Derivierte Mengen in der multikriteriellen Optimierung

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium (Dr. rer. nat.)

vorgelegt der

Mathematisch–Naturwissenschaftlich–Technischen Fakultät
(mathematisch–naturwissenschaftlicher Bereich)
der Martin–Luther–Universität Halle–Wittenberg

von Herrn Dipl.-Math. Matthias Sekatzek

geb. am 5. Januar 1972 in Merseburg

Gutachterin bzw. Gutachter:

1. Prof. Dr. Alfred Göpfert, Martin–Luther–Universität Halle–Wittenberg
2. Prof. Dr. Wolfgang W. Breckner, Babeş–Bolyai–Universität Cluj–Napoca
3. Prof. Dr. Johannes Jahn, Friedrich–Alexander–Universität Erlangen–Nürnberg
4. Prof. Dr. Christiane Tammer, Martin–Luther–Universität Halle–Wittenberg

Halle (Saale), den 29. November 1999

Inhaltsverzeichnis

Kapitel 1

Einführung

Seit langem beschäftigt sich die Mathematik mit dem Auffinden extremaler Werte von Funktionen und der Lage der entsprechenden Extremalstellen. Neben dem rein theoretischen Aspekt war diese Entwicklung immer auch getrieben von konkreten Anforderungen, etwa aus der Physik oder der Ökonomie, bei der nach Parametern gesucht wurde, für die der Wert einer konstruierten Größe einen maximalen oder minimalen Betrag erreicht. Die mathematische Disziplin, die sich mit derartigen Aufgaben beschäftigt, bezeichnen wir als Optimierung.

Häufig verlangt die Situation vom Anwender aber auch, daß nicht nur eine solche Forderung, sondern mehrere Bewertungskriterien bestens erfüllt werden sollen. Hierbei spricht man von Problemen der multikriteriellen oder Mehrziel- oder auch Vektoroptimierung.

Das Wesen solcher Probleme besteht darin, daß es für gewöhnlich keine Lösungen gibt, die alle Ziele gleichzeitg zum Optimum führen. Vielmehr ist es so, daß sich die Zielkriterien stark widersprechen, so daß der Optimalpunkt der einen Zielfunktion sehr schlecht in Bezug auf ein anderes Kriterium ist. (Ein Problem, bei dem die Ziele simultan erreicht werden können, ist auch ohne die Methoden der Mehrzieloptimierung zu lösen. Solche Aufgaben seien daher für unsere Zwecke nicht „wesentlich".)

Den Ausweg aus dieser Lage liefert eine Vorgehensweise, die auch außerhalb der Wissenschaften zur Schlichtung in Streitfällen gern strapaziert wird: Wenn es schon keine Möglichkeit gibt, allen Anforderungen gerecht zu werden, dann soll wenigstens eine Lösung gefunden werden, die einen Kompromiß zwischen den sich widersprechenden Zielen liefert. Eine solche Lösung soll so beschaffen sein, daß in einem noch festzulegenden Sinne alle Ziele gleichzeitig möglichst gut erfüllt werden. In der Vektoroptimierung bezeichnet man solche Kompromißlösungen auch als *effiziente* Punkte.

Eine sehr populäre, weil auch naheliegende, Definition von Kompromißlösungen stammt von dem italienischen Volkswirtschaftler und Soziologen V. Pareto (1848–1923). Demnach wird eine Alternative von einer anderen *dominiert*, wenn die letztere bezüglich aller Ziele nicht schlechter und hinsichtlich mindestens eines Zieles sogar besser als die erste ist. Eine Alternative, die von keiner anderen erlaubten Möglichkeit dominiert wird, heißt dann effizient.

Diese Begriffe sollen am folgenden Beispiel verdeutlich werden. Weiterhin zeigt dieses Beispiel, daß auch im täglichen Leben ständig multikriterielle Entscheidungsfindungen vorgenommen werden müssen.

Beispiel 1.1 *Ein Reisender sucht eine optimale Zugverbindung für die Fahrt von Halle (S) Hbf. nach Passau Hbf. In Betracht kommen alle Züge, die Halle zwischen 7.00 Uhr und 10.00 Uhr verlassen. Er stellt an die „Güte“ der Verbindung die folgenden Anforderungen.*

1. *Die Abfahrtszeit AB liegt möglichst spät.*
2. *Die Ankunftszeit AN liegt möglichst früh.*
3. *Die Zahl der Umstiege UM sei möglichst gering.*
4. *Der Fahrpreis PR sei möglichst gering.*

Ein Blick in den Fahrplan liefert ihm diese Alternativen:

Nr.	Abfahrt (AB)	Ankunft (AN)	Fahrzeit (FZ)	umsteigen (UM)	benutzte Züge	Strecke (ST)	Preis (PR)
1	7.32	14.34	7:02 h	1	IR, ICE	706 km	209,00 DM
2	7.32	14.34	7:02 h	2	IR, EC, ICE	507 km	155,00 DM
3	7.58	14.34	6:36 h	2	SE, IR, ICE	518 km	146,00 DM
4	8.31	20.39	12:08 h	0	IC	1230 km	342,00 DM
5	9.32	15.50	6:18 h	2	IR, IC, IC	507 km	144,00 DM

Quelle: Deutsche Bahn AG, Winterfahrplan 98/99.

Die Verbindung 2 wird als von der Alternative 3 dominiert. Letztere hat eine spätere Abfahrtszeit, erfordert die selbe Zahl an Umstiegen und ist darüberhinaus nicht teurer als erstere. Die Verbindungen 1,3,4 und 5 hingegen sind effizient. Selbst die offenbar absurde Reisemöglichkeit 4 (über Köln) ist in der Hinsicht eine effiziente Alternative, als daß es keine schnellere oder billigere Verbindung gibt, die ebenfalls ohne Umsteigen auskommt.

Beispiel 1.2 *Im vorangegangenen Beispiel werden die Kriterien 1 und 2 durch die folgende Forderung ersetzt.*

Die Fahrzeit $FZ = AN - AB$ *sei möglichst kurz.*

Wie im Beispiel 1.1 wird 2 von 3 dominiert. Desweiteren dominiert die Alternative 5 die Verbindung 3 und mithin auch 2. Effiziente Alternativen sind 1,4 und 5.

