satz 9.3 *Sei X Banachraum, $K \subset X$ konvex, $f : K \to \mathbb{R}$ konvex. Ist* $\mathrm{int}\,(K) \neq \emptyset$ *oder gibt es ein $\tilde{x} \in K$, so daß f stetig ist in $\tilde{x}$, so ist $x \in X$ Lösung von Problem 9.1 genau dann, wenn*

$$0 \in \partial f(x) + \partial I_K(x)\,. \tag{9.5}$$

Beweis: Ein $x \in X$ ist Lösung von 9.1 genau dann, wenn x Lösung ist von

$$\min_{x \in X} f(x) + I_K(x)\,, \tag{9.6}$$

was wiederum äquivalent ist mit $0 \in \partial(f + I_K)(x)$ nach Satz 7.5. Aus der Summenregel Satz 7.11 folgt die Behauptung. □

Problem 9.4
Sei X Banachraum, $f : X \to \mathbb{R}$ konvex und $g_i : X \to \mathbb{R}$ konvex, $1 \leq i \leq m$. Wir betrachten

$$\min f(x)\,, \tag{9.7}$$

unter den Nebenbedingungen

$$g_i(x) \leq 0\,, \quad 1 \leq i \leq m\,. \tag{9.8}$$

Wir betrachten die dem Problem 9.4 zugeordnete Lagrange-Funktion $L : X \times \mathbb{R}^m \to \mathbb{R}$, definiert durch

$$L(x, \lambda) = f(x) + \langle \lambda, g(x) \rangle = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) \,. \tag{9.9}$$

Satz 9.5 *Sei X Banachraum, $f : X \to \mathbb{R}$ und $g_i : X \to \mathbb{R}$ konvex, $1 \le i \le m$, es gebe ein $\tilde{x} \in X$ mit $g_i(\tilde{x}) < 0$ für $1 \le i \le m$ (Slater-Bedingung). Sei $x_* \in X$. Dann sind äquivalent*

(i) *x_* ist Lösung von Problem 9.4.*

(ii) *Es ist $g_i(x_*) \le 0$ für alle i, und es gibt $\lambda_i \ge 0$ mit $\lambda_i g_i(x_*) = 0$ für alle $1 \le i \le m$, so daß x_* Lösung ist von*

$$\min_{x \in X} L(x, \lambda) \,. \tag{9.10}$$

Beweis: "(ii)⇒(i)": Sei x_* Lösung von (9.10) für ein $\lambda \in \mathbb{R}^m$ mit $\lambda_i \ge 0$ für alle i. Dann gilt für alle $x \in X$, welche (9.8) erfüllen,

$$f(x) \ge f(x) + \langle \lambda, g(x) \rangle \ge f(x_*) + \langle \lambda, g(x_*) \rangle = f(x_*) \,. \tag{9.11}$$

"(ii)⇐(i)": Mit dem Trennungssatz, angewendet auf den Punkt 0 und die Menge $C \subset \mathbb{R}^{m+1}$,

$$C = \{(\mu_0, \ldots, \mu_m) : \text{ es gibt } x \text{ mit } f(x) - f(x_*) \le \mu_0,\ g_i(x) \le \mu_i \text{ für alle } i\} \,. \tag{9.12}$$

Einzelheiten siehe Übungsaufgabe. □

Satz 9.6 (Karush-Kuhn-Tucker)
Es seien die Voraussetzungen von Satz 9.5 erfüllt, außerdem gebe es ein $\hat{x} \in X$, so daß f und alle g_i stetig sind in $\hat{x}$. Dann sind äquivalent:

(i) *x_* ist Lösung von Problem 9.4.*

(ii) *Es ist $g_i(x_*) \le 0$ für alle i, und es gibt $\lambda \in \mathbb{R}^m$, $\lambda \ge 0$ mit $\lambda_i g_i(x_*) = 0$ für alle i und*

$$0 \in \partial f(x_*) + \sum_{i=1}^{m} \lambda_i \partial g_i(x_*) \,. \tag{9.13}$$

