

VERALLGEMEINERTE KONVEXE FUNKTIONEN

ERNST-AUGUST WEISS JR.

1. Einleitung.

X und Y seien zwei Vektorräume über \mathbb{R} , die ein Dualsystem bilden. Auf X und Y werden zulässige lokalkonvexe Topologien eingeführt. Einer reellwertigen Funktion f mit Definitionsbereich $D(f) \subset X$ ist nach Fenchel [4], Brøndsted [2], Dieter [3] die konjugierte Funktion f^\wedge zugeordnet. Sie ist definiert auf

$$D(f^\wedge) := \{y \in Y \mid \sup\{\langle x, y \rangle - f(x) \mid x \in D(f)\} < \infty\}$$

durch

$$f^\wedge(y) := \sup\{\langle x, y \rangle - f(x) \mid x \in D(f)\}.$$

Die Konstruktion von f^\wedge geschieht offenbar durch Hüllenbildung mit Hilfe der durch $y \mapsto \langle x, y \rangle - r$ gegebenen Hyperebenen, die jedem $(x, r) \in X \times \mathbb{R}$ zugeordnet sind. In Paragraph 2 können wir diese Konstruktion beträchtlich verallgemeinern: Wir setzen X und Y lediglich als topologische Räume voraus und ersetzen die Hyperebenen ganz allgemein durch stetige Funktionen $Y \rightarrow \mathbb{R}$, die jedem $(x, r) \in X \times \mathbb{R}$ zugeordnet sind, analog für $(y, s) \in Y \times \mathbb{R}$. Wir stellen an sie einige in diesem Zusammenhang recht selbstverständliche Forderungen. Dann gibt es $K: X \times Y \rightarrow \mathbb{R}$, so daß die Funktionen durch $y \mapsto K(x, y) - r$ bzw. $x \mapsto K(x, y) - s$ gegeben sind. Wir können dann die Konjugierte f^\wedge durch

$$f^\wedge(y) := \sup\{K(x, y) - f(x) \mid x \in D(f)\}$$

angeben. »Konvexe« Funktionen sind dann im Wesentlichen die Funktionen, die von den durch K gegebenen Funktionen eingehüllt werden.

In Paragraph 3 vergleichen wir Funktionen, die bezüglich verschiedener K konvex sind; speziell ergibt sich daraus, welche K dieselben Mengen von konvexen Funktionen definieren. Wir finden eine weitere Charakterisierung unserer konvexen Funktionen.

Erfüllt die Konjugierte der Funktion f spezielle Voraussetzungen, dann können wir in Paragraph 4 die verallgemeinerten Teile des Epigraphen

$$[f]_+ := \{(x, r) \in X \times \mathbb{R} \mid r \geq f(x)\}$$

betrachten.

Eingegangen am 17. Januar 1974, in revidierter Form am 6. Mai 1974.

In Paragraph 5 ergibt eine Anwendung unserer Theorie eine allgemeine » $\text{inf sup} = \text{sup inf}$ Gleichung« und einen Sattelpunkt-Satz für im üblichen Sinne nicht konvex-konkave Funktionen. Sätze dieser Art sind nützlich für viele Anwendungen in der Spieltheorie, bei denen an eine Vektorraumsstruktur garnicht zu denken ist.

Ich danke Herrn A. Brøndsted für seine hilfreichen Bemerkungen zu Satz 12 und der Deutschen Forschungsgemeinschaft für ihre finanzielle Unterstützung während des Entstehens dieser Arbeit.

2. Definition und Eigenschaften der konjugierten Funktion.

2.1. Im Folgenden sind X und Y zwei topologische Räume. Jedem Paar $(x, r) \in X \times \mathbb{R}$ sei die auf ganz Y stetige reellwertige Funktion $(x, r)^\wedge$ zugeordnet. Analog sei $(y, s) \in Y \times \mathbb{R}$ die auf ganz X stetige reellwertige Funktion $(y, s)^\wedge$ zugeordnet.

BEMERKUNG. X und Y können auch mit der diskreten Topologie versehen sein. Daher gilt alles Folgende auch für Mengen und reellwertige Funktionen ohne topologische Einschränkungen.

VERABREDUNG. Im Folgenden sollen alle Eigenschaften, die sich auf X beziehen, sinngemäß auch für Y gelten. Wir schreiben dies nicht gesondert auf.

Sei

$$F(X) := \{f \mid f: D(f) \rightarrow \mathbb{R}, \emptyset \neq D(f) \subset X\},$$

und für $f \in F(X)$ definieren wir

$$[f]_+ := \{(x, r) \in X \times \mathbb{R} \mid x \in D(f), r \geq f(x)\}.$$

DEFINITION 1. Sei $f \in F(X)$. Die (konvex) konjugierte Funktion f^\wedge von f wird definiert durch

$$\begin{aligned} D(f^\wedge) &:= \{y \in Y \mid \sup\{(x, r)^\wedge(y) \mid (x, r) \in [f]_+\} < \infty\}, \\ f^\wedge(y) &:= \sup\{(x, r)^\wedge(y) \mid (x, r) \in [f]_+\}. \end{aligned}$$

In Abschwächung der Eigenschaften der Hyperebenen in dualen topologischen Vektorräumen fordern wir von der obigen Zuordnung der Funktionen folgende Eigenschaften:

EIGENSCHAFT (E1). Sei $(x, r) \in X \times \mathbb{R}$, $y \in Y$ und $s := (x, r)^\wedge(y)$. Dann soll gelten $(y, s)^\wedge(x) = r$.

EIGENSCHAFT (E2). Sei $x \in X$ und $r_1, r_2 \in \mathbb{R}$. Dann soll gelten $r_1 - r_2 \leq (x, r_2)^\wedge - (x, r_1)^\wedge$.

