Kapitel 2

Zur Anwendung derivierter Mengen in Lagrangeschen Multiplikatorenregeln

2.1 Derivierte Mengen als Verallgemeinerung des Ableitungsbegriffs

Um eine möglichst große Klasse von Optimierungsaufgaben behandeln zu können, verfolgt man seit langem das Ziel, die Anforderungen an die Eingangsdaten so weit wie möglich abzuschwächen. Eine der in der modernen Wissenschaft mit Erfolg angewandten Methoden ist die sogenannte Bildraumtechnik. Hierbei geht es darum, alle durchzuführenden Untersuchungen soweit wie möglich in den Raum, in welchen die auftretenden Funktionen abbilden, zu verlegen. Der Vorteil dieser Methode ist es, im Urbildraum nur ein Mindestmaß an topologischen Eigenschaften voraussetzen zu müssen.

Bereits seit den Anfängen der Differentialrechnung konzentriert man sich bei der Suche nach Minimalstellen von nichtkonvexen Funktionen zunächst auf Punkte, die eine gewisse Stationaritätsbedingung erfüllen. Stationäre Punkte werden für gewöhnlich durch das Verhalten von Ableitungen der auftretenden Funktionen beschrieben. Das Formulieren von Ableitungsbegiffen allerdings fordert recht weitreichende Kenntnisse über die topologische Struktur des Urbildraumes in der Nähe des Punktes, an dem die Ableitung berechnet werden soll. So ist es verständlicherweise eines der größten Hindernisse bei der Arbeit mit der Bildraumtechnik, einen geeigneten Ableitungsbegriff zu formulieren, der ohne viel Strukturvoraussetzungen im Urbildraum auskommt.

Eine wichtige Entwicklung auf diesem Gebiet war die Einführung der derivierten Mengen durch Hestenes [18] um 1965. Betrachtet werden Optimierungsprobleme der Gestalt

$$f_1(x) \longrightarrow Min!$$

$$\text{über} \quad x \in \{x \in X |\ f_2(x) \le 0^{m_2}, f_3(x) = 0^{m_3}\}$$

mit einer nichtleeren Menge X und gewissen Funktionen $f_1 : X \to R^1$, $f_2 : X \to R^{m_2}$, $f_3 : X \to R^{m_3}$. Hierbei fassen wir unter Verwendung von $m = 1 + m_2 + m_3$ den Raum R^m als $R^1 \times R^{m_2} \times R^{m_3}$ auf, mithin jeden Vektor $y \in R^m$ als $(y_1, y_2, y_3) \in R^1 \times R^{m_2} \times R^{m_3}$. Die Relation $\le$ ist, sofern sie zwischen zwei Vektoren auftritt, wie allgemein üblich als zeilenweiser Vergleich zu verstehen. Indem wir $m_2 = 0$ und $m_3 = 0$ zulassen, werden durch obige Problemstellung auch Aufgaben ohne Ungleichungs- beziehungsweise ohne Gleichungsnebenbedingungen erfaßt. Für diese Probleme wird in [19] definiert:

Definition 2.1 *Eine Menge $D \subseteq R^m$ heißt derivierte Menge (oder Ableitungsmenge) von f an der Stelle x_0, wenn für alle $n \in N$ und jedes n–Tupel $\{d^1, \dots, d^n\} \subset D$ eine Zahl $r > 0$ und eine Abbildung $\omega : [0, r]^n \to X$ mit $\omega(0) = x_0$ existieren, so daß die Funktion $f \circ \omega$ stetig ist und*

$$f(\omega(t)) = f(x_0) + \sum_{i=1}^{n} t_i d^i + \rho(t) \qquad \forall t \in [0, r]^n, \tag{2.1}$$

gilt, wobei die Funktion $\rho :\ [0, r]^n \to R^m$ die Eigenschaft

$$\frac{\rho(t)}{\|t\|} \to 0, \ \ wenn\ t \to 0$$

besitzt.

Durch Vergleich von (2.1) mit einer üblichen Bestimmungsgleichung für Ableitungen, etwa

$$f(x_0 + h) = f(x_0) + A(h) + o(\|h\|), \tag{2.2}$$

erkennt man, daß bei der Hestenesschen Ableitungsmenge gewisse Linearkombinationen von Elementen dieser Menge die Aufgabe der Ableitung A übernehmen.

Aus einer Arbeit von Gittleman [14] aus dem Jahr 1971 stammt die folgende Erweiterung der Definition.

Definition 2.2 *Eine Menge $D \subseteq R^m$ heißt derivierte Menge von f an der Stelle x_0, wenn für jede Zahl $C > 0$ und jedes n–Tupel $\{d^1, \ldots, d^n\} \subset D$ $(n \in N)$ eine Zahl $r(C) > 0$ und eine Abbildung $\omega : Q^n(r) \to X$ existieren ($Q^n(r) = \{t \in R^n | \, t_i \geq 0, \sum_{i=1}^n t_i \leq r\}$), so daß*

$$
\begin{aligned}
f_1(\omega(t)) &\leq f_1(x_0) + \sum_{i=1}^{n} t_i d_1^i + \rho_1(t) \qquad \forall t \in Q^n(r), &(2.3)\\
f_2(\omega(t)) &\leq f_2(x_0) + \sum_{i=1}^{n} t_i d_2^i + \rho_2(t) \qquad \forall t \in Q^n(r), &(2.4)\\
f_3(\omega(t)) &= f_3(x_0) + \sum_{i=1}^{n} t_i d_3^i + \rho_3(t) \qquad \forall t \in Q^n(r), &(2.5)
\end{aligned}
$$

gilt, wobei die Funktion $\rho = (\rho_1, \rho_2, \rho_3) : \; Q^n(r) \to R^1 \times R^{m_2} \times R^{m_3}$ die Eigenschaften

$$
\frac{|\rho_1(t)|}{\sum_{i=1}^n t_i} \leq C, \quad \frac{\|\rho_2(t)\|}{\sum_{i=1}^n t_i} \leq C, \quad \frac{\|\rho_3(t)\|}{\sum_{i=1}^n t_i} \leq C \qquad \forall t \in Q^n(r) \setminus \{0^n\}
$$

besitzt und darüberhinaus ρ_3 stetig ist.