Beweis: Nach Satz 9.5 und Satz 7.5 ist (i) äquivalent zur Existenz von $\lambda \ge 0$ mit $\lambda_i g_i(x_*) = 0$ für alle i und

$$0 \in \partial_x L(x_*, \lambda) \,, \tag{9.14}$$

wobei $\partial_x L(x_*, \lambda) := \partial \tilde{L}(x_*)$ mit $\tilde{L}(x_*) = L(x_*, \lambda)$. Die Behauptung folgt nun aus der Summenregel 7.11. □

Satz 9.7
Es seien die Voraussetzungen von Satz 9.6 erfüllt, sei $x_* \in X$, seien f und alle g_i Gâteaux-differenzierbar in x_*. Dann sind äquivalent:

(i) *x_* ist Lösung von Problem 9.4.*

(ii) *Es ist $g_i(x_*) \leq 0$ für alle i, und es gibt $\lambda \in \mathbb{R}^m$, $\lambda \geq 0$ mit $\lambda_i g_i(x_*) = 0$ für alle i und*

$$0 = f'(x_*) + \sum_{i=1}^{m} \lambda_i g_i'(x_*)\,. \tag{9.15}$$

Beweis: Folgt aus Satz 9.6 und Satz 8.7. □

Die Charakterisierung des Minimums mit (9.15) wird auch als Multiplikatorenregel von Lagrange bezeichnet. Sind die beteiligten Funktionen differenzierbar, aber nicht mehr konvex, so gilt immer noch die Implikation von (i) nach (ii), aber nicht mehr die von (ii) nach (i), die Multiplikatorenregel ist nur noch eine notwendige Optimalitätsbedingung.

10 Dualität in der konvexen Optimierung

Wir gehen aus von dem Optimierungsproblem

$$\min_{x \in X} f(x) \,, \tag{10.1}$$

wobei X Banachraum und $f : X \to (-\infty, \infty]$. Problem (10.1) heißt das **primale Problem (P)**. Wir konstruieren auf folgende Weise ein sogenanntes duales Problem. Wir betten das primale Problem ein in eine Schar von Optimierungsproblemen

$$\min_{x \in X} F(x,p) \,, \quad p \in P \,, \tag{10.2}$$

wobei P Banachraum und $F : X \times P \to (-\infty, \infty]$. Dabei sei F so gewählt, daß gilt

$$F(x,0) = f(x) \,. \tag{10.3}$$

Wir betrachten nun das Problem

$$\max_{p^* \in P^*} -F^*(0,p^*) \,, \tag{10.4}$$

wobei $F^* : X^* \times P^* \to [-\infty, \infty]$ die zu F konjugierte Funktion ist. Problem (10.4) heißt das zu (P) **duale Problem (D)**. (Man müßte eigentlich genauer sagen "ein zu (10.1) duales Problem" oder "das zu (10.2) duale Problem".) Mit

$$V_P = \inf_{x \in X} f(x) = \inf_{x \in X} F(x,0) \,, \quad V_D = \sup_{p^* \in P^*} -F^*(0,p^*) \,, \tag{10.5}$$

bezeichnen wir den **Optimalwert** von Problem (P) bzw. (D). Durch

$$V(p) = \inf_{x \in X} F(x,p) \tag{10.6}$$

definieren wir eine Funktion $V : P \to [-\infty, \infty]$, die sogenannte **Optimalwertfunktion** der Schar (10.2). Nach Definition gilt

$$V_P = V(0) \,. \tag{10.7}$$

Satz 10.1 *Seien X, P Banachräume, $F : X \times P \to (-\infty, \infty]$. Dann gilt*

$$V_D \le V_P \,. \tag{10.8}$$

Für alle $x \in X$ und $p^* \in P^*$ sind äquivalent:

(i) *x löst* (P)*, p^* löst* (D)*, $V_D = V_P$,*

(ii) $F(x,0) + F^*(0,p^*) = 0$,

(iii) $(0,p^*) \in \partial F(x,0)$.