Beispiel 1.3 *Um sich die Entscheidung zu erleichtern, entschließt sich der Reisende aus den beiden vorangegangenen Beispiele, die von ihm aufgestellten Anforderungen zu bewerten. Unterstellen wir ihm eine materialistische Denkweise und sagen, er ordne jedem dieser Kriterien einen Geldwert zu. Eine im Zug verbrachte Stunde ist ihm 100,00 DM wert. Die Kosten für einen Umstieg beziffert er mit 20,00 DM. Das einzige Kriterium für die Qualität einer Verbindung lautet dann:*

Der Gesamtpreis $GP = 100FZ + 20UM + PR$ *sei möglichst gering.*

Aus diesen Vorgaben errechnet man:

Nr. Nr.	Kosten der Fahrzeit ($100FZ$)	Kosten des Umsteigens ($20UM$)	Kosten des Fahrscheins ($1PR$)	Gesamtpreis (GP)
1	703,33 DM	20,00 DM	209,00 DM	932,33 DM
2	703,33 DM	40,00 DM	155,00 DM	898,33 DM
3	660,00 DM	40,00 DM	146,00 DM	846,00 DM
4	1213,33 DM	0,00 DM	349,00 DM	1562,33 DM
5	630,00 DM	40,00 DM	144,00 DM	814,00 DM

Offensichtlich ist unter dieser Bewertung die Verbindung 5 die günstigste. Die auf diese Weise ermittelte Lösung ist im übrigen eine der effizienten Alternativen von Beispiel 1.2.

Die im Beispiel 1.3 vorgenommene „Bewertung“ ist unabhänig von dem hier betrachteten Problem eine oft verwendete Methode zur Lösung multikriterieller Aufgaben. Aus mathematischer Sicht wurde dabei aus dem Vektor, bestehend aus den drei Zielen (FZ, UM, PR) und dem Vektor der Bewertungen $(100, 20, 1)$ das Skalarprodukt

$$\left\langle \begin{pmatrix} 100 \\ 20 \\ 1 \end{pmatrix}, \begin{pmatrix} FZ \\ UM \\ PR \end{pmatrix} \right\rangle = 100FZ + 20UM + PR$$

gebildet und dieses minimiert. Eine solche Vorgehensweise, bei der anstelle eines Vektors von Zielen eine daraus gebildete skalare Größe optimiert wird, heißt *Skalarisierung*.

Wie wir später zeigen werden, ist für bestimmte Klassen von Skalarisierungen die Lösung der skalarisierten Aufgabe stets effizient für das vektorielle Optimierungsproblem. Im vorliegenden Fall hätte die Bewertung von FZ, UM und PR mit einem beliebigen Preis größer als Null immer eine der in Beispiel 1.2 als effizient ermittelt Alternativen ergeben.

Wir wollen nun die bis hierher intuitiv eingeführten Begriffe in eine exakte mathematische Sprechweise übertragen. Dazu bedarf es zunächst einiger grundlegender Festlegungen.

Wir bezeichnen eine Teilmenge M eines linearen Raumes $\mathcal{X}$ als *konvex*, wenn für je zwei Punkte $x, y \in M$ gilt

$$\lambda x + (1-\lambda)y \in M \qquad \forall \lambda \in [0,1]. \tag{1.1}$$

Mit $conv\,(M)$ beschreiben wir die konvexe Hülle der Menge M, also

$$conv\,(M) := \{z \in \mathcal{X} \mid \exists x, y \in M; \exists \lambda \in [0,1] :\ z = \lambda x + (1-\lambda)y\}\,. \tag{1.2}$$

Für Teilmengen M eines topologischen Raumes verstehen wir unter $int\,M$ das Innere, unter $cl\,M$ den Abschluß sowie unter $bd\,M$ den Rand der Menge.

Es sei R^p der p–dimensionale euklidische Raum aller p–Tupel $x = (x_1, \ldots, x_p)$. Zur Verdeutlichung bezeichnen wir das Nullelement des R^p in allen Zweifelsfällen mit 0^p.

Der Raum R^p ist mit der Norm

$$\|x\| := \sqrt{\sum_{i=1}^{p} (x_i)^2} \qquad \forall x \in R^p \tag{1.3}$$

und dem Skalarprodukt

$$\langle x, y\rangle := \sqrt{\sum_{i=1}^{p} x_i y_i} \qquad \forall x, y \in R^p \tag{1.4}$$

versehen.

Eine Teilmenge $K \subseteq R^p$ heißt *Kegel*, wenn für jeden Punkt $x \in K$ gilt

$$\alpha x \in K \qquad \forall \alpha \geq 0. \tag{1.5}$$

Mit $cone\,(M)$ bezeichnen wir den Kegel, der durch die Menge M aufgespannt wird, das heißt

$$cone\,(M) := \{y \in R^p \mid \exists \alpha \geq 0, \exists x \in M :\ y = \alpha x\}\,. \tag{1.6}$$

Mit *convcone* (M) bezeichnen wir den durch M aufgespannten konvexen Kegel

$$\mathit{convcone}\,(M) := \left\{ y \in R^p \;\middle|\; \exists k, \exists \alpha_1, \ldots, \alpha_k \geq 0, \exists x^1, \ldots, x^k \in M : \; y = \sum_{i=1}^{k} \alpha_i x^i \right\}. \tag{1.7}$$

Zu einem Kegel $K \subseteq R^p$ wird mit

$$K^* := \{x \in R^p |\; \langle x, y \rangle \geq 0 \;\forall y \in K\} \tag{1.8}$$

der zugehörige *duale Kegel* definiert.

Ein häufig betrachteter Kegel im Raum R^p ist der sogenannte positive Orthant

$$R^p_+ := \{y \in R^p |\; y_i \geq 0 \; (i = 1, \ldots, p)\}.$$

Eine reellwertige Funktion ϕ, definiert auf einer konvexen Menge M, bezeichnen wir als konvex, wenn für alle $x, y \in M$ gilt

$$\phi(\lambda x + (1 - \lambda)y) \leq \lambda\phi(x) + (1 - \lambda)\phi(y) \qquad \forall \lambda \in [0, 1]. \tag{1.9}$$

Vektorwertige Funktionen $\Phi : \; M \longrightarrow R^p$ werden konvex genannt, wenn für $x, y \in M$ gilt

$$\Phi(\lambda x + (1 - \lambda)x) \in \lambda\Phi(x) + (1 - \lambda)\Phi(y) - R^p_+ \qquad \forall \lambda \in [0, 1]. \tag{1.10}$$

Dies bedeutet lediglich, daß $\Phi = (\Phi_1, \ldots, \Phi_p)$ genau dann konvex ist, wenn jede Komponente Φ_i $(i = 1, \ldots, p)$ von Φ eine reellwertige konvexe Funktion ist.