SATZ 1. Für $(x, r_1), (x, r_2) \in X \times \mathbb{R}$ gilt

$$(x, r_2)^\wedge \equiv (x, r_1)^\wedge + (r_1 - r_2),$$

speziell

$$(x, r_2)^\wedge \equiv (x, 0)^\wedge - r_2.$$

BEWEIS. Es gebe ein $y_0 \in Y$, so daß mit $h_1 := (x, r_1)^\wedge$ und $h_2 := (x, r_2)^\wedge$ gilt

$$r_1 - r_2 < h_2(y_0) - h_1(y_0).$$

Nach (E2) findet man für $(y_0, h_1(y_0)), (y_0, h_2(y_0)) \in Y \times \mathbb{R}$ die Ungleichung

$$h_2(y_0) - h_1(y_0) \leq (y_0, h_1(y_0))^\wedge - (y_0, h_2(y_0))^\wedge.$$

Mit (E1) ergibt sich also

$$h_2(y_0) - h_1(y_0) \leq (y_0, h_1(y_0))^\wedge(x) - (y_0, h_2(y_0))^\wedge(x) = r_1 - r_2$$

im Widerspruch zur Annahme über y_0 . Mit (E1) folgt die Behauptung.

KOROLLAR. Für $f \in F(X)$ ist f^\wedge einfacher beschrieben durch

$$f^\wedge(y) = \sup \{ (x, 0)^\wedge(y) - f(x) \mid x \in D(f) \}.$$

BEWEIS. Für $(x, r) \in [f]_+$ hat man nach Satz 1

$$(x, r)^\wedge \equiv (x, 0)^\wedge - r \leq (x, 0)^\wedge - f(x).$$

SATZ 2. Für $(x_0, r_0) \in X \times \mathbb{R}$ gilt $(x_0, r_0)^\wedge \wedge (x_0) = r_0$.

BEWEIS. Benutzt man Satz 1, so schreibt sich (E1) mit den dortigen Bezeichnungen $(y, 0)^\wedge(x) = (x, 0)^\wedge(y)$. Nach dem Korollar ist

$$\begin{aligned} (x_0, r_0)^\wedge \wedge (x_0) &= \sup \{ (y, 0)^\wedge(x_0) - (x_0, r_0)^\wedge(y) \mid y \in Y \} \\ &= \sup \{ (y, 0)^\wedge(x_0) - (x_0, 0)^\wedge(y) + r_0 \mid y \in Y \} \\ &= r_0 \end{aligned}$$

nach obiger Bemerkung.

Speziell gilt

$$0 = (x_0, 0)^\wedge \wedge (x_0) = \sup \{ (y, 0)^\wedge(x_0) - (x_0, 0)^\wedge(y) \mid y \in Y \}.$$

Also

$$(y, 0)^\wedge(x) \leq (x, 0)^\wedge(y) \quad \text{für alle } x \in X \text{ und } y \in Y.$$

Analog

$$(y, 0)^\wedge(x) \geq (x, 0)^\wedge(y) \quad \text{für alle } x \in X \text{ und } y \in Y.$$

Wir definieren nun $K: X \times Y \rightarrow \mathbb{R}$ durch

$$K(x, y) := (y, 0)^\wedge(x) = (x, 0)^\wedge(y) \quad \text{für } x \in X, y \in Y.$$

Dann ist also sowohl für $f \in F(X)$ die Funktion f^\wedge durch K gegeben:

$$f^\wedge(y) = \sup\{K(x, y) - f(x) \mid x \in D(f)\},$$

als auch für $g \in F(Y)$ die Funktion g^\wedge durch K gegeben:

$$g^\wedge(x) = \sup\{K(x, y) - g(y) \mid y \in D(g)\}.$$

Mit dieser Funktion K hatten wir in [7] in einem spezielleren Fall eine Theorie verallgemeinerter konvexer Funktionen ohne tatsächliche Verwendung linearer Struktur begonnen, allerdings, ohne die hier vorliegende geometrische Begründung zu bringen. Moreau gibt in [6, 4c] nach Vorgabe der Funktion K die Definition der konjugierten Funktion f^\wedge von f in der wenige Zeilen weiter oben stehenden Form mit Hilfe von K .

Wir können offenbar auch statt (E1) und (E2) die Geltung der beiden Sätze fordern und kommen dann ebenfalls auf die Funktion K .

2.2 Für $f \in F(X)$ führen wir außer $[f]_+$ noch ein

$$[f]_- := \{(x, r) \mid x \in D(f), r \leq f(x)\}$$

und definieren auf $F(X)$ zwei Halbordnungen:

$$f_1 \leq_+ f_2 \Leftrightarrow [f_1]_+ \supseteq [f_2]_+,$$

$$f_1 \leq_- f_2 \Leftrightarrow [f_1]_- \supseteq [f_2]_-.$$

DEFINITION 2. Die (konkav) konjugierte Funktion f^\sim von $f \in F(X)$ wird definiert durch

$$D(f^\sim) := \{y \in Y \mid \inf\{-K(x, y) - f(x) \mid x \in D(f)\} > -\infty\},$$

$$f^\sim(y) := \inf\{-K(x, y) - f(x) \mid x \in D(f)\}.$$

BEMERKUNG. In Zusammenhang mit Optimierungsaufgaben, wo neben f^\wedge auch f^\sim benötigt wird, definiert man sonst die Funktion f_s^\sim durch

$$f_s^\sim(y) := \inf\{\langle x, y \rangle - f(x) \mid x \in D(f)\}$$

$$= \inf\{\langle -x, -y \rangle - f(x) \mid x \in D(f)\}$$

$$= f^\sim(-y).$$

Wir können nun analog zum Fenchelschen Fall (siehe Paragraph 1) folgende beiden Sätze beweisen:

Eine (untere) Schranke von $f \in F(X)$ ist eine »Hyperebene« $x \mapsto K(x, y) - s$ mit $[f]_+ \subset [x \mapsto K(x, y) - s]_+$. Es gilt:

SATZ 3. Die Punkte von $[f^\wedge]_+$ und die unteren Schranken von $f \in F(X)$ entsprechen einander eineindeutig. $(y, s) \in [f^\wedge]_+$ ist dabei die Schranke $x \mapsto K(x, y) - s$ zugeordnet.

Speziell ist $D(f^\wedge)$ genau dann leer, wenn f keine Schranke besitzt.

BEWEIS. Sei $(y, s) \in [f^\wedge]_+$. Dann gilt für alle $(x, r) \in [f]_+$

$$s \geq f^\wedge(y) \geq K(x, y) - r,$$

d.h. $x \mapsto K(x, y) - s$ ist Schranke von f .

Umgekehrt sei $x \mapsto K(x, y) - s$ eine gegebene Schranke von f , d.h. für alle $(x, r) \in [f]_+$ sei $r \geq K(x, y) - s$ oder $s \geq K(x, y) - r$. Dann ist aber $y \in D(f^\wedge)$ und $s \geq f^\wedge(y)$, also $(y, s) \in [f^\wedge]_+$.