Beweis: Aus der Definition von F^* folgt für alle $x \in X$ und alle $p^* \in P^*$

$$F^*(0,p^*) \geq \langle (0,p^*),(x,0)\rangle - F(x,0) = \langle x,0\rangle + \langle 0,p^*\rangle - F(x,0) = -F(x,0)\,, \tag{10.9}$$

also $-F^*(0,p^*) \leq F(x,0)$ für alle $x \in X$, $p^* \in P^*$, also $V_D \leq V_P$. Zur Äquivalenz: (i) ist äquivalent zu

$$F(x,0) = V_P = V_D = -F^*(0,p^*)\,, \tag{10.10}$$

woraus (ii) folgt. Umgekehrt impliziert (ii), daß

$$V_D \geq -F^*(0,p^*) = F(x,0) \geq V_P\,, \tag{10.11}$$

und aus $V_D \leq V_P$ folgt (i). Die Äquivalenz von (ii) und (iii) ist in Satz 7.6 bewiesen worden. □

Bemerkung 10.2
Ist $V_D < V_P$, so spricht man von einer Dualitätslücke. Falls keine Dualitätslücke vorliegt, so gilt

$$-F^*(0,p^*) \leq V_D = V_P \leq F(x,0) \tag{10.12}$$

für alle $x \in X$ und alle $p^* \in P^*$. In jedem Fall liefert die Berechnung eines Paares $(F(x,0),F^*(0,p^*))$ eine Einschließung des Optimalwerts. Dieser Sachverhalt wird von numerischen Verfahren ausgenutzt, sowohl in der Linearen Optimierung als auch in der Kombinatorischen Optimierung. So kann etwa heuristischen Verfahren zur Lösung NP-vollständiger Probleme die Optimalität über die Bedingung (ii) verifiziert werden.

Man beachte, daß in Satz 10.1 nicht behauptet wird, daß Lösungen von (P) oder (D) existieren.

Lemma 10.3 *Seien X, P Banachräume, $F : X \times P \to (-\infty,\infty]$ konvex. Dann ist die Optimalwertfunktion $V : P \to [-\infty,\infty]$ konvex.*

Beweis: Analog zum Beweis von Lemma 5.11. □

Lemma 10.4 *Seien X, P Banachräume, $F : X \times P \to (-\infty,\infty]$. Dann gilt für alle $p^* \in P^*$:*

(i) $V^*(p^*) = F^*(0,p^*)$.

(ii) $V_D = V^{**}(0)$.

Beweis: Zu (i):

$$\begin{aligned} V^*(p^*) &= \sup_{p\in P}(\langle p^*,p\rangle - V(p)) = \sup_{p\in P}\sup_{x\in X}(\langle p^*,p\rangle - F(x,p)) \\ &= \sup_{\substack{p\in P\\ x\in X}}(\langle 0,x\rangle + \langle p^*,p\rangle - F(x,p)) = F^*(0,p^*)\,. \end{aligned} \tag{10.13}$$

Zu (ii):

$$V^{**}(0) = \sup_{p^*\in P^*}(\langle p^*,0\rangle - V^*(p^*)) = \sup_{p^*\in P^*} -F^*(0,p^*) = V_D\,. \tag{10.14}$$

□

Definition 10.5 (Stabilität) *Das primale Problem* (P) *heißt stabil, wenn* $\partial V(0) \neq \emptyset$.
□

Satz 10.6 *Seien* X, P *Banachräume, sei* $F : X \times P \to (-\infty, \infty]$. *Dann sind äquivalent:*

(i) (P) *ist stabil.*

(ii) (D) *ist lösbar und* $V_P = V_D$.

In diesem Fall gilt, dass

$$\partial V(0) = \{p^* : p^* \in P^*\,,\ p^* \text{ löst (D)}\}\,.$$

Beweis: "(ii)⇒(i)": Es ist $V_P = V_D \in \mathbb{R}$. Ist p^* Lösung von (D), so gilt

$$V(0) + V^*(p^*) = V_P + F^*(0, p^*) = V_P - V_D = 0\,, \tag{10.15}$$

also $p^* \in \partial V(0)$ nach Satz 7.6.
"(i)⇒(ii)": Da $\partial V(0) \neq \emptyset$, ist $V_P = V(0) \in \mathbb{R}$ und $V(0) = V^{**}(0)$ nach Satz 7.7, also