Eine reellwertige Funktion ϕ, definiert auf einem linearen normierten Raum $\mathcal{X}$, heißt Fréchet–differenzierbar im Punkt $x_0 \in \mathcal{X}$, wenn es auf $\mathcal{X}$ ein lineares, stetiges Funktional $\phi'(x_0)$ gibt mit

$$\phi(x_0 + x) = \phi(x_0) + \phi'(x_0)(x) + \|x\| s(x) \qquad \forall x \in \mathcal{X}, \tag{1.11}$$

wobei $s(x)$ der Bedingung $\lim_{x \to 0} s(x) = 0$ genügt. Das Funktional $\phi'(x_0)$ heißt Fréchet–Ableitung von ϕ im Punkt x_0. Analog definieren wir die Fréchet–Ableitung vektorwertiger Funktionen $\Phi : \; \mathcal{X} \longrightarrow R^p$ durch

$$\Phi(x_0 + x) = \Phi(x_0) + \Phi'(x_0)(x) + \|x\| S(x) \qquad \forall x \in \mathcal{X}, \tag{1.12}$$

Hierbei verwenden wir als Ableitung eine lineare, stetige Funktion $\Phi'(x_0) : \; \mathcal{X} \longrightarrow R^p$. Für das Restglied $S(x) \in R^p$ gelte entsprechend $\lim_{x \to 0} S(x) = 0^p$.

Für lineare, stetige Funktionale ϕ auf einem linearen Raum mit der Norm $\|.\|$ definieren wir die *duale Norm* durch

$$\|\phi\|_* := \sup_{\|x\|=1} |\phi(x)|. \tag{1.13}$$

Mit den so eingeführten Bezeichnungen wollen den Effizienzbegriff von Pareto konkretisieren: Ein Punkt $y \in R^p$ wird von $x \in R^p$ *dominiert*, wenn

$$x \in y - R^p_+ \setminus \{0^p\}. \tag{1.14}$$

Ein Punkt $y \in Y$ einer Teilmenge $Y \subseteq R^p$ heißt *Pareto–effizient für die Menge Y*, wenn er von keinem anderen Punkt aus Y dominiert wird, das heißt

$$(y - R^p_+ \setminus \{0^p\}) \cap Y = \emptyset. \tag{1.15}$$

Durch den Kegel R^p_+ wird im Raum R^p eine Vergleichsrelation eingeführt. Diese wird auch als die „natürliche“ Vergleichsrelation bezeichnet, weil hier die Vektoren in natürlicher Weise, das heißt komponentenweise verglichen werden. Eine allgemeinere Formulierung der Effizienz wird erreicht, indem wir eine solche Relation durch einen beliebigen konvexen Kegel $K \subseteq R^p$ festlegen.

Ein Punkt $y \in R^p$ wird von $x \in R^p$ *bezüglich K dominiert*, wenn

$$x \in y - K \setminus \{0^p\}. \tag{1.16}$$

In dieser Sprechweise heißt ein Punkt $y \in Y$ *effizient für die Menge Y bezüglich K*, wenn

$$(y - K \setminus \{0^p\}) \cap Y = \emptyset. \tag{1.17}$$

Offenbar ist die Pareto–Effizienz ein Spezialfall der so eingeführten Effizienz bezüglich eines Kegels, wenn man sich davon überzeugt, daß R^p_+ ein konvexer Kegel ist.

Bemerkung 1.1 *Ist der konvexe Kegel K darüberhinaus spitz, das heißt*

$$K \cap (-K) = \{0^p\},$$

so ist die durch K festgelegte Vergleichsrelation $\leq_K$, definiert als

$$x \leq_K y \Longleftrightarrow y - x \in K,$$

reflexiv, transitiv und antisymmetrisch, mithin eine Halbordnung. Wir können in dieser Arbeit auf die Spitzheit der Ordnungskegel verzichten. Man beachte, daß dadurch die Antisymmetrie der Vergleichsrelation im allgemeinen nicht gegeben ist.
Die „natürliche“ Vergleichsrelation ist eine Halbordnung, da R^p_+ ein spitzer konvexer Kegel ist.

Eine Abschwächung des letzteren Begriffs stellt die schwache Effizienz dar: Es sei $K \subseteq R^p$ ein konvexer Kegel mit nichtleerem Inneren. Ein Punkt $y \in Y \subseteq R^p$ heißt *schwach effizient für die Menge Y bezüglich K*, wenn

$$(y - int\, K) \cap Y = \emptyset. \tag{1.18}$$

Tatsächlich ist jeder effiziente Punkt einer Menge auch schwach effizient für diese.

Mit dem anfangs angeführten Beispiel 1.3 haben wir die Methode der Skalarisierung multikriterieller Optimierungsaufgaben vorgestellt. Hierbei wird anstelle eines Problems

> Bestimme alle Punkte $y \in Y$, die effizient für die Menge Y bezüglich K sind!

das Problem

> Bestimme alle Punkte $y \in Y$, die den Ausdruck $\langle \lambda, y \rangle$ über Y minimieren!

mit einem geeignet gewählten Skalarisierungvektor $\lambda \in K^*$ betrachtet.

Am Beispiel zeigte sich, daß mit der dort gewählten Skalarisierung der drei Kriterien eine Lösung gefunden wurde, die auch ein effizienter Punkt der entsprechenden nichtskalarisierten Aufgabe war. Dies ist in der Tat ein allgemeines Phänomen.