SATZ 4. Sei $f \in F(X)$ mit $D(f^\wedge) \neq \emptyset$.

Dann gilt $f^{\wedge\wedge} \leq_+ f$ und $f^{\wedge\wedge\wedge} = f^\wedge$, analog für \leq_- und \sim .

BEWEIS. Sei $x \in D(f)$, dann gilt für alle $y \in D(f^\wedge)$, daß $f(x) \geq K(x, y) - f^\wedge(y)$, woraus folgt $x \in D(f^{\wedge\wedge})$ und $f(x) \geq f^{\wedge\wedge}(x)$, daher ist $D(f^{\wedge\wedge}) \neq \emptyset$ und $f^{\wedge\wedge} \leq_+ f$. Indem man dies auf f^\wedge anwendet, folgt $f^{\wedge\wedge\wedge} \leq_+ f^\wedge$.

Andererseits folgt aber einfach aus $f^{\wedge\wedge} \leq_+ f$, daß $f^\wedge \leq_+ f^{\wedge\wedge\wedge}$.

DEFINITION 3. \bar{f} (\underline{f}) heißt oberer (bzw. unterer) Abschluß von $f \in F(X)$ genau dann, wenn:

Der Abschluß von $[f]_+$ (bzw. $[f]_-$) in $X \times \mathbb{R}$ ist $[\bar{f}]_+$ (bzw. $[\underline{f}]_-$).

Ist $f = \bar{f}$ ($f = \underline{f}$), so heißt f von oben (bzw. von unten) abgeschlossen.

Offenbar braucht nicht $\bar{f} \in F(X)$ bzw. $\underline{f} \in F(X)$ zu gelten. Bei dem späteren Gebrauch etwa des Ausdrucks \bar{f}^\wedge ist $\bar{f} \in F(X)$ vorausgesetzt.

Es gilt $f^\wedge = \bar{f}^\wedge$, da f und \bar{f} dieselben (stetigen) Schranken $x \mapsto K(x, y) - s$ besitzen.

Aus Satz 3 folgt die Beziehung

$$[f^\wedge]_+ = \bigcap_{(x,r) \in [f]_+} [y \mapsto K(x, y) - r]_+,$$

$[f^\wedge]_+$ ist also als Durchschnitt abgeschlossener Mengen abgeschlossen.

f^\wedge , und ebenso $f^{\wedge\wedge}$ sind also von oben abgeschlossen, speziell unterhalb stetig.

Nun können wir unsere wichtige Definition geben:

DEFINITION 4. $f \in F(X)$ ist *konvex* $:\Leftrightarrow$ der Abschluß \bar{f} hat die Eigenschaften $f = \bar{f}|D(f)$ und $\bar{f} = \bar{f}^{\wedge\wedge}$.

$f \in F(X)$ ist *konkav* $:\Leftrightarrow$ der Abschluß \underline{f} hat die Eigenschaften $f = \underline{f}|D(f)$ und $f = \underline{f}^{\sim\sim}$.

BEMERKUNG. Die hier gegebene Definition fordert im Fall dualer topologischer Vektorräume X, Y mit Bilinearform gegeben durch $K(x, y) := \langle x, y \rangle$ einerseits weniger als die übliche Definition, denn $D(f)$ braucht Punkte nicht zu enthalten, die den Abschluß \bar{f} bzw. \underline{f} nicht beeinflussen. Andererseits fordert sie aber mehr, da z. B. im konvexen Fall Unstetigkeiten nach unten nicht zugelassen sind, d.h. Punkte $x \in D(f)$ mit $f(x) \neq \bar{f}(x)$. Die vorliegende Definition 4 ist aber natürlich in diesem Zusammenhang, wenn wir an das Ergebnis von Fenchel [4] denken: f konvex und abgeschlossen genau dann, wenn $f = f^{\wedge\wedge}$.

SATZ 5. f ist konkav genau dann, wenn $-f$ konvex ist.

BEWEIS. Man zeigt die Gleichung $-f^{\sim} = (-f)^\wedge$ und folgert $f^{\sim\sim} = -(-f)^{\wedge\wedge}$. Daher gilt $\underline{f} = \underline{f}^{\sim\sim} = f^{\sim\sim}$ genau dann, wenn $\overline{-f} = (-f)^{\wedge\wedge} = (-f)^{\wedge\wedge}$. Man berücksichtigt noch $-f = \overline{-f}$.

3. Vergleich und Charakterisierung konvexer Funktionen.

In diesem Paragraphen legen wir die topologischen Räume X und Y_1, Y_2 zugrunde, sowie die in jeder Variablen partiell stetigen Funktionen $K_1: X \times Y_1 \rightarrow \mathbb{R}$ und $K_2: X \times Y_2 \rightarrow \mathbb{R}$.

Wir bezeichnen für $i = 1, 2$

$$E(X, Y_i, K_i) := \{h \in F(X) \mid \exists (y, s) \in Y_i \times \mathbb{R} : h = K_i(\cdot, y) - s\}.$$

3.1. Offenbar läßt sich eine Funktion $f \in F(X)$ genau dann zu einer K_i -konvexen Funktion erweitern, wenn:

Für jedes $x \in D(f)$ und jedes $\varepsilon > 0$ gibt es $h = K_i(\cdot, y) - s \in E(X, Y_i, K_i)$ mit $(y, s) \in [f^\wedge]_+$, so daß $h|D(f) \leq f$ und $f(x) - h(x) < \varepsilon$.

Im folgenden Spezialfall können wir mehr zeigen:

SATZ 6. Ist $f \in F(X)$ unterhalb stetig auf dem abgeschlossenen Definitionsbereich $D(f)$, dann ist f zu einer K_i -konvexen Funktion erweiterbar genau dann, wenn gilt:

Für jede kompakte Teilmenge $T \subset X$ und jedes $\varepsilon > 0$ gibt es endlich viele Funktionen $h_1, \dots, h_n \in E(X, Y_i, K_i)$, so daß

$$h_j | D(f) \leq f, \quad j = 1, \dots, n,$$

und

$$f - \sup\{h_1, \dots, h_n\} < \varepsilon \quad \text{auf } T \cap D(f).$$

BEWEIS. Wir verallgemeinern Satz 7 aus [1] auf die hier vorliegende allgemeinere Situation.