$$V_P = V(0) = V^{**}(0) = V_D \tag{10.16}$$

nach Lemma 10.4. Ist $p^* \in \partial V(0)$, so ist

$$0 = \langle p^*, 0\rangle = V(0) + V^*(p^*) = V_D + F^*(0, p^*)\,, \tag{10.17}$$

also ist p^* Lösung von (D). □

Man kann nun (D) als primales Problem auffassen und das dazu duale Problem untersuchen. Dadurch wird es möglich, in Satz 10.6 die Rollen von (P) und (D) zu vertauschen und die Lösbarkeit von (P) mit der Stabilität von (D) zu charakterisieren. Wir gehen aus von

$$\min_{p^* \in P^*} F^*(0, p^*)\,, \tag{10.18}$$

eingebettet in die Schar

$$\min_{p^* \in P^*} F^*(x^*, p^*)\,, \quad x^* \in X^*\,. \tag{10.19}$$

Wir bezeichnen mit W die zugehörige Optimalwertfunktion, also

$$W(x^*) = \min_{p^* \in P^*} F^*(x^*, p^*)\,, \quad W : X^* \to [-\infty, \infty]\,. \tag{10.20}$$

Das Dualproblem zu (10.19) ist im Raum X^{**} angesiedelt; wir wollen aber zum Ausgangsraum X zurück. Wir nehmen also an, daß X ein reflexiver Banachraum ist; wir können dann vermittels des kanonischen Isomorphismus

$$i : X \to X^{**}\,, \quad (i(x))(x^*) = \langle x^*, x\rangle\ , \tag{10.21}$$

die Räume X und X^{**} identifizieren. Als Dualproblem zu (10.19) betrachten wir

$$\max_{x \in X} -F^{**}(x, 0)\,. \tag{10.22}$$

Falls $F^{**} = F$, so stimmt (10.22) mit (P) überein.

Definition 10.7 (Stabilität) *Das duale Problem* (D) *heißt stabil, wenn* $\partial W(0) \neq \emptyset$. □

Satz 10.8 *Seien* X, P *reflexive Banachräume, sei* $F : X \times P \to (-\infty, \infty]$ *konvex, unterhalbstetig und eigentlich. Dann sind äquivalent:*

(i) (D) *ist stabil.*

(ii) (P) *ist lösbar und* $V_P = V_D$.

In diesem Fall gilt, dass

$$\partial W(0) = \{x : x \in X\,,\ x \text{ löst (P)}\}\,.$$

Beweis: Verläuft analog zum Beweis von Satz 10.6.
Es gilt $F^{**} = F$ nach Folgerung 6.8 und

$$\begin{aligned} W^*(x) &= \sup_{x^* \in X^*} (\langle x^*, x\rangle - W(x^*)) = \sup_{x^* \in X^*, p^* \in P^*} (\langle x^*, x\rangle - F^*(x^*, p^*)) = F^{**}(x,0) \\ &= F(x,0)\,. \end{aligned} \tag{10.23}$$

"(ii)⇒(i)": Sei $V_P = V_D$ und $x \in X$ Lösung von (P). Dann gilt nach (10.23)

$$W(0) + W^*(x) = -V_D + F(x,0) = -V_D + V_P = 0\,, \tag{10.24}$$

also $x \in \partial W(0)$ nach Satz 7.6.
"(i)⇒(ii)": Aus $\partial W(0) \neq \emptyset$ folgt $W(0) = -V_D \in \mathbb{R}$ und $W(0) = W^{**}(0)$ sowie nach (10.23)

$$W^{**}(0) = -\inf_{\xi \in X} W^*(\xi) = -\inf_{\xi \in X} F(\xi, 0) = -V_P\,, \tag{10.25}$$

also ist $V_D = V_P$. Ist $x \in \partial W(0)$, so gilt

$$0 = \langle 0, x\rangle = W(0) + W^*(x) = -V_D + F(x,0) = -V_P + F(x,0)\,, \tag{10.26}$$

also ist x Lösung von (P). □

Bemerkung 10.9
Anstatt für reflexive Banachräume kann man die Dualitätssätze auch für sogenannte duale Paare von lokalkonvexen Räumen herleiten. Sind X und Y lokalkonvex, so heißt (X, Y) duales Paar, wenn X^* isomorph zu Y und Y^* isomorph zu X ist. Für einen beliebigen lokalkonvexen Raum X läßt sich diese Situation immer herstellen, indem man $Y = X^*$ wählt, X mit der schwachen und X^* mit der schwach-$*$-Topologie versieht.