Satz 1.1 *Sei Y eine nichtleere Teilmenge von R^p und K ein konvexer Kegel in R^p. Gegeben sei ein Funktional $\lambda \in K^*$.*
Dann ist jeder Punkt $y_0 \in Y$ mit

$$\langle \lambda, y_0 \rangle \leq \langle \lambda, y \rangle \qquad \forall y \in Y \tag{1.19}$$

a) schwach effizient für die Menge Y bezüglich K, falls $\lambda \neq 0^p$;

b) effizient für die Menge Y bezüglich K, falls λ im Quasi–Inneren von K^* liegt, das heißt

$$\langle \lambda, y \rangle > 0 \qquad \forall y \in K \setminus \{0^p\}.$$

Beweis. a) Angenommen, y_0 wäre nicht schwach effizient für die Menge Y bezüglich K. Dann gibt es einen Punkt $y_1 \in Y$ mit

$$y_1 \in y_0 - int\, K.$$

Weil für $\lambda \neq 0^p$ gilt

$$\langle \lambda, y \rangle > 0 \qquad \forall y \in int\, K,$$

folgt hieraus

$$\langle \lambda, y_1 \rangle < \langle \lambda, y_0 \rangle,$$

was der Voraussetzung (1.19) widerspricht.
b) Angenommen, y_0 wäre nicht effizient für die Menge Y bezüglich K. Dann gibt es einen Punkt $y_1 \in Y$ mit

$$y_1 \in y_0 - K \setminus \{0^p\}.$$

Weil λ im Quasi–Inneren von K^* liegt, folgt hieraus

$$\langle \lambda, y_1 \rangle < \langle \lambda, y_0 \rangle,$$

was der Voraussetzung (1.19) widerspricht. ■

Von besonderem Interesse ist allerdings die Umkehrung dieses Satzes. Es stellt sich die Frage, ob alle effizienten Punkte eines multikriteriellen Optimierungsproblems auch Lösung eines entsprechenden skalarisierten Problems sind, wenn man den Multiplikator λ nur hinreichend oft variiert. Gelingt es, dieses zu zeigen, so können wir die Gesamtheit aller effizienten Punkte durch Lösung einer Schar von skalaren Aufgaben finden, wofür uns die Methoden der gewöhnlichen Optimierung zur Verfügung stehen.

Das unterstreicht die Bedeutung von Aussagen des folgenden Typs:

Sei $y_0 \in Y$ ein effizienter Punkt für die Menge Y bezüglich K. Dann gibt es (unter gewissen weiteren Voraussetzungen) ein $\lambda \in K^*$, $\lambda \neq 0^p$, so daß y_0 den Ausdruck $\langle \lambda, y \rangle$ über Y minimiert.

Im Verlauf dieser Arbeit werden desöfteren Resultate vorgestellt, die vom Prinzip her solchen „Skalarisierungsregeln“ gleichkommen.

Eine andere Vorgehensweise, auf die wir zurückgreifen werden, ist die unter dem Namen von J. de Lagrange (1736–1813) bekannt gewordenen Methode der Multiplikatorenregeln. Diese dienen zur Behandlung von Optimierungsproblemen, wenn die Bedingungnen für die Zulässigkeit einer Lösung von spezieller funktionaler Gestalt sind. In ihrer ursprünglichen Form galten sie für die Optimierung einer reellen Funktion ϕ über einer Menge der Form

$$S = \{x \in X |\ \psi_1(x) = 0, \ldots, \psi_k(x) = 0\}. \tag{1.20}$$

Unter gewissen Voraussetzungen wurde nachgewiesen, daß x_0 genau dann eine Extremalstelle von ϕ

über S ist, wenn auch die „Lagrange–Funktion“

$$L(\lambda, x) = \phi(x) + \sum_{i=1}^{k} \lambda_i \psi_i(x) \tag{1.21}$$

in x_0 ein Extremum über $R^k \times X$ annimmt. Dadurch war es möglich, anstelle eines restringierten ein freies Optimierungsproblem zu bearbeiten. Gelöst werden konnte dieses Problem mit Hilfe der Differentialrechnung, wenn man die Differenzierbarkeit aller auftretenden Funktionen unterstellte.

Im Verlauf weiterer Arbeiten wurde die Idee, die Zielfunktion auf die oben beschriebene Weise zu erweitern, in vielfältigen Formen angewandt. Zusätzlich zu den Gleichungsrestriktionen konnten bald auch Nebenbedingungen in Gestalt von Ungleichungen $\psi_i(x) \leq 0$ $(i = k+1, \ldots, m)$ betrachtet werden, wenn man die entsprechenden Multiplikatoren $\lambda_{k+1}, \ldots, \lambda_m$ als nichtnegativ voraussetzte. Es war desweiteren möglich, die in den Ungleichungen auftretende Vergleichsrelation mit Hilfe eines allgemeinen Ordnungskegels auszudrücken, so wie wir es bei der Verallgemeinerung des Effizienzbegriffes von Pareto vorgeführt hatten. Andererseits erlaubten es die Einwicklungen in der Analysis, die Optimalität einer Lösung auf andere Weise als durch die klassische Ableitung zu charaktierisieren, wodurch die Voraussetzung der Differenzierbarkeit der Funktionen abgeschwächt werden konnte. Mit der Vielfalt der in der modernen Optimierungtheorie betrachteten Modelle und den unterschiedlichsten Chrakaterisierungsmöglichkeiten von Extremalität entstand eine solche Fülle von „Lagrangeschen Multiplikatorenregeln“, daß selbst ein Überblick über all diese einen sehr großen Aufwand erfordern würde.

Die in dieser Arbeit vorgestellten Optimalitätsbedingungen werden eine Kombination aus Skalarisierungs- und Multiplikatorenregeln, wie sie oben vorgestellt wurden, sein. Auch hierbei sollen auftretende Nebenbedingungen in funktionaler Form nach Art und Weise der Lagrangeschen Multiplikatorenregeln behandelt werden, indem wir anstelle der Zielfunktion eine die Nebenbedingungen umfassende Lagrange–Funktion betrachten.

Die Charakterisierung der Optimalität einer Lösung geschieht mit Hilfe sogenannter derivierter Mengen. Diese stellen einen recht allgemeinen Ableitungsbegriff dar. Im Kapitel 2 werden wir die entsprechenden Definitionen einführen. Ferner werde ich einen neuen Beweis für die grundlegende Multiplikatorenregel über derivierte Mengen in der multikriteriellen Optimierung vorstellen.

Im Kapitel 3 soll darauf eingangen werden, daß derivierte Mengen oft in Form konvexer Kegel auftreten. Es liegt daher nahe, gewisse Beziehungen zu anderen Kegeln, die ebenfalls Eigenschaften einer Ableitung aufweisen, zu untersuchen.

Kapitel 4 ist der Betrachtung von Multiplikatorenregeln für Näherungslösungen von Optimierungsaufgaben gewidmet. Ich werde unter Verwendung der Variationsprinzips von Ekeland eine Multiplikatorenregel über derivierte Mengen für Näherungslösungen herleiten. Dieses allgemeine Ergebnis wird in Spezialfällen diskutiert.