Sei zunächst $f, T \subset X$ und ε wie im Satz gegeben. Zu jedem $x \in T \cap D(f)$ gibt es $h_x \in E(X, Y_i, K_i)$ mit $h_x | D(f) \leq f$ und $f(x) - h_x(x) < \varepsilon$. Da die Funktionen $f - h_x | D(f)$ unterhalb stetig sind, gibt es zu jedem $x \in T \cap D(f)$ eine in $D(f)$ offene Umgebung U_x mit $f - h_x | D(f) < \varepsilon$ auf U_x . Da $T \cap D(f)$ kompakt ist, gibt es endlich viele $x_1, \dots, x_n \in X$, so daß

$$T \cap D(f) \subset U_{x_1} \cap \dots \cap U_{x_n}.$$

Die Funktionen $h_j := h_{x_j}$ genügen dann unserer Behauptung.

Die Umkehrung gilt offenbar, weil einpunktige Mengen kompakt sind.

Wir wollen nun Funktionen aus $F(X)$ vergleichen, die bezüglich K_1 konvex (kurz: K_1 -konvex) sind, und bezüglich K_2 konvex sind.

SATZ 7. Die folgende Aussagen sind äquivalent:

- (a) Alle $f \in F(X)$, die K_1 -konvex sind, sind auch K_2 -konvex.
- (b) Ist $h \in E(X, Y_1, K_1)$, so ist h K_2 -konvex.

BEWEIS. (a) \Rightarrow (b): Jedes $h \in E(X, Y_1, K_1)$ ist K_1 -konvex.

(b) \Rightarrow (a): Sei $f \in F(X)$ K_1 -konvex. Zu $x \in D(\bar{f})$ gibt es $h_x \in E(X, Y_1, K_1)$, so daß $h_x | D(\bar{f}) \leq \bar{f}$ und $\bar{f}(x) - h_x(x) < \frac{1}{2}\varepsilon$. Da h_x nach (b) auch K_2 -konvex ist, gibt es $h_{x'} \in E(X, Y_2, K_2)$, so daß $h_{x'} \leq h_x$ und $h_{x'}(x) - h_{x'}(x) < \frac{1}{2}\varepsilon$. Zusammen gibt es also zu jedem $x \in D(\bar{f})$ ein $h_{x'} \in E(X, Y_2, K_2)$ mit $h_{x'} | D(\bar{f}) \leq \bar{f}$ und $\bar{f}(x) - h_{x'}(x) < \varepsilon$. Damit ist $\bar{f} = \bar{f}^{\wedge \wedge}$ bezüglich K_2 .

Durch Wahl von $K_2(x, y) := \langle x, y \rangle$ im Fall dualer topologischer Vektorräume gewinnt man:

KOROLLAR. Die K_1 -konvexen $f \in F(X)$ lassen sich durch Abschluß zu im üblichen Sinn konvexen Funktionen fortsetzen genau dann, wenn die $h \in E(X, Y_1, K_1)$ im üblichen Sinn konvex sind.

3.2. Wir gewinnen weitere Aussagen durch folgende »beidseitige« Voraussetzung über unsere beiden » K -Systeme«, die in 3.2. gelten soll:

FORDERUNG (F1). Für $i = 1, 2$ soll gelten:

Für alle $(y_i, s_i) \in Y_i \times \mathbb{R}$ gibt es $(y_i', s_i') \in Y_i \times \mathbb{R}$, so daß

$$-K_i(\cdot, y_i) - s_i = K_i(\cdot, y_i') - s_i',$$

d.h. es ist $E(X, Y_i, K_i) = -E(X, Y_i, K_i)$.

Die Forderung ist z. B. zu erfüllen durch die folgende schärfere Forderung: Es gibt Involutionen $I_{Y_i}: Y_i \rightarrow Y_i$ mit

$$-K_i(\cdot, y_i) = K_i(\cdot, I_{Y_i}y_i) \quad \text{für alle } y_i \in Y_i.$$

DEFINITION 5. (a) Sei $f: X \rightarrow \mathbb{R}$. f heißt K_i -beschränkt; $f \in B(X, Y_i, K_i)$; genau dann, wenn es $h_1, h_2 \in E(X, Y_i, K_i)$ gibt, so daß $h_1 \leq f \leq h_2$.

(b) Die Elemente von

$$B^\wedge(X, Y_i, K_i) := \{f \in B(X, Y_i, K_i) \mid f = f^{\wedge\wedge} = f^{\sim\sim} \text{ bzgl. } K_i\}$$

heißen $E(X, Y_i, K_i)$ -affine Funktionen.

Wegen (F1) können wir nun alle $E(X, Y_i, K_i)$ -affinen Funktionen folgendermaßen charakterisieren:

SATZ 8. $f: X \rightarrow \mathbb{R}$ ist $E(X, Y_i, K_i)$ -affin genau dann, wenn: Für jede kompakte Teilmenge $T \subset X$ und jedes $\varepsilon > 0$ gilt: Es gibt endlich viele Funktionen h_1, \dots, h_n und $h_1', \dots, h_n' \in E(X, Y_i, K_i)$, so daß $h_j \leq f \leq h_j'$ für $j = 1, \dots, n$ und

$$\inf\{h_1', \dots, h_n'\} - \sup\{h_1, \dots, h_n\} < \varepsilon \quad \text{auf } T.$$

BEWEIS, (vgl. [1, Satz 4]). Statt der Funktionen $f - h_x \mid D(f)$ im Beweis zu Satz 6 hat man die stetigen Funktionen $h_x' - h_x$ zu betrachten.

SATZ 9. Die folgenden Aussagen sind äquivalent:

(a) Alle $f \in F(X)$, die K_1 -konvex sind, sind auch K_2 -konvex.

(b) $B^\wedge(X, Y_1, K_1) \subset B^\wedge(X, Y_2, K_2)$.

BEWEIS. (a) \Rightarrow (b). Sei $f \in B^\wedge(X, Y_1, K_1)$. Da f K_1 -konvex ist, gilt $f = f^{\wedge\wedge}$ bzgl. K_2 . Da f K_1 -konkav ist, ist auch $f = f^{\sim\sim}$ bzgl. K_2 . Also ist $B^\wedge(X, Y_1, K_1) \subset B^\wedge(X, Y_2, K_2)$.