Offensichtlich genügt es nach Gittleman, wenn die Gleichung (2.1) lediglich bei den Zeilen, die mit den Gleichungsrestriktionen korrespondieren, erfüllt ist, während man sie in allen anderen Fällen zur entsprechenden Ungleichung abschwächen kann.

Im Jahr 1980 paßte Nieuwenhuis [37] die Definition für Probleme mit allgemeiner formuluierten Nebenbedingungen an:

$$
f_1(x) \longrightarrow Min!
$$

$$
\text{über} \quad x \in \{x \in X | \, f_2(x) \in -K_2, f_3(x) \in -K_3\}.
$$

Hierbei seien K_2 und K_3 zwei konvexe abgeschlossene Kegel des R^{m_2} bzw. R^{m_3}, von denen K_2 ein nichtleeres Inneres habe.

Definition 2.3 *Eine Menge $D \subseteq R^m$ heißt derivierte Menge von f an der Stelle x_0, wenn für jede Zahl $C > 0$ und jedes $(m+1)$–Tupel $\{d^1, \ldots, d^{m+1}\} \subset D$ eine Zahl $r(C) > 0$ und eine Abbildung $\omega : Q^{m+1}(r) \to X$ existieren, so daß*

$$
f_1(\omega(t)) \leq f_1(x_0) + \sum_{i=1}^{m+1} t_i d_1^i + \rho_1(t) \qquad \forall t \in Q^{m+1}(r), \tag{2.6}
$$

$$f_2(\omega(t)) \in f_2(x_0) + \sum_{i=1}^{m+1} t_i d_2^i + \rho_2(t) - K_2 \qquad \forall t \in Q^{m+1}(r), \tag{2.7}$$

$$f_3(\omega(t)) \in f_3(x_0) + \sum_{i=1}^{m+1} t_i d_3^i + \rho_3(t) - K_3 \qquad \forall t \in Q^{m+1}(r), \tag{2.8}$$

gilt, wobei die Funktion $\rho = (\rho_1, \rho_2, \rho_3) : Q^{m+1}(r) \to R^1 \times R^{m_2} \times R^{m_3}$ *die Eigenschaften*

$$\frac{|\rho_1(t)|}{\sum_{i=1}^n t_i} \leq C, \quad \frac{\|\rho_2(t)\|}{\sum_{i=1}^n t_i} \leq C, \quad \frac{\|\rho_3(t)\|}{\sum_{i=1}^n t_i} \leq C \qquad \forall t \in Q^{m+1}(r) \setminus \{0^{m+1}\}$$

besitzt.

Bei dieser Definition fällt neben der Einführung der durch die Ordnungskegel beschriebenen Vergleichsrelationen weiterhin auf, daß Nieuwenhuis sich auf die Betrachung von $(m+1)$–Tupeln aus D beschränken kann, und nicht, wie bisher, alle endlichen Kombinationen von Punkten aus D untersucht.

Die letzte Modifikation des Begriffs stammt von Breckner [3] aus dem Jahre 1994. Anstelle der reellwertigen Zielfunktion bei Nieuwenhuis wird nun eine Funktion f_1 mit Werten im R^{m_1} betrachtet. Die Ordnung im R^{m_1} wird durch einen konvexen Kegel $K_1 \subseteq R^{m_1}$ mit nichtleerem Inneren gegeben. Weiterhin wird verlangt, daß X Teilmenge eines topologischen Raumes $\mathcal{X}$ sei.

$$f_1(x) \longrightarrow K_1 - Min!$$

$$\text{über} \quad x \in S := \{x \in \mathcal{X} |\ f_2(x) \in -K_2, f_3(x) \in -K_3, x \in X\}.$$

Analog zur anfangs eingeführten Vereinbarung sei $m := m_1 + m_2 + m_3$; das heißt, wir verstehen den Raum R^m als $R^{m_1} \times R^{m_2} \times R^{m_3}$. Es sei $K := K_1 \times K_2 \times K_3 \subseteq R^m$.

Definition 2.4 *Eine Menge* $D \subseteq R^m$ *heißt* K*–derivierte Menge (oder* K*–Ableitungsmenge) von* f *an der Stelle* x_0*, wenn für jedes* n*–Tupel* $\{d^1, \ldots, d^n\} \subset D$ *eine Zahl* $r > 0$*, eine Abbildung* $\omega : B_+^n(r) \to X$ $(B_+^n(r) := \{x \in R_+^n |\ \|x\| \leq r\})$*, und eine Abbildung* $\varrho = (\varrho_1, \varrho_2, \varrho_3) : B_+^n(r) \to R^{m_1} \times R^{m_2} \times R^{m_3}$ *existieren, so daß die folgenden Bedingungen erfüllt sind:*

(A) Es gilt

$$f_1(\omega(t)) \in f_1(x_0) + \sum_{i=1}^{n} t_i d_1^i + \|t\| \varrho_1(t) - K_1 \qquad \forall t \in B_+^n(r), \tag{2.9}$$

$$f_2(\omega(t)) \in f_2(x_0) + \sum_{i=1}^{n} t_i d_2^i + \|t\| \varrho_2(t) - K_2 \qquad \forall t \in B_+^n(r), \tag{2.10}$$

$$f_3(\omega(t)) \in f_3(x_0) + \sum_{i=1}^{n} t_i d_3^i + \|t\| \varrho_3(t) - K_3 \qquad \forall t \in B_+^n(r). \tag{2.11}$$

(B) $\omega(0) = x_0$ *und* ω *ist stetig im Nullpunkt.*

(C) Es gibt einen Punkt $y^0 = (y_1^0, y_2^0) \in K_1 \times K_2$*, so daß zu jedem* $\varepsilon > 0$ *ein* $r_\varepsilon \in (0, r]$ *existiert mit*

$$\varrho_1(t) \in \varepsilon y_1^0 - K_1 \qquad \forall t \in B_+^n(r_\varepsilon), \tag{2.12}$$

$$\varrho_2(t) \in \varepsilon y_2^0 - K_2 \qquad \forall t \in B_+^n(r_\varepsilon). \tag{2.13}$$