Satz 10.10 *Seien* X, P *Banachräume,* $F : X \times P \to (-\infty, \infty]$ *konvex, es gebe* $x_0 \in X$ *mit* $f(x_0) = F(x_0, 0) < \infty$*, es gebe* $p_0^* \in P^*$ *mit* $F^*(0, p_0^*) < \infty$*. Dann gilt:*

(i) *Ist die durch* $f_0(p) = F(x_0, p)$ *definierte Funktion stetig in* $p = 0$*, so ist* (P) *stabil.*

(ii) *Ist die durch* $g_0(x^*) = F^*(x^*, p_0^*)$ *definierte Funktion stetig in* $x^* = 0$*, so ist* (D) *stabil.*

Beweis: Zunächst gilt $-\infty < V_D \le V_P = V(0) < +\infty$ nach Voraussetzung an F und F^*.
Zu (i): V ist konvex nach Lemma 10.3. Da $V(p) \le f_0(p)$ und f_0 in $p = 0$ stetig ist, ist V auf einer Umgebung von 0 nach oben beschränkt. Aus Satz 5.20 und der Bemerkung im Anschluß daran folgt, daß V den Wert $-\infty$ nicht annimmt und in 0 stetig ist. Aus Satz 7.9 folgt, daß $\partial V(0)$ nichtleer ist, also ist (P) stabil.
Zu (ii): Analog. □

Beispiel 10.11
Wir betrachten für $X = \mathbb{R}$

$$\min x^2\,, \quad \text{Nebenbedingung } x \ge 1\,. \tag{10.27}$$

Problem (P) lautet dann

$$\min f(x)\,, \quad f(x) = x^2 + I_K(x)\,, \quad K = \{x : x \ge 1\}\,. \tag{10.28}$$

Wir setzen $P = \mathbb{R}$ und wählen

$$F(x,p) = x^2 + I_K(x-p)\,. \tag{10.29}$$

F ist konvex, unterhalbstetig und eigentlich, und es gilt

$$\begin{aligned} F^*(x^*,p^*) &= \sup_{x,p}(x^*x + p^*p - x^2 - I_K(x-p)) \\ &= \sup_x(x^*x - x^2 + \sup_{p\le x-1} p^*p)\,. \end{aligned} \tag{10.30}$$

Für $p^* \ge 0$ wird (10.30) zu

$$F^*(x^*,p^*) = \sup_x(x^*x - x^2 + p^*(x-1))\,. \tag{10.31}$$

Das Supremum wird angenommen, falls $x^* - 2x + p^* = 0$, also in

$$x = \frac{x^*+p^*}{2}\,, \tag{10.32}$$

so daß

$$F^*(x^*,p^*) = \frac{x^*+p^*}{2}\left(x^* + p^* - \frac{x^*+p^*}{2}\right) - p^*\,, \tag{10.33}$$

also insgesamt

$$F^*(x^*,p^*) = \begin{cases} \infty\,, & p^* < 0\,, \\ \frac{(x^*+p^*)^2}{4} - p^*\,, & p^* \ge 0\,. \end{cases} \tag{10.34}$$

Es wird $-F^*(0,p^*)$ maximal für $p^* = 2$, f minimal für $x = 1$, und es gilt $V_D = V_P = 1$.