(b) \Rightarrow (a). Es genügt offenbar zu zeigen: f ist K_i -konvex genau dann, wenn $f = \bar{f} \mid D(f)$ und

$$\bar{f} = \sup\{h \in B^\wedge(X, Y_i, K_i) \mid h \mid D(\bar{f}) \leq \bar{f}\}.$$

Dies folgt durch ein Argument wie in Satz 7, da $E(X, Y_i, K_i) \subset B^\wedge(X, Y_i, K_i)$ wegen (F1).

Es ist nun interessant, unter welcher Bedingung die Funktionen aus $E(X, Y_i, K_i)$ die einzigen auf X definierten Funktionen sind, die gleichzeitig konvex und konkav sind.

Es ist

$$E(X, Y_i, K_i) = B(X, Y_i, K_i) \quad \text{und daher} \quad E(X, Y_i, K_i) = B^\wedge(X, Y_i, K_i),$$

falls für alle $y_i, y_i' \in Y_i$ mit $y_i \neq y_i'$ gilt

$$\sup\{K_i(\cdot, y_i) - K_i(\cdot, y_i')\} = \infty.$$

Diese Bedingung garantierte uns bereits in [7], daß Funktionen mit einpunktigem Definitionsbereich enthalten in Y_i in unserer Theorie konvex und abgeschlossen sind.

Ist X, Y_1 ein Paar dualer topologischer Vektorräume mit $K_1 = \langle \cdot, \cdot \rangle$, so ist daher $B^\wedge(X, Y_1, K_1) = E(X, Y_1, K_1)$.

4. Die Teile des Epigraphen einer konvexen und abgeschlossenen Funktion.

Zur Vorbereitung der Paragraphen 4 und 5 bemerken wir zunächst, daß wir im Fall der konvexen Konjugation immer auf dem ganzen Raum, etwa X , definierte Funktionen betrachten können, indem wir den Funktionen in den bisher nicht definierten Punkten den Wert $+\infty$ geben; im Fall der konkaven Konjugation analog den Wert $-\infty$.

Wir kommen also ganz natürlich dazu, Funktionen aus

$$\bar{F}(X) := \{f \mid f: X \rightarrow \langle -\infty, \infty \rangle = \bar{\mathbb{R}}\}$$

zu betrachten. Nimmt $f \in \bar{F}(X)$ an einer Stelle den Wert $-\infty$ an, so ist $f^\wedge \equiv +\infty$. Ist $f \equiv +\infty$, so ist $f^\wedge \equiv -\infty$. Wir betrachten demnach die Funktion, die konstant gleich $+\infty$, bzw., die konstant gleich $-\infty$ ist, als konvex und abgeschlossen; und ebenso auch als konkav und abgeschlossen.

DEFINITION 6. Sei $f \in \bar{F}(X)$ und $x_0 \in X$.

$y_0 \in Y$ heißt *Subgradient von f in x_0* , falls $f(x_0) \in \mathbb{R}$ und

$$f \geq K(\cdot, y_0) - K(x_0, y_0) + f(x_0).$$

$\partial f(x_0)$ sei die Menge dieser Subgradienten in x_0 .

Speziell kann also eine Funktion, die einen Subgradienten besitzt, nicht den Wert $-\infty$ annehmen.

In dem restlichen Teil dieses Paragraphen setzen wir X und Y als hausdorffsch voraus.

DEFINITION 7. $g: Y \rightarrow \mathbb{R} \cup \{+\infty\}$ heißt *inf-kompakt*, falls für alle $s \in \mathbb{R}$ die Menge $\{y \in Y \mid g(y) \leq s\}$ kompakt ist.

g heißt *inf-kompakt in Richtung* $x \in X$, falls $g - K(x, \cdot)$ inf-kompakt ist.

SATZ 10. Wenn $g: Y \rightarrow \mathbb{R} \cup \{+\infty\}$ inf-kompakt in Richtung $x_0 \in X$ ist, dann gibt es zu der maximalen Schranke $K(x_0, \cdot) - \tau$ von g ein $y_0 \in Y$, so daß $K(x_0, y_0) - \tau = g(y_0)$.

BEWEIS. Nach Voraussetzung ist

$$\inf \{g(y) - K(x_0, y) + \tau \mid y \in Y\} = 0,$$

und in der Klammer ist $g - K(x_0, \cdot) + \tau =: h$ eine inf-kompakte Funktion. Nun zeigt man, daß h ihr Minimum annimmt, indem man in einer nicht leeren, kompakten Menge $\{y \in Y \mid h(y) \leq s_0\}$ die abgeschlossenen Mengen $\{y \in Y \mid h(y) \leq s\}$ mit $s \leq s_0$ betrachtet.

SATZ 11. Sei $f: X \rightarrow \mathbb{R} \cup \{+\infty\}$ eine Funktion mit $f = f^{\wedge\wedge}$ und f^{\wedge} inf-kompakt in allen Richtungen. Dann gibt es zu jedem $x_0 \in X$ eine Schranke $K(\cdot, y_0) - s$ von f mit $K(x_0, y_0) - s = f(x_0)$.

BEWEIS. Eine Funktion $h: X \rightarrow \mathbb{R} \cup \{+\infty\}$ besitzt in x_0 den Subgradienten y_0 genau dann, wenn $h(x_0) \in \mathbb{R}$ und

$$K(x_0, y_0) - h(x_0) \geq \sup \{K(x, y_0) - h(x) \mid x \in X\},$$

d.h. $h(x_0) + h^{\wedge}(y_0) - K(x_0, y_0) = 0$.

In unserem Fall hat $f^{\wedge}: Y \rightarrow \mathbb{R} \cup \{+\infty\}$ die Eigenschaft

$$\text{dom } f^{\wedge} := \{y \in Y \mid f^{\wedge}(y) \in \mathbb{R}\} \neq \emptyset.$$

Daher gibt es zu jedem $x_0 \in X$ eine maximale Schranke $K(x_0, \cdot) - \tau$ von f^{\wedge} . Nach Satz 10 schreibt sich diese in der Form $K(x_0, \cdot) - K(x_0, y_0) + f^{\wedge}(y_0)$ mit $y_0 \in \text{dom } f^{\wedge}$; d.h. f^{\wedge} besitzt den Subgradienten x_0 in y_0 . Nach unserer Bemerkung zu Beginn des Beweises gilt also: Zu jedem $x_0 \in X$ gibt es $y_0 \in Y$, so daß $f^{\wedge}(y_0) + f^{\wedge\wedge}(x_0) - K(x_0, y_0) = 0$. Wegen $f(x_0) = f^{\wedge\wedge}(x_0)$ und unserer anfänglichen Bemerkung folgt die Behauptung.