Es gilt $\varrho_3(0^n) = 0^{m_3}$ *und* ϱ_3 *ist stetig.*

Bei den vorangegangenen Definitionen war gefordert worden, daß die Restgliedfunktion mit kleiner werdendem Argument t gegen Null strebt oder zumindestens beschränkt bleibt. An dieser Stelle sei besonders darauf hingewiesen, daß in Breckners Definition $\varrho_1(t)$ und $\varrho_2(t)$ in der Nähe von $t = 0$ durchaus weit vom Nullpunkt entfernt sein können. In (C) wird nur verlangt, daß sich diese Abweichung in „unproblematische“ Richtungen, das heißt jeweils in Richtung der negativen Ordnungskegel, erstreckt. Der Vollständigkeit halber sei erwähnt, daß, falls $\varrho_1(0^n) = 0^{m_1}$, $\varrho_2(0^n) = 0^{m_2}$ und sowohl ϱ_1 als auch ϱ_2 stetig sind, die Bedingungen (2.12) und (2.13) stets erfüllt werden (vgl. [3], Remark 3.1).

Wegen der leichteren Anwendbarkeit von Trennungssätzen im Bildraum liegt es nahe, besonders solche K–Ableitungsmengen zu betrachten, die darüberhinaus konvexe Kegel sind.

Definition 2.5 *Eine Menge* $D \subseteq R^m$ *heißt* K*–derivierter Kegel (oder* K*–Ableitungskegel) von* f *an der Stelle* x_0*, wenn* D *ein konvexer Kegel ist und für jedes* $(m+1)$*–Tupel von Elementen aus* D *eine Zahl* $r > 0$*, eine Abbildung* $\omega : B_+^{m+1}(r) \to X$ *und eine Abbildung* $\varrho = (\varrho_1, \varrho_2, \varrho_3) : B_+^{m+1}(r) \to R^{m_1} \times R^{m_2} \times R^{m_3}$ *existieren, so daß die Bedingungen (A), (B) und (C) von Definition 2.4 mit* $n = m + 1$ *erfüllt sind.*

Dies stellt keine wesentlich schärfere Anforderung an die derivierten Mengen dar. Vielmehr läßt sich aus einer gegebenen K–Ableitungsmenge einfach ein K–Ableitungskegel konstruieren (vgl. [3], Proposition 3.1).

Lemma 2.1 *Sei* $D \subseteq R^m$ *eine* K*–Ableitungsmenge einer gegebenen Funktion* $f : X \longrightarrow R^m$ *an der Stelle* $x_0 \in X$*.*
Dann ist die Menge

$$convcone\,(D)$$

stets ein K*–Ableitungskegel von* f *an der Stelle* x_0*.*

In diesem Zusammenhang sei angemerkt, daß die originale Definition der K–derivierten Mengen in [3] auf Maximierungsaufgaben ausgerichtet ist. Definition 2.5 (in Verbindung mit Definition 2.4) beschreibt also, im Grunde genommen, einen $(-K)$–Ableitungskegel im Brecknerschen Sinne. Ich habe mich für diese Umformulierung aus Gründen der Kompatibilität entschieden, da in der sonstigen Literatur die Minimierungsaufgaben dominieren.

2.2 Notwendige Optimalitätsbedingungen für multikriterielle Aufgaben

Mit den im vorangehenden Abschnitt eingeführten Begriffen gelingt es den Autoren, notwendige Optimalitätsbedingungen für Lösungen der jeweils entsprechenden Aufgaben nachzuweisen. Die so erhaltenen Optimalitätsbedingungen haben die Gestalt von Multiplikatorenregeln und entsprechen in ihrer Struktur in etwa dem im folgenden angegebenen Theorem 2.2. Daher sei es gestattet, der interessierten Leser an dieser Stelle auf die aufgeführte Litertur zu verweisen.

Wir wollen uns im folgenden auf ein Resultat aus [3] konzentrieren, welches eine notwendige Bedingung für lokale schwache Lösungen des angegebenen multikriteriellen Optimierungsproblems liefert. Der Punkt $x_0 \in S$ heißt lokale schwache Lösung dieses Problems, wenn eine Umgebung U von x_0 existiert, so daß $f_1(x_0)$ ein schwach effizienter Punkt für die Menge $f_1(S \cap U)$ bezüglich K_1 ist.

Theorem 2.2 (Brecknersche Multiplikatorenregel) *Gegeben sei eine lokale schwache Lösung $x_0 \in S$ des betrachteten Problems. Sei $D \subseteq R^m$ ein K–Ableitungskegel für f an der Stelle x_0. Dann existiert ein Multiplikator $\lambda \in K_1^* \times K_2^* \times K_3^* \setminus \{0^m\}$ mit*

$$\langle \lambda, d \rangle \geq 0 \qquad \forall d \in D, \tag{2.14}$$

$$\langle \lambda_2, f_2(x_0) \rangle = 0. \tag{2.15}$$

Dieses Theorem kann, wie in [45] gezeigt wird, um eine weitere Komplementaritätsbedingung

$$\langle \lambda_3, f_3(x_0) \rangle = 0 \tag{2.16}$$

ergänzt werden. Eine Möglichkeit des Beweises ist in Abschnitt 2.3 angegeben.

In [5] gelingt es Breckner und Göpfert, in konkreten Fällen unter weiteren Voraussetzungen bestimmte Ableitungskegel zu ermitteln. Diese sollen im folgenden dargestellt werden.

Sei $\mathcal{X}$ ein linearer, topologischer Raum. Wir betrachten das Problem

$$f_1(x) \longrightarrow Pareto - Min!$$

$$\text{über} \quad x \in S := \{x \in X | \; f_2(x) \leq 0^{m_2}, f_3(x) \in 0^{m_2}, x \in X\},$$

das wir durch $K_1 = R_+^{m_1}$, $K_2 = R_+^{m_2}$, $K_3 = \{0^{m_3}\}$ aus dem ursprünglichen erhalten. Der Punkt $x_0 \in S$ heißt lokale schwache Pareto–Lösung des obigen Problems, wenn eine Umgebung U von x_0 existiert, so daß $f_1(x_0)$ ein schwach Pareto–effizienter Punkt für die Menge $f_1(S \cap U)$ ist. Dann gilt das folgende Resultat.