Problem 10.12
Wir betrachten nun das primale Problem (P) in der Form

$$\min_{x\in X} f(x)\,, \quad f(x) = J(x) + h(Ax-b)\,, \tag{10.35}$$

wobei X, P Banachräume, $J : X \to (-\infty,\infty]$, $h : P \to (-\infty,\infty]$, $A : X \to P$ linear und stetig, $b \in P$ gegeben. Die Standardkonstruktion für F ist

$$F(x,p) = J(x) + h(Ax - b - p)\,. \tag{10.36}$$

Um das zu (10.35) duale Problem aufzustellen, müssen wir $F^*(x^*, p^*)$ berechnen. Es ist

$$\begin{aligned} F^*(x^*, p^*) &= \sup_{x\in X}\sup_{p\in P} [\langle x^*, x\rangle + \langle p^*, p\rangle - J(x) - h(Ax - b - p)] \\ &= \sup_{x\in X} \left[\langle x^*, x\rangle - J(x) + \sup_{p\in P} (\langle p^*, p\rangle - h(Ax - b - p)) \right] , \end{aligned} \tag{10.37}$$

und

$$\begin{aligned} &\sup_{p\in P} (\langle p^*, p\rangle - h(Ax - b - p)) = \sup_{q\in P} (\langle p^*, Ax - b - q\rangle - h(q)) \\ &\qquad = \langle p^*, Ax - b\rangle + \sup_{q\in P}(\langle -p^*, q\rangle - h(q)) = \langle p^*, Ax - b\rangle + h^*(-p^*) . \end{aligned} \tag{10.38}$$

Einsetzen von (10.38) in (10.37) ergibt

$$\begin{aligned} F^*(x^*, p^*) &= \sup_{x\in X} [\langle x^*, x\rangle - J(x) + \langle p^*, Ax - b\rangle + h^*(-p^*)] \\ &= h^*(-p^*) - \langle p^*, b\rangle + \sup_{x\in X} [\langle x^*, x\rangle + \langle p^*, Ax\rangle - J(x)] . \end{aligned} \tag{10.39}$$

Es ist weiter

$$\begin{aligned} \sup_{x\in X} [\langle x^*, x\rangle + \langle p^*, Ax\rangle - J(x)] &= \sup_{x\in X} [\langle x^* + A^*p^*, x\rangle - J(x)] \\ &= J^*(x^* + A^*p^*) , \end{aligned} \tag{10.40}$$

also insgesamt

$$F^*(x^*, p^*) = J^*(x^* + A^*p^*) + h^*(-p^*) - \langle p^*, b\rangle . \tag{10.41}$$

Das zu (10.35) duale Problem lautet also

$$\max_{p^*\in P^*} -J^*(A^*p^*) - h^*(-p^*) + \langle p^*, b\rangle . \tag{10.42}$$

Satz 10.13 *Seien X, P reflexive Banachräume, $J : X \to (-\infty, \infty]$ und $h : P \to (-\infty, \infty]$ konvex, unterhalbstetig und eigentlich, $A : X \to P$ linear und stetig, $b \in P$. Es gebe $x_0 \in X$ mit $J(x_0) < \infty$ und $h(Ax_0 - b) < \infty$, es gebe $p_0^* \in P^*$ mit $J^*(A^*p_0^*) < \infty$ und $h^*(-p_0^*) < \infty$. Dann gilt*

$$-\infty < V_D \le V_P < \infty . \tag{10.43}$$

Es sind äquivalent

(i) *x löst* (P), *p^* löst* (D) *und $V_D = V_P$.*

(ii) *$A^*p^* \in \partial J(x)$, $-p^* \in \partial h(Ax - b)$.*

Ist außerdem h stetig in $Ax_0 - b$, so ist (D) *lösbar und $V_D = V_P$; ist J^* stetig in $A^*p_0^*$, so ist* (P) *lösbar und $V_D = V_P$.*

Beweis: Die Aussage (10.43) folgt aus Satz 10.1 und den Eigenschaften von x_0 und p_0^*. Zur Äquivalenz: (i) ist wegen (10.43) äquivalent zu

$$J(x) + h(Ax - b) = F(x, 0) = -F^*(0, p^*) = -J^*(A^*p^*) - h^*(-p^*) + \langle p^*, b \rangle \ , \qquad (10.44)$$

was wiederum äquivalent ist zu

$$\Big[J(x) + J^*(A^*p^*) - \langle A^*p^*, x\rangle\Big] = -\Big[h(Ax - b) + h^*(-p^*) - \langle -p^*, Ax - b\rangle\Big] . \qquad (10.45)$$