Abschließend werden wir uns in Kapitel 5 mit Regularitätsbedingungen beschäftigen. Diese sichern, daß der Multiplikator, der in der Lagrange–Funktion die Zielfunktion bewertet, nicht Null wird. In der Literatur fand sich keine solche Aussage für die im Kaptiel 2 betrachteten Multiplikatorenregeln. Ich habe solche Bedingungen hergeleitet. Bemerkenswert an diesen ist, daß die Regularität einer Lösung hierbei auch unter Einbeziehung der Zielfunktion festgestellt wird, während sich Regularitätsbedingungen für gewöhnlich ausschließlich auf Eigenschaften der zulässigen Menge stützen. Zwischen diesen und den von mir gefundenen Resultaten sollen Vergleiche gezogen werden.

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

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

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

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 $\leq$ 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 $\{d^1, \ldots, d^n\} \subset D$ 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

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

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*

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

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* $\omega : B_+^n(r) \to X$ $(B_+^n(r) := \{x \in R_+^n |\ \|x\| \le r\})$*, 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

$$\mathit{convcone}\,(D)$$

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.

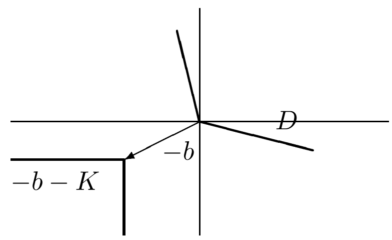


Abbildung 2.1: Trennbarkeit der Kegel D und K

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 $\{C_1, \dots, C_n\}$, wenn $x_0 \in \bigcap_{i=1}^n C_i$ gilt und Folgen $\{a_{ik}\}$ $(k = 1, 2, \dots)$ mit $a_{ik} \to 0$ für $k \to \infty$ existieren, so daß*

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

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 $f_{3l} = 0$ $(l = 1, \dots, m_3)$ 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 $\{b_k\} \subset R^m$ $(k = 1, 2, \dots)$, 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

$$f_1(x_n) - f_1(x_0) \in -int\, K_1 \qquad \forall n, \tag{2.24}$$

$$f_2(x_n) \in -K_2 \qquad \forall n, \tag{2.25}$$

$$f_3(x_n) = 0^{m_3} \qquad \forall n. \tag{2.26}$$

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 ($i = 0, 1, \dots, m_3$) zu D. Gemäß Definition 2.5 finden wir für das $(m+1)$-Tupel $\{a^0, a^1, \dots, a^{m_3}, 0^{m_1}, 0^{m_2}\}$ 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

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

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)

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

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. ■

Kapitel 3

Derivierte Kegel im Vergleich mit anderen Ableitungskegeln

3.1 Grundlegende Betrachtungen

Derivierte Mengen, so hatten wir im Kapitel 2 festgestellt, stellen eine Verallgemeinerung des Ableitungbegriffs dar. Dabei war allerdings noch nichts darüber ausgesagt worden, welche Beziehungen zwischen derivierten Mengen und speziellen Typen von Ableitungen bestehen. Insbesondere stellt sich die Frage, mit Hilfe welcher Ableitungsbegriffe man Punkte konstruieren kann, die sich als Elemente einer $K-$Ableitungsmenge (vgl. Definition 2.4) erweisen, um aus diesen nach Lemma 2.1 einen $K-$Ableitungskegel der gegebenen Funktion zu erzeugen. Wir wollen zunächst einige elementare Eigenschaften von $K-$Ableitungskegeln festhalten.

Aus der Definition eines $K-$Ableitungskegels folgt sofort, daß jeder Teilkegel eines solchen Kegels wiederum selbst ein $K-$Ableitungskegel ist. Wir müssen folglich davon ausgehen, daß es zu einer vorgegebenen Funktion mehrere verschiedene derivierte Kegel geben kann.

Berücksichtigt man die Verwendung der derivierten Mengen in notwendigen Optimalitätsbedingungen vom Typ des Theorems 2.2, so erkennt man die Notwendigkeit, zu einer gegebenen Aufgabe einen möglichst großen derivierten Kegel zu finden, um die Optimalitätsbedingung möglichst scharf formulieren zu können. Zwar ist es recht einfach einzusehen, daß, falls $D_1 \in R^m$ und $D_2 \in R^m$ zwei $K-$Ableitungskegel ein und derselben Funktion sind, auch deren Durchschnitt $D_1 \cap D_2$ ein solcher Kegel ist. Im allgemeinen ist jedoch ihre Vereinigung $D_1 \cup D_2$ keine K–Ableitungsmenge (und ebenso *convcone* $(D_1 \cup D_2)$ kein K–Ableitungskegel) mehr, wie Beispiel 3.1 zeigen wird. Auch kann man nicht davon ausgehen, daß in jedem Fall ein maximaler derivierter Kegel existiert, also ein K–Ableitungskegel, der alle anderen K–Ableitungskegel umfaßt.

Die Frage nach der Existenz von Ableitungskegeln kann leicht beantwortet werden: Grundsätzlich ist die Menge $\{0^m\}$ ein K–Ableitungskegel für eine Funktion $f: X \longrightarrow R^m$ im Punkt $x_0 \in X$, wie der Ordnungskegel K auch immer beschaffen ist. Mit diesem trivialen Kegel verliert die notwendige Optimalitätsbedingung von Theorem 2.2 allerdings jegliche Aussagekraft. Einen Ableitungskegel mit von Null verschiedenen Elementen findet man durch folgende Aussage.

Lemma 3.1 *Der (konvexe) Kegel $K \subseteq R^m$ ist stets ein K–Ableitungskegel für eine beliebige Funktion $f: X \longrightarrow R^m$ im Punkt $x_0 \in X$.*

Beweis. Es sei $\{d^1, d^2, \ldots, d^{m+1}\}$ ein $(m+1)$–Tupel von Punkten aus K. Wir setzen

$$\omega(t) := x_0 \qquad \forall t \in R_+^{m+1}$$

und

$$\varrho(t) := 0 \qquad \forall t \in R_+^{m+1}.$$

Dann gilt

$$f(\omega(t)) - f(x_0) - \sum_{i=1}^{m+1} t_i d^i - \|t\|\varrho(t) = -\sum_{i=1}^{m+1} t_i d^i \in -K \qquad \forall t \in R_+^{m+1}.$$

■

Mit Hilfe des Ordnungskegels K läßt sich ein bereits gefundener derivierter Kegel vergrößern.