Theorem 2.3 *Gegeben sei eine lokale schwache Pareto–Lösung $x_0 \in S$ des obigen Problems. Es existieren eine Menge $X_0 \subseteq \mathcal{X}$ sowie eine Funktion $F : X_0 \longrightarrow R^m$, die die folgenden Bedingungen erfüllen:*

(a) X_0 ist nichtleer und konvex.

(b) F_1 und F_2 sind konvex, F_3 ist affin.

(c) Für jede Zahl $n \in N$ und jedes n–Tupel $\{x_1, \ldots, x_n\} \subset X_0$ gibt es ein $r > 0$, so daß

$$x_0 + \sum_{i=1}^{n} t_i x_i \in X \quad \forall (t_1, \ldots, t_n) \in B_+^n(r)$$

gilt und die Funktion

$$(t_1, \ldots, t_n) \in B_+^n(r) \longmapsto f_3 \left(x_0 + \sum_{i=1}^{n} t_i x_i \right) \in R^{m_3}$$

stetig ist.

(d) Für alle $x \in X_0$ gilt

$$\begin{aligned}
\limsup_{a \downarrow 0} \frac{f_1(x_0 + ax) - f_1(x_0)}{a} &\leq F_1(x), \\
\limsup_{a \downarrow 0} \frac{f_2(x_0 + ax) - f_2(x_0)}{a} &\leq F_2(x), \\
\lim_{a \downarrow 0} \frac{f_3(x_0 + ax) - f_3(x_0)}{a} &= F_3(x),
\end{aligned}$$

und für jedes konvexe Polytop $P \subseteq X_0$ ist die Konvergenz gleichmäßig für alle $x \in P$.

Dann existiert ein Multiplikator $\lambda \in R^m \setminus \{0^m\}$ *mit* $\lambda_1 \geq 0$, $\lambda_2 \geq 0$ *und*

$$\begin{aligned} \langle \lambda, F(x) \rangle &\geq 0 \qquad \forall x \in X_0, \\ \langle \lambda_2, f_2(x_0) \rangle &= 0. \end{aligned}$$

Unter den Voraussetzungen a) - d) ist nämlich die Menge

$$D = F(X_0) \tag{2.17}$$

eine $(R_+^{m_1} \times R_+^{m_2} \times \{0^{m_3}\})$–Ableitungsmenge von f an der Stelle x_0.

Weitere Resultate, die ähnlich wie Theorem 2.3 eine Konvexifizierung des behandelten Problems durchführen, findet man beispielsweise in Arbeiten von Neustadt [36] und Tuy [52]. Der Zusammenhang der Neustadtschen Resultate mit der Theorie der derivierten Mengen ist in [14] gegeben; ein Vergleich von Tuys Multiplikatorenregel mit Theorem 2.3 wird in [43] vorgenommen.

Hinter Theorem 2.3 verbergen sich, trotz seiner augenscheinlichen Inpraktikabilität, bei weiterer Spezialisierung zwei wohlbekannte Optimalitätsbedingungen für den konvexen und für den differenzierbaren Fall. Dies unterstreicht, wie umfassend dieses Resultat ist.

Unterstellt man, daß die Menge X konvex ist, und daß die Funktionen f_1 und f_2 konvex und die Funktion f_3 affin sind, so zeigt sich, daß durch

$$\begin{aligned} X_0 &:= X - x_0, \\ F(x) &:= f(x + x_0) - f(x_0) \qquad \forall x \in X_0 \end{aligned}$$

die Bedingungen a) - d) erfüllt sind. So folgt (vgl. [5]):

Satz 2.4 *Gegeben sei eine lokale schwache Pareto–Lösung* $x_0 \in S$ *des betrachteten Problems. Sei die Menge* X *konvex, seien die Funktionen* f_1 *und* f_2 *konvex und sei* f_3 *affin.*
Dann existiert ein Multiplikator $\lambda \in R^m \setminus \{0^m\}$ *mit* $\lambda_1 \geq 0$, $\lambda_2 \geq 0$ *und*

$$\begin{aligned} \langle \lambda, f(x) \rangle &\geq \langle \lambda_1, f_1(x_0) \rangle \qquad \forall x \in X, \\ \langle \lambda_2, f_2(x_0) \rangle &= 0. \end{aligned}$$

Im Fall, daß die Funktionen f_1, f_2 und f_3 sämtlich Fréchet-differenzierbar in x_0 sind, ergeben sich a) - d) durch

$$\begin{aligned} X_0 &:= \mathcal{X}, \\ F(x) &:= f'(x_0)(x) \qquad \forall x \in X_0. \end{aligned}$$

In einem linearen, normierten Raum $\mathcal{X}$ erhält man also (vgl. [5]):

Satz 2.5 *Gegeben sei eine lokale schwache Pareto–Lösung* $x_0 \in S$ *des betrachteten Problems. Sei* $x_0 \in int\ X$ *und sei* f *im Punkt* x_0 *Fréchet–differenzierbar. Darüberhinaus sei die Funktion* f_3 *auf einer Menge* $\{x \in X;\ \|x - x_0\| \leq r\}$ $(r > 0)$ *stetig.*
Dann existiert ein Multiplikator $\lambda \in R^m \setminus \{0^m\}$ *mit* $\lambda_1 \geq 0$, $\lambda_2 \geq 0$ *und*

$$\sum_{i=1}^{m_1} \lambda_{1i} f'_{1i}(x_0) + \sum_{k=1}^{m_2} \lambda_{2k} f'_{2k}(x_0) + \sum_{l=1}^{m_3} \lambda_{3l} f'_{3l}(x_0) = 0,$$

$$\langle \lambda_2, f_2(x_0) \rangle = 0.$$

2.3 Extremaleigenschaften derivierter Mengen

Eine wesentliche Eigenschaft von derivierten Kegeln für Funktionen in einem ihrer Minimalpunkte ist es, so stellte sich bei meinen Untersuchungen heraus, daß die Öffnung dieser Kegel überwiegend in eine andere Richtung zeigt als die Öffnung der negativen Ordnungskegel. Mit „Öffnung“ eines Kegels wollen wir hier die Gesamtheit aller Richtungen verstehen, in denen der Kegel unbeschränkt ist. Jedoch war eine Aussage der Form

$$D \cap (-K) = \{0^m\} \tag{2.18}$$

für einen K–derivierten Kegel D nicht zu beweisen. Es zeigten sich Fälle, in denen D und $-K$ sich in einigen ihrer Randstrahlen berührten.