Wegen Lemma 6.5 sind beide Klammern nichtnegativ, also ist (10.45) äquivalent zu

$$J(x) + J^*(A^*p^*) - \langle A^*p^*, x\rangle = 0 \, , \quad h(Ax - b) + h^*(-p^*) - \langle -p^*, Ax - b\rangle = 0 \, . \quad (10.46)$$

Wegen Satz 7.6 ist (10.46) äquivalent zu

$$A^*p^* \in \partial J(x) \, , \quad -p^* \in \partial h(Ax - b) \, . \qquad (10.47)$$

Die Zusatzannahmen an h bzw. J^* implizieren wegen Satz 10.10 die Stabilität von (P) bzw. (D). Aus den Sätzen 10.6 und 10.8 folgen die restlichen Behauptungen. □

Lineare Optimierungsprobleme liefern Spezialfälle von Problem (P). Sie lassen sich (im nichtganzzahligen Fall) folgendermaßen formulieren.

Problem 10.14 (Allgemeines lineares Optimierungsproblem)
Wir betrachten

$$\min \langle c, x\rangle \ , \quad \text{wobei } x \in K_X \subset X \, , \quad Ax - b \in K_P \subset P \, . \qquad (10.48)$$

Hierbei sind X, P Banachräume, $A : X \to P$ linear und stetig, $c \in X^*$, $b \in P$, K_X und K_P abgeschlossene konvexe Kegel.
Falls X und P endlichdimensional sind, spricht man vom finiten Fall; falls einer der beiden Räume unendlichdimensional ist, vom semiinfiniten Fall. □

Ein Beispiel im finiten Fall ist etwa $X = \mathbb{R}^n$, $P = \mathbb{R}^m$,

$$\min \sum_{i=1}^{n} c_i x_i \, , \quad \text{wobei } x \geq 0 \, , \quad Ax \leq b \, , \qquad (10.49)$$

wobei die Ungleichungen komponentenweise zu verstehen sind. Durch geeignete Wahl der konvexen Kegel K_X und K_P lassen sich auch kombinierte Gleichungs- und Ungleichungsnebenbedingungen erfassen.

Wir gewinnen das duale Problem über die Standardkonstruktion aus Problem 10.12. Problem 10.14 wird unter

$$\min_{x \in X} f(x) \, , \quad f(x) = J(x) + h(Ax - b) \, , \qquad (10.50)$$

subsumiert, indem wir setzen

$$J(x) = \langle c, x\rangle + I_{K_X}(x) \, , \quad h(p) = I_{K_P}(p) \, . \qquad (10.51)$$

Lemma 10.15 (Dualkegel) *Sei X Banachraum, $K \subset X$. Dann wird durch*

$$K^* = \{x^* : x^* \in X^*,\ \langle x^*, x\rangle \ge 0 \textit{ für alle } x \in K\} \tag{10.52}$$

ein abgeschlossener konvexer Kegel definiert. K^* heißt der Dualkegel von K.

Beweis:

$$K^* = \bigcap_{x\in X} K_x^*, \quad K_x^* = \{x^* : \langle x^*, x\rangle \ge 0\}. \tag{10.53}$$

□

Es ist

$$J^*(x^*) = \sup_{x\in X} \langle x^*, x\rangle - \langle c, x\rangle - I_{K_X}(x) = \sup_{x\in K_X} -\langle c - x^*, x\rangle = I_{K_X^*}(c - x^*), \tag{10.54}$$

$$h^*(p^*) = \sup_{p\in K_P} \langle p^*, p\rangle = I_{K_P^*}(-p^*). \tag{10.55}$$

Da das duale Problem zu (10.50) die Form

$$\max_{p^*\in P^*} -J^*(A^*p^*) - h^*(-p^*) + \langle p^*, b\rangle \tag{10.56}$$

hat, erhalten wir als duales Problem zu Problem 10.14

$$\max_{p^*\in P^*} \langle p^*, b\rangle , \quad \text{wobei } p^* \in K_P^*,\ c - A^*p^* \in K_X^*. \tag{10.57}$$