Lemma 3.2 *Sei $D \subseteq R^m$ ein K–Ableitungskegel für eine Funktion $f: X \longrightarrow R^m$ im Punkt $x_0 \in X$. Dann ist auch die Menge $D + K$ ein K–Ableitungskegel für f in x_0.*

Beweis. Es sei $\{\bar{d}^1, \bar{d}^2, \ldots, \bar{d}^{m+1}\}$ ein $(m+1)$–Tupel von Punkten aus $D + K$. Dann gibt es Punkte $d^i \in D$ und $k^i \in K$ $(i = 1, \ldots, m+1)$ mit

$$\bar{d}^i = d^i + k^i \qquad \forall (i = 1, \ldots, m+1).$$

Für das $(m+1)$–Tupel $\{d^1, d^2, \ldots, d^{m+1}\} \subset D$ gibt es Größen r, ω und ϱ mit Eigenschaften gemäß Definition 2.5. Dann gilt aber auch

$$\begin{aligned} & f(\omega(t)) - f(x_0) - \sum_{i=1}^{m+1} t_i \bar{d}^i - \|t\|\varrho(t) \\ = \; & -\sum_{i=1}^{m+1} t_i k^i + \left[f(\omega(t)) - f(x_0) - \sum_{i=1}^{m+1} t_i d^i - \|t\|\varrho(t) \right] \\ \in \; & -\sum_{i=1}^{m+1} t_i k^i - K \\ \subseteq \; & -K \qquad \forall t \in B_+^{m+1}(r). \end{aligned}$$

■

Im abschließenden Lemma wird eine Grenze für die Größe von nichttrivialen Ableitungskegeln gegeben. Enthält eine Ableitungsmenge nämlich eine vollständige Gerade in ihrem Inneren, so ist der gesamte Raum ein Ableitungskegel. Oder anders ausgedrückt: Falls sich zeigen läßt, daß nicht der ganze Raum ein Ableitungskegel für die gegebene Funktion sein kann, so darf keine Ableitungsmenge dieser Funktion eine Gerade im Inneren enthalten.

Wir werden diese Aussage später verwenden, um für gewisse Ableitungsmengen nachzuweisen, daß eine weitere Vergrößerung der Menge nicht möglich ist, weil sonst der oben beschriebene Fall eintreten würde.

Lemma 3.3 *Enhält eine K–Ableitungsmenge $D \subseteq R^m$ einer Funktion $f : X \longrightarrow R^m$ im Punkt $x_0 \in X$ ein Element d mit $d \in int\, D$ und $-d \in D$, so ist der gesamte Raum R^m K–Ableitungskegel für f im Punkt x_0.*

Beweis. Nach Lemma 2.1 ist *convcone* (D) ein K–Ableitungskegel für f im Punkt x_0. Wegen $d \in int\, D$ gibt es eine Nullumgebung $V \in R^m$ mit

$$V = (d + V) + (-d) \in D + D \subseteq convcone\,(D).$$

Daraus folgt $convcone\,(D) = R^m$. ■

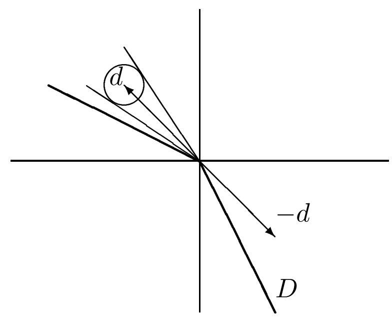


Abbildung 3.1: „überstumpfer“ Ableitungskegel

3.2 Vergleich mit speziellen Ableitungsbegriffen

3.2.1 Ableitungen in Form von Funktionalen

Wir wollen im folgenden untersuchen, ob man für eine Funktion spezielle K–Ableitungsmengen angeben kann, wenn von dieser Funktion gewisse Differenzierbarkeitseigenschaften bekannt sind. In der Tat lassen sich aus der Existenz einiger Ableitungstypen Schlußfolgerungen auf die zugehörigen derivierten Mengen ziehen.

Wir beginnen mit der Untersuchung von Funktionen, für die eine Richtungsableitung existiert.

Satz 3.4 *Sei X eine nichtleere Teilmenge eines linearen, topologischen Raumes $\mathcal{X}$. Die Funktion $f:\ X \longrightarrow R^m$ besitze in $x_0 \in X$ für ein $x \in \mathcal{X}$ die (einseitige) Richtungsableitung*

$$f'_+(x_0, x) := \lim_{\tau \downarrow 0} \frac{f(x_0 + \tau x) - f(x_0)}{\tau}.$$

Dann ist die (einelementige) Menge $\{f'_+(x_0, x)\}$ eine K–Ableitungsmenge für f in x_0 und

$$cone\,(f'_+(x_0, x)) + K \tag{3.1}$$

ein K–Ableitungskegel für f in x_0.

Bemerkung 3.1 *Wegen der positiven Homogenität der Richtungsableitung läßt sich der konische Abschluß in (3.1) ausdrücken als*

$$cone\,(f'_+(x_0, x)) = \{f'_+(x_0, \mu x) |\ \mu \geq 0\}.$$

Beweis. Zu zeigen ist, daß das n–Tupel $\{d, d, \ldots, d\}$ mit $d := f'_+(x_0, x)$ die Anforderungen von Definition 2.4 erfüllt.

Wir setzen

$$\omega(t) := x_0 + \left(\sum_{i=1}^{n} t_i\right) x \qquad \forall t \in R^m_+$$

und

$$\varrho(t) := \begin{cases} \dfrac{\sum_{i=1}^n t_i}{\|t\|} \left(\dfrac{f(x_0 + \sum_{i=1}^n t_i x) - f(x_0)}{\sum_{i=1}^n t_i} - d \right), & \text{wenn} \quad t \neq 0 \\ 0^m, & \text{wenn} \quad t = 0 \end{cases} \qquad \forall t \in R^m_+.$$

Offenbar sind ω und ϱ im Punkt $t = 0$ stetig und es gilt

$$f(\omega(t)) - f(x_0) - \sum_{i=1}^{n} t_i d - \|t\|\varrho(t) = 0^m \qquad \forall t \in R^m_+.$$

Somit sind die Kegel $cone(f'_+(x_0,x))$ (nach Lemma 2.1) und $cone(f'_+(x_0,x))+K$ (nach Lemma 3.2) K–Ableitungskegel für f in x_0. ■