Es gelang mir aber, eine leichte Modifikation der gewünschten Aussage nachzuweisen. Wenn auch (2.18) nicht zutrifft, so existiert immerhin ein Vektor $b \in R^m$, so daß gilt

$$D \cap (-b - K) = \emptyset. \tag{2.19}$$

Es ist sogar möglich, einen solchen Vektor b mit beliebig kleiner Norm zu finden. Als geometrische Interpretation möge man sich vorstellen, daß die Mengen D und $-K$ so liegen, daß sie nach einer nur marginalen „Verschiebung“ disjunkt sind.

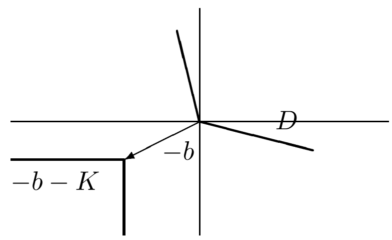

Abbildung 2.1: Trennbarkeit der Kegel D und K

Bemerkung 2.1 *Eine solche Eigenschaft wird bei Mordukhovic [29, 35] als Extremalität bezeichnet. Konkret heißt dort ein Punkt x_0 extremal für das System $\{C_1, \ldots, C_n\}$, wenn $x_0 \in \bigcap_{i=1}^n C_i$ gilt und Folgen $\{a_{ik}\}$ $(k = 1, 2, \ldots)$ mit $a_{ik} \to 0$ für $k \to \infty$ existieren, so daß*

$$\bigcap_{i=1}^{n} (C_i - a_{ik}) = \emptyset \qquad (k = 1, 2, \ldots).$$

In diesem Sinne ist also für einen K–Ableitungskegel D einer Funktion f an einer Minimalstelle der Punkt $0 \in R^m$ extremal für das System $\{D, -K\}$.

Diese Eigenschaft kann, wie im Anschluß gezeigt wird, unmittelbar zum Beweis von Breckner's Multiplikatorenregel (Theorem 2.2) genutzt werden. Zur Vereinfachung wollen wir dies unter der Voraussetzung $K_3 = \{0^{m_3}\}$ durchführen. Wie in [44] (vgl. auch [6]) gezeigt wird, kann der zulässige Bereich

$$S = \{x \in X |\ f_2(x) \in -K_2, f_3(x) \in -K_3\}$$

in jedem Fall in eine Menge

$$S = \{x \in X |\ \bar{f}_2(x) \in -\bar{K}_2, \bar{f}_3(x) = 0^k\}$$

mit $k \in [0, m_3]$, $\bar{f}_2 : X \to R^{m_2+k}$, $\bar{f}_3 : X \to R^k$ und $\bar{K}_2 \subseteq R^{m_2+k}$ ($int\, \bar{K}_2 \neq \emptyset$) transformiert werden. Daher stellt diese Spezialisierung keine Beschränkung der Allgemeinheit dar.

Diese Voraussetzung erlaubt es aber beispielsweise, beim Beweis auf den Rückgriff auf ein recht kompliziertes Linearisierungstheorem von Tuy, welches sowohl Breckner als auch Nieuwenhuis für die Behandlung der Nebenbedingung $f_3(x) \in -K_3$ benötigten, zu verzichten und statt dessen die Gleichungen $f_{3l} = 0$ $(l = 1, \ldots, m_3)$ mit Hilfe eines Satzes über implizite Funktionen aufzulösen, wie von Hestenes in [19] für skalare Zielfunktionen $(m = 1)$ vorgeführt wird.

Zunächst soll jedoch die angesprochene Behauptung formuliert werden.

Satz 2.6 *Gegeben sei eine lokale schwache Lösung $x_0 \in S$ des in Theorem 2.2 betrachteten Problems. Sei $D \subseteq R^m$ ein K–Ableitungskegel für f an der Stelle x_0. Dann existiert eine Nullfolge $\{b_k\} \subset R^m$ $(k = 1, 2, \ldots)$, so daß*

$$D \cap (-b_k - (K_1 \times K_2 \times \{0^{m_3}\})) = \emptyset \qquad \forall k. \tag{2.20}$$

Wir wollen nicht diesen, sondern einen etwas schärferen Satz beweisen. Dieser ist nötig, um auch die Komplementaritätsbedingung (2.15) in Theorem 2.2 zu erhalten. Es sei

$$L := \{y \in K_2^* |\ \langle y, f_2(x_0) \rangle = 0\}. \tag{2.21}$$

Satz 2.7 *Gegeben sei eine lokale schwache Lösung $x_0 \in S$ des in Theorem 2.2 betrachteten Problems. Sei $D \subseteq R^m$ ein K–Ableitungskegel für f an der Stelle x_0. Dann existiert eine Nullfolge $\{b_k\} \subset R^m$ $(k = 1, 2, \ldots)$, so daß*

$$D \cap (-b_k - (K_1 \times L^* \times \{0^{m_3}\})) = \emptyset \qquad \forall k. \tag{2.22}$$

Beweis. Man wähle $b_1 \in int\, K_1$ und $b_2 \in int\, L^*$. Zu b_1 and b_2 findet man stets ein $b_3 \in R^{m_3}$, so daß

$$-\begin{pmatrix} b_1 + K_1 \\ b_2 + L^* \\ b_3 \end{pmatrix} \cap D = \emptyset. \tag{2.23}$$

Anderenfalls wäre

$$-\begin{pmatrix} b_1 + K_1 \\ b_2 + L^* \\ b_3 \end{pmatrix} \cap D \neq \emptyset \qquad \forall b_3 \in R^{m_3},$$

das heißt, es gäbe $d_1 \in -b_1 - K_1 \subseteq -int\, K_1$ und $d_2 \in -b_2 - L^* \subseteq -int\, L^*$, so daß

$$\begin{pmatrix} d_1 \\ d_2 \\ -b_3 \end{pmatrix} \in D \qquad \forall b_3 \in R^{m_3}.$$