KOROLLAR. Wenn f die Eigenschaften wie in Satz 11 hat, dann ist $f(X) \subset \mathbb{R}$.

BEISPIEL. Ist Y in Satz 11 kompakt, so ist $f^{\wedge}: Y \rightarrow \mathbb{R} \cup \{+\infty\}$ inf-kompakt in allen Richtungen, da f^{\wedge} unterhalb stetig ist.

Speziell: X', Y' bilde ein Dualsystem, wähle $X := X'$ und Y kompakt in Y' .

DEFINITION 8. Für zwei Punkte $(x_1, r_1), (x_2, r_2) \in [f]_+$ sei $(x_1, r_1) \sim_f (x_2, r_2)$ genau dann, wenn gilt:

Ist $h = K(\cdot, y_0) - s$ eine Schranke von f , so ist $h(x_1) = r_1$ genau dann, wenn $h(x_2) = r_2$ ist.

BEMERKUNG. Jeder Punkt $(x, r) \in [f]_+$ mit $r = f(x)$ besitzt bei Gültigkeit der Voraussetzungen von Satz 11 eine solche Schranke, ein Punkt $(x, r) \in [f]_+$ mit $r > f(x)$ jedoch nicht. Daher bilden die Punkte von $\{(x, r) \mid f(x) < r\}$ eine Äquivalenzklasse. Für die Punkte des Graphen von f gilt

$$(x_1, f(x_1)) \sim_f (x_2, f(x_2)) \text{ genau dann, wenn } \partial f(x_1) = \partial f(x_2).$$

DEFINITION 9. f habe die Eigenschaften wie in Satz 11.

a) $p(x, r)$ sei die Äquivalenzklasse, die $(x, r) \in [f]_+$ enthält.

b) Es gelte $p(x_1, f(x_1)) < p(x_2, f(x_2))$ genau dann, wenn $\partial f(x_1) \supset \partial f(x_2)$ und $\partial f(x_1) \neq \partial f(x_2)$.

c) Die minimalen Elemente in der Halbordnung heißen Extremalmengen.

Offenbar gilt $\partial f(x_2) \subset \cap \partial f(x_1)$, wobei der Durchschnitt über alle x_1 gebildet wird mit $p(x_1, f(x_1)) < p(x_2, f(x_2))$.

Die Extremalmengen brauchen nicht einelementig zu sein. Als Beispiel sei nämlich $Y = \{y\}$ einelementig und X nicht einelementig. Dann bildet $\{(x, f(x)) \mid x \in X\}$ eine Äquivalenzklasse, die Extremalmenge ist.

Es folgt ein Satz vom Typ Krein–Milman:

SATZ 12. Sei $M \subset X$ und $f: X \rightarrow \mathbb{R} \cup \{+\infty\}$ eine Funktion mit $f = f^{\wedge\wedge}$. f und f^{\wedge} seien inf-kompakt in allen Richtungen.

Es gilt $f = (f|_M)^{\wedge\wedge}$, wenn die Projektion jeder Extremalmenge auf X mit dem Abschluß \bar{M} von M in X einen nicht leeren Durchschnitt hat.

(Wir erlauben uns, $f|_M: X \rightarrow \mathbb{R} \cup \{+\infty\}$ gleich f auf M und $+\infty$ sonst zu setzen.)

BEWEIS. Nach Satz 3 gilt $f^{\wedge} = (f|_M)^{\wedge}$, falls f und $f|_M$ dieselben Schranken der Form $K(\cdot, y_0) - s$ haben. Dies ist der Fall:

Sei $K(\cdot, y_0) - s$ nicht Schranke von f . Nach Satz 10 (f ist inf-kompakt in Richtung y_0) gibt es $x_1 \in X$ mit $y_0 \in \partial f(x_1)$. Da $K(\cdot, y_0) - s$ nicht Schranke ist, muß $f(x_1) < K(x_1, y_0) - s$ gelten.

Die Mengen $q(x, f(x)) := \cup p(x', f(x'))$ (wobei die Vereinigung über alle $x' \in X$ genommen wird mit $p(x', f(x')) \leq p(x, f(x))$) sind kompakt und abgeschlossen wegen der inf-Kompaktheit von f in allen Richtungen. Man definiert $q(x'_1, f(x'_1)) < q(x'_2, f(x'_2))$ durch $p(x'_1, f(x'_1)) < p(x'_2, f(x'_2))$.

Satz 11 und Definition 9 liefern daher mit Hilfe des Zornschen Lemmas ein minimales Element $q(x_2, f(x_2))$ mit $q(x_1, f(x_1)) \geq q(x_2, f(x_2))$ und damit die Extremalmenge $p(x_2, f(x_2))$ mit $p(x_1, f(x_1)) \geq p(x_2, f(x_2))$, und für jedes $(x, f(x)) \in p(x_2, f(x_2))$ ist $y_0 \in \partial f(x)$ und $f(x) < K(x, y_0) - s$. Mit der Voraussetzung über \bar{M} folgt, daß $K(\cdot, y_0) - s$ nicht Schranke von $f|_{\bar{M}}$ ist, d.h., daß

$$\inf \{f(x) - K(x, y_0) + s \mid x \in \bar{M}\} < 0.$$

Da $f - K(\cdot, y_0) + s$ unterhalb stetig ist, gilt auch

$$\inf \{f(x) - K(x, y_0) + s \mid x \in M\} < 0,$$

d.h. $K(\cdot, y_0) - s$ ist nicht Schranke von $f|M$.