Beispiel 3.1 *In diesem Beispiel werden aus zwei Richtungsableitungen zwei Ableitungskegel erzeugt, von denen keiner den anderen enthält. Wir sehen, daß ihre Vereinigung keine Ableitungsmenge mehr sein kann.*
$X = R^2$, $x_0 = (0,0)$; $m_1 = 1$, $m_2 = 1$, $m_3 = 0$; $K = R^2_+$;

$$\begin{aligned} f_1(x) &:= \begin{cases} x_1^2, & \textit{wenn} \quad x_1 \leq 0; \\ -x_1, & \textit{wenn} \quad x_1 > 0; \end{cases} \\ f_2(x) &:= x_1 + x_2^2 \end{aligned}$$

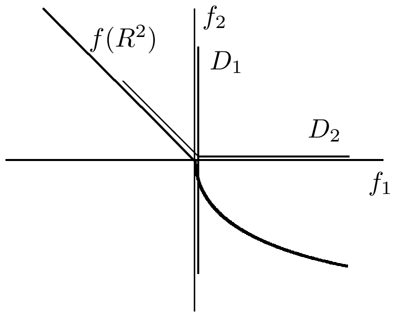


Abbildung 3.2: zu Beispiel 3.1

Es ergeben sich als Richtungsableitungen

$$f'_+((0,0),(-1,0)) = \begin{pmatrix} 0 \\ -1 \end{pmatrix},$$
$$f'_+((0,0),(1,0)) = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Mithin sind nach Satz 3.4 die Kegel

$$D_1 = \left\{ \mu \begin{pmatrix} 0 \\ -1 \end{pmatrix} \middle| \; \mu \geq 0 \right\} + R^2_+$$

und

$$D_2 = \left\{ \mu \begin{pmatrix} -1 \\ 1 \end{pmatrix} \middle| \; \mu \geq 0 \right\} + R^2_+,$$

beide für sich genommen, K–Ableitungskegel für f in x_0. Ihre Vereinigung

$$D_1 \cup D_2 = \left[\left\{ \mu \begin{pmatrix} 0 \\ -1 \end{pmatrix} \middle| \; \mu \geq 0 \right\} \cup \left\{ \mu \begin{pmatrix} -1 \\ 1 \end{pmatrix} \middle| \; \mu \geq 0 \right\} \right] + R^2_+$$

ist allerdings keine K–Ableitungsmenge für f in x_0. Wäre dies der Fall, so wäre wegen Lemma 3.3 der ganze Raum R^2 ein Ableitungskegel. Die Annahme, daß beispielsweise der Punkt $(-1,-1)$ Element eines Ableitungskegels für die gegebene Funktion ist, führt zu einem Widerspruch.

Letztes Beispiel zeigt, daß, obwohl die Strahlen $\{\mu(0,-1)\}$ und $\{\mu(-1,1)\}$ $(\mu \geq 0)$ Bilder von Richtungsableitungen und somit derivierte Kegel sind, ihre Vereinigung keine Ableitungsmenge sein muß. Richtungsableitungen, die für unterschiedliche Richtungen berechnet wurden, „passen“ also unter Umständen nicht in ein und dieselbe derivierte Menge.

Damit unterschiedliche Richtungsableitungen in einer gemeinsamen derivierten Menge liegen können, sind zusätzliche Voraussetzungen daran zu stellen, wie der Wert Ableitung von der Richtung abhängt. Das oben beschriebene Phänomen der Unverträglichkeit der einseitigen Richtungsableitungen tritt beispielsweise bei Gâteaux–differenzierbaren Funktionen nicht mehr auf. In diesem Fall ist die Menge $\{f'(x_0, x) |\ x \in \mathcal{X}\}$ aller Punkte, die sich als Bild der Gâteaux–Ableitung ergeben, insgesamt ein derivierter Kegel, wie Satz 3.5 zeigen wird. Dies liegt daran, daß bei der Gâteaux–Ableitung eine lineare Abhängigkeit von der Richtung besteht.

Beispiel 3.2 *Derivierte Kegel können auch existieren, wenn keine Richtungsableitungen vorhanden sind.*
$X = R^2$, $x_0 = (0,0)$; $m_1 = 1$, $m_2 = 1$, $m_3 = 0$; $K = R^2_+$;

$$\begin{aligned} f_1(x) &:= x_1; \\ f_2(x) &:= \begin{cases} x_1 \sin\left(\frac{1}{x_1}\right) + x_2^2, & \textit{wenn} \quad x_1 \neq 0; \\ x_2^2, & \textit{wenn} \quad x_1 = 0. \end{cases} \end{aligned}$$

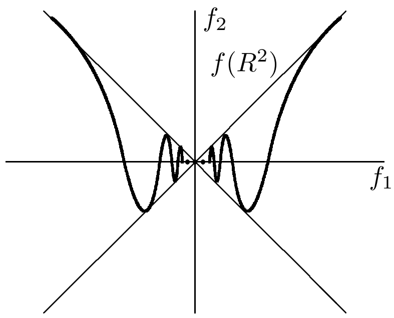


Abbildung 3.3: zu Beispiel 3.2

Für alle Richtungen (x_1, x_2) mit $x_1 \neq 0$ existiert im Punkt $(0,0)$ keine Richtungsableitung. Es läßt sich lediglich

$$f'_+((0,0),(0,x_2)) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \forall x_2 \in R$$

berechnen. Mit Hilfe von Satz 3.4 erhält man also nur R^2_+ als K–Ableitungskegel für f in x_0. Hingegen ist die Menge $\{\mu(-1,1)\}$ $(\mu \geq 0)$ eine K–Ableitungsmenge für diese Funktion und somit

$$\left\{ \mu \begin{pmatrix} -1 \\ 1 \end{pmatrix} \middle|\ \mu \geq 0 \right\} + R^2_+$$

ein K–Ableitungskegel für f in x_0. Dieser Kegel enthält Elemente, die nicht in R^2_+ enthalten sind.

Satz 3.5 *Sei X eine nichtleere Teilmenge eines linearen, topologischen Raumes $\mathcal{X}$. Die Funktion $f:\ X \longrightarrow R^m$ besitze in $x_0 \in X$ die Gâteaux–Ableitung*

$$f'(x_0, x) := \lim_{\tau \to 0} \frac{f(x_0 + \tau x) - f(x_0)}{\tau} \qquad (x \in \mathcal{X}).$$

Dann ist die Menge

$$f'(x_0, \mathcal{X}) + K \tag{3.2}$$

ein K–Ableitungskegel für f in x_0.