Dann wäre nach Lemma 2.8 (folgt im Anschluß) der Punkt x_0 keine lokale schwache Lösung der betrachten Aufgabe. Also gibt es ein $b_3 \in R^{m_3}$, so daß (2.23) gilt. Die Folge $\{b_k\} \subset R^m$ definiert durch

$$b_k := \begin{pmatrix} \frac{1}{k} b_1 \\ \frac{1}{k} b_2 \\ \frac{1}{k} b_3 \end{pmatrix} \qquad (k = 1, 2, \ldots)$$

strebt gegen Null und erfüllt wegen (2.23) und der Kegeleigenschaft der auftretenden Mengen die Behauptung (2.22). ■

Wie bereits angekündigt, folgen aus dem vorangegangenen Satz die Aussagen von Theorem 2.2. Nach einem einfachen Trennungssatz (etwa Theorem 3.2.3 in [32]), der wegen der endlichen Dimension des Raumes ohne Voraussetzungen über innere Punkte oder an die Kompaktheit der Mengen auskommt, können wir die konvexen und nach (2.22) disjunkten Mengen D und $-b_k - (K_1 \times L^* \times \{0^{m_3}\})$ für alle k $(k = 1, 2, \ldots)$ trennen. Somit existiert für jedes k ein Multiplikator $\lambda_k = (\lambda_{k1}, \lambda_{k2}, \lambda_{k3}) \in R^{m_1} \times R^{m_2} \times R^{m_3}$, $\|\lambda_k\| = 1$ mit der Eigenschaft

$$-\langle \lambda_k, b_k \rangle - \langle \lambda_k, y \rangle \leq \langle \lambda_k, d \rangle \qquad \forall y \in K_1 \times L^* \times \{0^{m_3}\} \quad \forall d \in D,$$

woraus bei geeigneter Wahl von y und d folgt

$$\begin{aligned}
\langle \lambda_k, d \rangle &\geq -\langle \lambda_k, b_k \rangle \qquad \forall d \in D, \\
\langle \lambda_{k1}, y_1 \rangle &\geq -\langle \lambda_{k1}, b_{k1} \rangle \qquad \forall y_1 \in K_1, \\
\langle \lambda_{k2}, y_2 \rangle &\geq -\langle \lambda_{k2}, b_{k2} \rangle \qquad \forall y_2 \in L^*.
\end{aligned}$$

Wegen der Kompaktheit der endlichdimensionalen Einheitskugel konvergiert die Folge $\{\lambda_k\}$ (oder zumindest eine Teilfolge) gegen ein $\lambda = (\lambda_1, \lambda_2, \lambda_3) \in R^{m_1} \times R^{m_2} \times R^{m_3}$, $\lambda \neq 0$, mit

$$\begin{aligned}
\langle \lambda, d \rangle &\geq 0 \qquad \forall d \in D, \\
\langle \lambda_1, y_1 \rangle &\geq 0 \qquad \forall y_1 \in K_1, \\
\langle \lambda_2, y_2 \rangle &\geq 0 \qquad \forall y_2 \in L^*.
\end{aligned}$$

Daraus folgt $\lambda_1 \in K_1^*$ und $\lambda_2 \in L \subseteq K_2^*$. Wegen der Konstruktion von L ergibt sich weiterhin

$$\langle \lambda_2, f_2(x_0) \rangle = 0.$$

Übrig bleibt der Beweis des angekündigten Lemmas.

Lemma 2.8 *Sei $D \subseteq R^m$ ein K–Ableitungskegel für eine Funktion $f : X \to R^m$ an der Stelle $x_0 \in X$. Falls es Vektoren $d_1 \in -int\, K_1$ und $d_2 \in -int\, L^*$ gibt, so daß*

$$\begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} \in D \qquad \forall d_3 \in R^{m_3},$$

so existiert eine Folge $\{x_n\} \in X$ $(n = 1, 2, \ldots)$, die gegen x_0 strebt und für die gilt

$$\begin{aligned}
f_1(x_n) - f_1(x_0) &\in -int\, K_1 \qquad \forall n, && (2.24) \\
f_2(x_n) &\in -K_2 \qquad \forall n, && (2.25) \\
f_3(x_n) &= 0^{m_3} \qquad \forall n. && (2.26)
\end{aligned}$$

Beweis. Es seien

$$a^i := \begin{pmatrix} d_1 \\ d_2 \\ 0^{m_3} \end{pmatrix} + \begin{pmatrix} 0^{m_1} \\ 0^{m_2} \\ e_i \end{pmatrix} \qquad (i = 1, \ldots m_3)$$

und

$$a^0 := \begin{pmatrix} d_1 \\ d_2 \\ 0^{m_3} \end{pmatrix} - \sum_{i=1}^{m_3} \begin{pmatrix} 0^{m_1} \\ 0^{m_2} \\ e_i \end{pmatrix}.$$

(e_i steht für den i-ten Einheitsvektor des Raumes R^{m_3}.) Nach Voraussetzung gehören alle a^i ($i = 0, 1, \ldots, m_3$) zu D. Gemäß Definition 2.5 finden wir für das $(m+1)$-Tupel $\{a^0, a^1, \ldots, a^{m_3}, 0^{m_1}, 0^{m_2}\}$ eine Konstante r und Funktionen ω und ϱ mit den Eigenschaften (A), (B) und (C). Zur Verkürzung sei

$$\begin{aligned} \bar{\omega}(t) &:= \omega(t, 0^{m_1}, 0^{m_2}) : \ B_+^{m_3+1}(r) \to X, \\ \bar{\varrho}(t) &:= \varrho(t, 0^{m_1}, 0^{m_2}) : \ B_+^{m_3+1}(r) \to R^m. \end{aligned}$$