Herrn A. Brøndsted verdanke ich folgendes Beispiel, welches zeigt, daß die Bedingung des Satzes 12 nicht notwendig ist:

BEISPIEL. Sei X das abgeschlossene Intervall $[-1, 1]$, und das zwei-punktige $Y := \{y_1, y_2\}$ sei mit der diskreten Topologie versehen. K sei gegeben durch $K(x, y_1) := -x$ und $K(x, y_2) := x$ für $x \in X$. Die Funktion f mit $f(x) := |x|$ ist, ebenso wie f^\wedge , inf-kompakt in allen Richtungen. Mit $M := \{-1, 1\}$ gilt $f = (f|M)^\wedge^\wedge$, aber die Projektion $\{0\}$ der Extremalmenge $\{(0, 0)\}$ schneidet $M = \bar{M}$ nicht.

BEMERKUNG. Sind X und Y topologische Vektorräume, die ein Dualsystem bilden, dann ist die vorliegende Einteilung in Äquivalenzklassen die bekannte Einteilung in Teile der konvexen Menge $[f]_+$.

5. Ein allgemeiner Sattelpunkt-Satz.

In diesem Paragraphen setzen wir zwei » K -Systeme« voraus:

X, Y sind topologische Räume und $K_1: X \times Y \rightarrow \mathbb{R}$ ist in jeder Variablen partiell stetig; U, V sind topologische Räume und $K_2: U \times V \rightarrow \mathbb{R}$ ist in jeder Variablen partiell stetig.

Hieraus können wir weitere » K -Systeme« ableiten, z. B. das bestehend aus den Räumen $X \times U$ und $Y \times V$ und der Funktion

$$((x, u), (y, v)) \rightarrow K_1(x, y) + K_2(u, v),$$

oder dasjenige bestehend aus den Räumen $X \times V$ und $Y \times U$ und der Funktion

$$((x, v), (y, u)) \rightarrow K_1(x, y) + K_2(u, v).$$

Die folgenden drei Sätze sind Verallgemeinerungen von Sätzen, die für den Fall dualer topologischer Vektorräume von Moreau und Joffe-Tikomirow gefunden wurden, vgl. [5, II 1.3].

Wir weisen nochmals darauf hin, daß das Folgende auch ohne topologische Voraussetzungen gilt; die Begriffe »konvexe Funktion« und »konvexe und abgeschlossene Funktion« fallen dann zusammen.

SATZ 13. Sei $f \in \bar{F}(X \times U)$ und für alle $x \in X$ sei $u \rightarrow f(x, u)$ konkav und abgeschlossen (bzgl. U, V, K_2). Dann ist $f_1^\wedge \in \bar{F}(Y \times U)$ mit

$$f_1^\wedge(y, u) := \sup_{x \in X} \{K_1(x, y) - f(x, u)\}$$

konvex und abgeschlossen (bzgl. $X \times V, Y \times U, ((x, v), (y, u)) \rightarrow K_1(x, y) + K_2(u, v)$).

BEWEIS. Es ist zu zeigen, daß es ein $g \in \bar{F}(X \times V)$ gibt mit $g^\wedge = f_1^\wedge$; wir nehmen $g := -f_2^\sim$ (gebildet bzgl. K_2). Nach Voraussetzung ist für alle $x \in X$

$$f(x, u) = \inf_{v \in V} \{-K_2(u, v) - f_2^\sim(x, v)\}.$$

Wir setzen ein:

$$\begin{aligned} f_1^\wedge(y, u) &= \sup_{x \in X} \{K_1(x, y) - \inf_{v \in V} \{-K_2(u, v) - f_2^\sim(x, v)\}\} \\ &= \sup_{x \in X} \sup_{v \in V} \{K_1(x, y) + K_2(u, v) + f_2^\sim(x, v)\}. \end{aligned}$$

Für $f \in \bar{F}(X \times U)$ bezeichnen wir

$$\text{dom} f := \{(x, u) \mid f(x, u) \in \mathbb{R}\}$$

und z. B. die Projektion von $\text{dom} f$ auf X mit $\text{dom}_X f$. Die Voraussetzung von Satz 13 besagt dann für $x \notin \text{dom}_X f$, daß entweder $f(x, \cdot) \equiv +\infty$ oder $f(x, \cdot) \equiv -\infty$.

Für das Weitere benötigen wir folgende

FORDERUNG (F2). $(X, \bar{x}_0), (Y, \bar{y}_0), (U, \bar{u}_0), (V, \bar{v}_0)$ sind punktierte Räume. K_1 und K_2 mögen die Eigenschaft haben:

Für alle $x \in X$ und $y \in Y$ gilt $K_1(x, \bar{y}_0) = K_1(\bar{x}_0, y) = 0$.

Für alle $u \in U$ und $v \in V$ gilt $K_2(u, \bar{v}_0) = K_2(\bar{u}_0, v) = 0$.

Solche Funktionen erhält man leicht: Sei $K': X \times Y \rightarrow \mathbb{R}$ und sei gewählt $\bar{x}_0 \in X, \bar{y}_0 \in Y$. Dann ist

$$(x, y) \mapsto K'(x, y) - K'(x, \bar{y}_0) - K'(\bar{x}_0, y) + K'(\bar{x}_0, \bar{y}_0)$$

eine solche Funktion.

Auch Fixpunkte der in 3.2. nach (F1) angegebenen Involutionen hätten die in (F2) geforderte Eigenschaft.

Mit (F2) folgt dann für $f \in \bar{F}(X)$

$$(1) \quad \begin{aligned} f \wedge (\bar{y}_0) &= \sup \{ K_1(x, \bar{y}_0) - f(x) \mid x \in X \} \\ &= \sup \{ -f(x) \mid x \in X \} \\ &= -\inf \{ f(x) \mid x \in X \}. \end{aligned}$$

Dies wird benutzt im Beweis von

SATZ 14. Für jedes $u \in \text{dom}_U f$ sei $x \mapsto f(x, u)$ konvex und abgeschlossen. Bezeichne

$$h(y) := \inf \{ f_1 \wedge (y, u) \mid u \in \text{dom}_U f \}, \quad h \in \bar{F}(Y).$$

Dann gilt

$$\inf_{x \in \text{dom}_X f} \sup_{u \in \text{dom}_U f} f(x, u) = \sup_{u \in \text{dom}_U f} \inf_{x \in \text{dom}_X f} f(x, u)$$

genau dann, wenn $h(\bar{y}_0) = h \wedge (\bar{y}_0)$.