Satz 10.16 *Seien X, P reflexive Banachräume, $K_X \subset X$ und $K_P \subset P$ abgeschlossene konvexe Kegel, $A : X \to P$ linear und stetig, $b \in P$ und $c \in X^*$. Es gebe $x_0 \in X$ mit $x_0 \in K_X$ und $Ax_0 - b \in K_P$, es gebe $p_0^* \in P^*$ mit $p_0^* \in K_P^*$ und $c - A^*p_0^* \in K_X^*$. Dann gilt*

$$-\infty < V_D \le V_P < \infty. \tag{10.58}$$

Es sind äquivalent

(i) *x löst* (P)*, p^* löst* (D) *und $V_D = V_P$.*

(ii) *$x \in K_X$, $Ax - b \in K_P$, $p^* \in K_P^*$, $c - A^*p^* \in K_X^*$ (das heißt, x und p^* erfüllen die Nebenbedingungen), sowie*

$$\langle c - A^*p^*, x\rangle = 0, \quad \langle p^*, Ax - b\rangle = 0. \tag{10.59}$$

Ist außerdem $Ax_0 - b \in \operatorname{int}(K_P)$, so ist (D) *lösbar und $V_D = V_P$; ist $c - A^*p_0^* \in \operatorname{int}(K_X^*)$, so ist* (P) *lösbar und $V_D = V_P$.*

Beweis: Es handelt sich um eine unmittelbare Konkretisierung von Satz 10.13. Die Äquivalenz von (i) und (ii) folgt aus den Äquivalenzen

$$\begin{aligned} A^*p^* \in \partial J(x) \quad &\Leftrightarrow \quad J(x) + J^*(A^*p^*) = \langle A^*p^*, x\rangle \\ &\Leftrightarrow \quad \langle c, x\rangle + I_{K_X}(x) + I_{K_X^*}(c - A^*p^*) = \langle A^*p^*, x\rangle \\ &\Leftrightarrow \quad x \in K_X, \quad c - A^*p^* \in K_X^*, \quad \langle c - A^*p^*, x\rangle = 0, \end{aligned} \tag{10.60}$$

sowie

$$\begin{aligned} -p^* \in \partial h(Ax-b) \;&\Leftrightarrow\; h(Ax-b) + h^*(-p^*) = \langle -p^*, Ax-b\rangle \\ &\Leftrightarrow\; I_{K_P}(Ax-b) + I_{K_P^*}(p^*) = \langle -p^*, Ax-b\rangle \\ &\Leftrightarrow\; Ax-b \in K_P\,, \quad p^* \in K_P^*\,, \quad \langle p^*, Ax-b\rangle = 0\,. \qquad (10.61) \end{aligned}$$

□

Bemerkung 10.17 (Slater-Bedingung)
Die Bedingung

$$\text{es gibt ein } x_0 \text{ mit } Ax_0 - b \in \operatorname{int}(K_X) \qquad (10.62)$$

bzw. die entsprechende Bedingung für p_0^* heißt Slater-Bedingung. Laut Satz 10.16 ist sie hinreichend für "$V_D = V_P$" und die Lösbarkeit von (D) bzw. (P). Sie ist aber nicht in allen Fällen notwendig. So kann man z.B. zeigen, daß im finiten Fall eine Dualitätslücke nicht auftreten kann.

Beispiel 10.18
Wir betrachten das finite Problem

$$\min\,\langle c,x\rangle\ , \quad x \ge 0\,, \quad Ax = b\,, \qquad (10.63)$$

mit $X = \mathbb{R}^n$, $P = \mathbb{R}^m$, $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{(m,n)}$, $b \in \mathbb{R}^m$. Das Dualproblem lautet (Übung)

$$\max\,\langle b,y\rangle\ , \quad A^T y \le c\,. \qquad (10.64)$$

Für das Problem

$$\min\,\langle c,x\rangle\ , \quad x \ge 0\,, \quad Ax \ge b\,, \qquad (10.65)$$

erhalten wir als Dualproblem

$$\max\,\langle b,y\rangle\ , \quad y \ge 0\,, \quad A^T y \le c\,. \qquad (10.66)$$