Aus (A) folgt speziell für $(t, 0^{m_1}, 0^{m_2})$ mit $t \in B_+^{m_3+1}(r)$

$$f_1(\bar{\omega}(t)) - f_1(x_0) - \left(\sum_{i=0}^{m_3} t_i\right) d_1 - \|t\|\bar{\varrho}_1(t) \in -K_1 \quad \forall t \in B_+^{m_3+1}(r), \tag{2.27}$$

$$f_2(\bar{\omega}(t)) - f_2(x_0) - \left(\sum_{i=0}^{m_3} t_i\right) d_2 - \|t\|\bar{\varrho}_2(t) \in -K_2 \quad \forall t \in B_+^{m_3+1}(r), \tag{2.28}$$

$$f_3(\bar{\omega}(t)) - \sum_{i=0}^{m_3} t_i a_{i3} - \|t\|\bar{\varrho}_3(t) = 0^{m_3} \quad \forall t \in B_+^{m_3+1}(r). \tag{2.29}$$

Nach (B) ist auch $\bar{\omega}$ stetig im Nullpunkt und es gilt $\bar{\omega}(0) = x_0$. Aus (C) folgt schließlich die Existenz von $(y_1^0, y_2^0) \in K_1 \times K_2$, so daß $\forall \varepsilon > 0 \ \exists r_\varepsilon \in (0, r]$ mit

$$\bar{\varrho}_1(t) \in \varepsilon y_1^0 - K_1 \qquad \forall t \in B_+^{m_3+1}(r_\varepsilon), \tag{2.30}$$

$$\bar{\varrho}_2(t) \in \varepsilon y_2^0 - K_2 \qquad \forall t \in B_+^{m_3+1}(r_\varepsilon), \tag{2.31}$$

und es gilt $\bar{\varrho}_3(0^{m_3+1}) = 0^{m_3}$ und $\bar{\varrho}_3$ ist stetig.

Wir werden zeigen, daß die folgenden Behauptungen erfüllt sind:

(i) Es gibt eine Konstante $r_1 \in (0, r]$, so daß

$$f_1(\bar{\omega}(t)) - f_1(x_0) \in -int\, K_1 \qquad \forall t \in B_+^{m_3+1}(r_1) \setminus \{0^{m_3+1}\}. \tag{2.32}$$

(ii) Es gibt eine Konstante $r_2 \in (0, r]$, so daß

$$f_2(\bar{\omega}(t)) \in -K_2 \qquad \forall t \in B_+^{m_3+1}(r_2). \tag{2.33}$$

(iii) Es gibt eine Folge $\{t_n\} \subset B_+^{m_3+1}(r) \setminus \{0^{m_3+1}\}$, die gegen Null strebt, mit der Eigenschaft

$$f_3(\bar{\omega}(t_n)) = 0^{m_3} \qquad \forall n. \tag{2.34}$$

Beweis von (i). Wegen $d_1 \in -int\, K_1$ existiert ein genügend kleines $\varepsilon_0 > 0$ so daß

$$\frac{\sum_{i=0}^{m_3} t_i}{\|t\|} d_1 + \varepsilon_0 y_1^0 \in -int\, K_1 \quad \forall t = (t_0, t_1, \ldots, t_{m_3}) \in R_+^{m_3+1} \setminus \{0^{m_3+1}\}.$$

Zu ε_0 existiert ein r_{ε_0}, so daß (2.30) erfüllt ist. Mit $r_1 := r_{\varepsilon_0}$ folgt hieraus und aus (2.27)

$$\begin{aligned} f_1(\bar{\omega}(t)) - f_1(x_0) &\in \left(\sum_{i=0}^{m_3} t_i\right) d_1 + \|t\| \bar{\varrho}_1(t) - K_1 \\ &\subseteq \left(\sum_{i=0}^{m_3} t_i\right) d_1 + \varepsilon_0 \|t\| y_1^0 - K_1 \\ &\subseteq -int\, K_1 \qquad \forall t \in B_+^{m_3+1}(r_1) \setminus \{0^{m_3+1}\}. \end{aligned}$$

Beweis von (ii). Analog zum Beweis von (i) folgt aus $d_2 \in -int\, L^*$ und (2.31), die Existenz eines $r' \in (0, r]$, so daß

$$\left(\sum_{i=0}^{m_3} t_i\right) d_2 + \|t\| \bar{\varrho}_2(t) \in -int\, L^* \qquad \forall t \in B_+^{m_3+1}(r') \setminus \{0^{m_3+1}\}.$$

Für jeden Multiplikator $\mu \in L$ gilt daher

$$\left\langle \mu, f_2(x_0) + \left(\sum_{i=0}^{m_3} t_i\right) d_2 + \|t\| \bar{\varrho}_2(t) \right\rangle \leq 0 \qquad \forall t \in B_+^{m_3+1}(r').$$

Im Fall $L = K_2^*$ können den Beweis hier mit $r_2 = r'$ beenden. Anderenfalls finden für jeden Multiplikator $\mu \in K_2^* \setminus L$ wegen $\langle \mu, f_2(x_0) \rangle < 0$ gleichfalls eine Konstante $r_\mu > 0$, so daß

$$\left\langle \mu, f_2(x_0) + \left(\sum_{i=0}^{m_3} t_i\right) d_2 + \|t\| \bar{\varrho}_2(t) \right\rangle \leq 0 \qquad \forall t \in B_+^{m_3+1}(r_\mu).$$

Somit haben wir

$$\forall \mu \in K_2^* \, \exists r_\mu \in (0, r] \quad \left\langle \mu, f_2(x_0) + \left(\sum_{i=0}^{m_3} t_i\right) d_2 + \|t\| \bar{\varrho}_2(t) \right\rangle \leq 0 \quad \forall t \in B_+^{m_3+1}(r_\mu).$$

Der Kegel K_2^* wird durch eine kompakte Basis $\mathcal{B} := \{y \in K_2^* |\ \|y\| = 1\}$ erzeugt. (Falls $K_2^* = \{0^{m_2}\}$, so wäre bereits oben der Fall $L = K_2^*$ eingetreten.) Daher kann man die Konstante r_μ unabhängig von μ fixieren. Es gilt nämlich

$$\forall \mu \in \mathcal{B} \, \exists r_\mu \in (0, r] \quad \left\langle \mu, f_2(x_0) + \left(\sum_{i=0}^{m_3} t_i\right) d_2 + \|t\| \bar{\varrho}_2(t) \right\rangle \leq 0 \quad \forall t \in B_+^{m_3+1}(r_\mu).$$