BEWEIS. Wie im Fenchelschen Fall zeigt man einerseits

$$\inf_{x \in \text{dom}_X f} \sup_{u \in \text{dom}_U f} f(x, u) = \inf_{x \in X} h \wedge (x) = -h \wedge (\bar{y}_0)$$

und andererseits

$$\sup_{u \in \text{dom}_U f} \inf_{x \in \text{dom}_X f} f(x, u) = \sup_{u \in \text{dom}_U f} -f_1 \wedge (\bar{y}_0, u) = -h(\bar{y}_0).$$

Für $f \in \bar{F}(X \times U)$ setzen wir analog zu [5]

$$g(y, v) := \sup_{u \in U} \{ K_2(u, v) - f_1 \wedge (y, u) \}.$$

Sei $\text{dom} f \neq \emptyset$ und: Für alle $x \in X$ sei $u \mapsto f(x, u)$ konkav und abgeschlossen und für alle $u \in U$ sei $x \mapsto f(x, u)$ konvex und abgeschlossen. Es berechnet sich dann

$$h(y) = -g(y, \bar{v}_0), \quad y \in Y,$$

und wir setzen noch

$$k(v) := g(\bar{y}_0, v), \quad v \in V.$$

Hat f einen Sattelpunkt in $(x_0, u_0) \in X \times U$, d.h. für alle $x \in X$, $u \in U$ gilt

$$f(x_0, u) \leq f(x_0, u_0) \leq f(x, u_0),$$

so ist $f(x_0, u_0) \in \mathbb{R}$.

SATZ 15. Sei $f \in \bar{F}(X \times U)$ wie oben angegeben. Die Menge der Sattelpunkte von f ist $\partial h(\bar{y}_0) \times \partial k(\bar{v}_0)$.

BEWEIS. Er wird aus dem Fenchelschen Fall ([5, Th. 2.3]) abgeleitet. Wir zeigen nur das Hinreichen der Bedingung und setzen demnach $x_0 \in \partial h(\bar{y}_0)$ und $u_0 \in \partial k(\bar{v}_0)$ voraus, d.h.

$$(2) \quad h(\bar{y}_0) \in \mathbb{R} \quad \text{und} \quad h \geq K_1(x_0, \cdot) + h(\bar{y}_0)$$

$$(3) \quad k(\bar{v}_0) \in \mathbb{R} \quad \text{und} \quad k \geq K_2(u_0, \cdot) + k(\bar{v}_0).$$

Wegen $\text{dom } f \neq \emptyset$ gibt es $u' \in U$, so daß $f(\cdot, u') \not\equiv +\infty$. Daher ist $f_1 \hat{\ }(\cdot, u') > -\infty$, und für jedes $y \in Y$ ist $f_1 \hat{\ } (y, \cdot)$ nicht identisch $-\infty$. Da diese Funktionen nach Satz 13 bzgl. K_2 konvex und abgeschlossen sind, gilt für jedes $y \in Y$, daß $f_1 \hat{\ } (y, \cdot) > -\infty$. Aus $f_1 \hat{\ } (\cdot, \cdot) > -\infty$ folgt für jedes $y \in Y$, daß

$$h(y) := \inf \{ f_1 \hat{\ } (y, u) \mid u \in \text{dom}_U f \} = \inf \{ f_1 \hat{\ } (y, u) \mid u \in U \},$$

und daher

$$(4) \quad h(y) \leq f_1 \hat{\ } (y, u) \quad \text{für alle } y \in Y, u \in U.$$

Satz 13 und (3) liefern

$$(5) \quad f_1 \hat{\ } (\bar{y}_0, u_0) = \sup_{v \in V} \{ K_2(u_0, v) - g(\bar{y}_0, v) \} = -k(\bar{v}_0) = h(\bar{y}_0).$$

Aus (4), (5) und (2) folgt für alle $y \in Y, u \in U$

$$(6) \quad f_1 \hat{\ } (y, u) - f_1 \hat{\ } (\bar{y}_0, u_0) \geq h(y) - h(\bar{y}_0) \geq K_1(x_0, y).$$

Da $f(\cdot, u)$ für jedes $u \in U$ konvex und abgeschlossen bzgl. K_1 ist, haben wir

$$f(x_0, u) = \sup_{y \in Y} \{ K_1(x_0, y) - f_1 \hat{\ } (y, u) \}.$$

Wegen (6) folgt hieraus

$$(7) \quad f(x_0, u) \leq -f_1 \hat{\ } (\bar{y}_0, u_0) \quad \text{für alle } u \in U.$$

(1) angewandt auf $f(\cdot, u_0)$ ergibt $-f_1 \hat{\ } (\bar{y}_0, u_0) \leq f(x, u_0)$ für alle $x \in X$. Dieses und (7) liefern die gesuchten Ungleichungen

$$f(x_0, u) \leq -f_1 \hat{\ } (\bar{y}_0, u_0) = f(x_0, u_0) \leq f(x, u_0) \quad \text{für } x \in X, u \in U.$$

LITERATURVERZEICHNIS

1. H. Bauer, *Šilovscher Rand und Dirichletsches Problem*, Ann. Inst. Fourier (Grenoble) 11 (1961), 89–136.
2. A. Brøndsted, *Conjugate convex functions in topological vector spaces*. Mat. Fys. Medd. Danske Vid. Selsk. 34 (1964), no. 2.
3. U. Dieter, *Optimierungsaufgaben in topologischen Vektorräumen I: Dualitätstheorie*, Z. Wahrscheinlichkeitstheorie Verw. Gebiete 5 (1966), 89–117.
4. W. Fenchel, *On conjugate convex functions*, Canad. J. Math. 1 (1949), 73–77.

5. A. D. Ioffe und V. M. Tikhomirow, *Dualität konvexer Funktionen und Extremalprobleme* (Russisch). Uspehi Mat. Nauk 23 (1968), 50–116.
6. J. J. Moreau, *Inf-convolution, sous-additivité, convexité des fonctions numériques*, J. Math. Pures Appl. 49 (1970), 109–154.
7. E.-A. Weiss, *Konjugierte Funktionen*, Arch. Math. (Basel) 20 (1969), 538–545.

MATHEMATISCHES INSTITUT DER UNIVERSITÄT ERLANGEN-NÜRNBERG
BISMARCKSTR. 1 1/2
852 ERLANGEN, BUNDESREPUBLIK DEUTSCHLAND