Wegen der Kompaktheit von $\mathcal{B}$ folgt

$$\exists r_2 \in (0, r]\, \forall \mu \in \mathcal{B} \quad \left\langle \mu, f_2(x_0) + \left(\sum_{i=0}^{m_3} t_i\right) d_2 + \|t\| \bar{\varrho}_2(t) \right\rangle \leq 0 \quad \forall t \in B_+^{m_3+1}(r_2).$$

Und weil sich jedes Element von K_2^* als nichtnegatives Vielfaches eines Elements aus $\mathcal{B}$ darstellen läßt, folgt schließlich

$$\left\langle \mu, f_2(x_0) + \left(\sum_{i=0}^{m_3} t_i\right) d_2 + \|t\| \bar{\varrho}_2(t) \right\rangle \leq 0 \quad \forall t \in B_+^{m_3+1}(r_2) \quad \forall \mu \in K_2^*.$$

Also ist

$$f_2(x_0) + \left(\sum_{i=0}^{m_3} t_i\right) d_2 + \|t\| \bar{\varrho}_2(t) \in -K_2 \qquad \forall t \in B_+^{m_3+1}(r_2),$$

und wegen (2.28)

$$f_2(\bar{\omega}(t)) \in -K_2 \qquad \forall t \in B_+^{m_3+1}(r_2).$$

Beweis von (iii). Sei $h = (1, \ldots, 1) \in R^{m_3+1}$. Sei $\delta > 0$ hinreichend klein, so daß

$$\|\tau(h + \sigma)\| \leq r \qquad \forall \sigma \in B^{m_3+1}(1) \quad \forall \tau \in [0, \delta].$$

Wegen $h + \sigma \in R_+^{m_3+1}$ für alle $\sigma \in B^{m_3+1}(1)$ gilt außerdem

$$\tau(h + \sigma) \in B_+^{m_3+1}(r) \qquad \forall \sigma \in B^{m_3+1}(1) \quad \forall \tau \in [0, \delta].$$

Wir betrachten die vektorwertige Funktion $\Phi:\ B^{m_3+1}(1) \times [0, \delta] \to R^{m_3}$, zeilenweise definiert durch

$$\Phi_l(\sigma, \tau) := \begin{cases} \frac{1}{\tau} f_{3l}(\bar{\omega}(\tau(h + \sigma))), & \text{wenn } \tau > 0 \\ \\ \sum_{i=0}^{m_3} a_{3l}^i \sigma_i, & \text{wenn } \tau = 0 \end{cases} \qquad (l = 1, \ldots, m_3).$$

Die Funktion Φ ist überall stetig; auch in $\tau = 0$, denn es gilt wegen (2.29)

$$\begin{aligned} \lim_{\tau \to +0} \Phi_l(\sigma, \tau) &= \lim_{\tau \to +0} \frac{f_{3l}(\bar{\omega}(\tau(h + \sigma)))}{\tau} \\ &= \lim_{\tau \to +0} \left(\frac{\sum_{i=0}^{m_3} \tau(h_i + \sigma_i) a_{3l}^i}{\tau} + \frac{\|\tau(h + \sigma)\|}{\tau} \bar{\varrho}_{3l}(\tau(h + \sigma)) \right) \\ &= \sum_{i=0}^{m_3} (h_i + \sigma_i) a_{3l}^i + \|(h + \sigma)\| \lim_{\tau \to +0} \bar{\varrho}_{3l}(\tau(h + \sigma)) \\ &= \sum_{i=0}^{m_3} \sigma_i a_{3l}^i \qquad (l = 1, \ldots, m_3) \end{aligned}$$

gleichmäßig für alle $\sigma \in B^{m_3+1}(1)$. Darüberhinaus gilt

$$\Phi(\sigma, 0) = A\sigma$$

mit

$$A = \begin{pmatrix} -1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & \cdots & 1 \end{pmatrix} \in R^{m_3 \times (m_3+1)},$$

wobei die Matrix A den Rang m_3 besitzt.

Somit ist ein Satz über implizite Funktionen (Theorem 8.2 in [19]) auf die Funktion Φ anwendbar. Mithin existieren eine Konstante $\delta' \in (0, \delta]$ und eine Funktion $\sigma : [0, \delta'] \to B^{m_3+1}(1)$ mit

$$\Phi(\sigma(\tau), \tau) = 0^{m_3} \qquad \forall \tau \in [0, \delta'] \tag{2.35}$$

und

$$\lim_{\tau \to +0} \sigma(\tau) = \sigma(0) = 0^{m_3+1}. \tag{2.36}$$

Sei $\{\tau_n\} \subset (0, \delta']$ $(n = 1, 2, \ldots)$ eine Nullfolge positiver Zahlen. Wir betrachten die Folge $\{t_n\} \subset R^{m_3+1}$ definiert durch

$$t_n = \tau_n(h + \sigma(\tau_n)) \qquad (n = 1, 2, \ldots).$$

Für die Glieder diese Folge gilt $t_n \in B_+^{m_3+1}(r)$ $\forall n$ und wegen $\tau_n > 0$ $\forall n$ und (2.36) sogar

$$t_n \in B_+^{m_3+1}(r) \setminus \{0^{m_3+1}\} \qquad \forall n.$$

Weiterhin gilt $\lim_{n\to\infty} t_n = 0$ und wegen (2.35) schließlich

$$\begin{aligned} f_3(\bar{\omega}(t_n)) &= f_3(\bar{\omega}(\tau_n(h + \sigma(\tau_n)))) \\ &= \tau_n \Phi(\sigma(\tau_n), \tau_n) \\ &= 0^{m_3} \qquad \forall n. \end{aligned}$$

Weil $\{t_n\}$ gegen Null strebt, gibt es einen Index n_0, so daß $t_n \in B_+^{m_3+1}(r_1)$ und $t_n \in B_+^{m_3+1}(r_2)$ für $n > n_0$. Die Folge $\{x_n\} \subset X$, definiert durch

$$x_n := \bar{\omega}(t_{n+n_0}) \qquad (n = 1, 2, \ldots),$$

strebt wegen $\bar{\omega}(0) = x_0$ und der Stetigkeit von $\bar{\omega}$ gegen x_0 und besitzt wegen (2.32), (2.33) und (2.34) die im Lemma behaupteten Eigenschaften. ■