Beweis. Wir zeigen zunächst, daß $f'(x_0, \mathcal{X})$ ein K–Ableitungskegel für f in x_0 ist. Dazu betrachten wir ein n–Tupel $\{d^1, d^2, \ldots, d^n\} \subset f'(x_0, \mathcal{X}) + K$ mit

$$d^i = f'(x_0, x^i) \qquad (i = 1, \ldots, n).$$

Wir setzen

$$\omega(t) := x_0 + \sum_{i=1}^{n} t_i x^i \qquad \forall t \in R^n_+$$

und

$$\varrho(t) := \begin{cases} \dfrac{f(x_0 + \sum_{i=1}^n t_i x^i) - f(x_0)}{\|t\|} - \dfrac{\sum_{i=1}^n t_i d^i}{\|t\|}, & \text{wenn} \quad t \neq 0 \\ 0^m, & \text{wenn} \quad t = 0 \end{cases} \qquad \forall t \in R^n_+.$$

Die Funktion ω ist im Punkt $t = 0$ stetig. Dasselbe gilt für ϱ: Wegen der Linearität von $f'(x_0, .)$ ist nämlich

$$\frac{1}{\|t\|} \sum_{i=1}^{n} t_i d^i = f'\left(x_0, \frac{1}{\|t\|} \sum_{i=1}^{n} t_i x^i\right) \qquad \forall t \neq 0.$$

Also folgt

$$\lim_{\tau \downarrow 0} \frac{f(x_0 + \tau \frac{1}{\|t\|} \sum_{i=1}^n t_i x^i) - f(x_0)}{\tau} - \frac{1}{\|t\|} \sum_{i=1}^{n} t_i d^i = 0^m \qquad \text{gleichmäßig für alle } t \in R^n_+ \setminus \{0\}.$$

Die Gleichmäßigkeit der Konvergenz folgt aus der Kompaktheit der Menge

$$\left\{ \frac{1}{\|t\|} \sum_{i=1}^{n} t_i d^i \,\middle|\, t \in R^n_+ \setminus \{0\} \right\} \in R^m.$$

Somit gibt es für jedes $\varepsilon > 0$ ein (von t unabhängiges) $\tau_0 > 0$, so daß für alle $\tau \in (0, \tau_0]$ gilt

$$\left\| \frac{f(x_0 + \tau \frac{1}{\|t\|} \sum_{i=1}^n t_i x^i) - f(x_0)}{\tau} - \frac{1}{\|t\|} \sum_{i=1}^{n} t_i d^i \right\| < \varepsilon \qquad \forall t \in R^n_+ \setminus \{0\}.$$

Also folgt für alle $t \in R^m_+$ mit $\|t\| \in (0, \tau_0]$

$$\begin{aligned} &\left\| \frac{f(x_0 + \sum_{i=1}^n t_i x^i) - f(x_0)}{\|t\|} - \frac{1}{\|t\|} \sum_{i=1}^{n} t_i d^i \right\| \\ = &\left\| \frac{f(x_0 + \|t\| \frac{1}{\|t\|} \sum_{i=1}^n t_i x^i) - f(x_0)}{\|t\|} - \frac{1}{\|t\|} \sum_{i=1}^{n} t_i d^i \right\| < \varepsilon \qquad \forall t \in B^n_+(\tau_0) \setminus \{0\}, \end{aligned}$$

wodurch die Stetigkeit von ϱ in $t = 0$ gezeigt ist. Aufgrund der Konstruktion haben wir

$$f(\omega(t)) - f(x_0) - \sum_{i=1}^{n} t_i d^i - \|t\| \varrho(t) = 0^m \qquad \forall t \in R^m_+.$$

Mit $f'(x_0, \mathcal{X})$ ist nach Lemma 3.2 auch $f'(x_0, \mathcal{X}) + K$ ein K–Ableitungskegel für f in x_0. ■

Korollar 3.6 *Für Gâteaux–differenzierbare Funktionen f ist der Kegel (3.2) der größte K–Ableitungskegel in x_0, das heißt, jeder K–Ableitungskegel für f in x_0 ist in (3.2) enthalten.*

Beweis. Im Fall $f'(x_0, \mathcal{X}) + K = R^m$ ist die Aussage trivial. Anderenfalls enthält der Dualkegel von $f'(x_0, \mathcal{X}) + K$ von Null verschiedene Punkte. Es existiert also ein $y \in [f'(x_0, \mathcal{X}) + K]^*$ mit $y \neq 0^m$.
Angenommen, es gäbe einen K–Ableitungskegel für f an der Stelle x_0, der den Kegel $f'(x_0, \mathcal{X}) + K$ umfaßt und darüberhinaus ein Element $d \notin f'(x_0, \mathcal{X}) + K$ enthält. Dann gilt $\langle y, d\rangle < 0$.
Für das n–Tupel $\{d, d, \ldots, d\}$ existieren Größen $r > 0$, $\omega : B^n_+(r) \longrightarrow X$ und $\varrho : B^n_+(r) \longrightarrow R^m$, die die in Definition 2.5 genannten Eigenschaften besitzen.

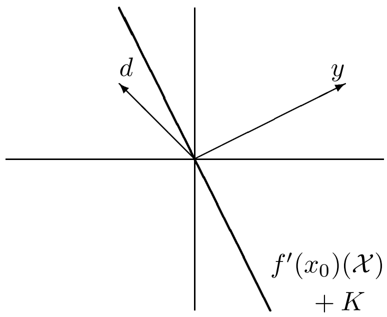


Abbildung 3.4: $d \notin f'(x_0, \mathcal{X}) + K$

Insbesondere gilt für hinreichend kleine $t \in R^n_+$

$$f(\omega(t)) - f(x_0) \in \sum_{i=1}^{n} t_i d + \vec{o}(\|t\|) - K,$$

wobei $\frac{\vec{o}(\|t\|)}{\|t\|} \longrightarrow 0^m$, wenn $t \downarrow 0$. Daraus folgt

$$\langle y, f(\omega(t)) - f(x_0)\rangle \leq \sum_{i=1}^{n} t_i \langle y, d\rangle + o(\|t\|),$$

wobei $\frac{o(\|t\|)}{\|t\|} \longrightarrow 0$, wenn $t \downarrow 0$. Also gilt für $t \neq 0$

$$\frac{\langle y, f(\omega(t)) - f(x_0)\rangle}{\|t\|} \leq \frac{\sum_{i=1}^{n} t_i}{\|t\|}\langle y, d\rangle + \frac{o(\|t\|)}{\|t\|}.$$

Weil $\frac{\sum_{i=1}^{n} t_i}{\|t\|}$ für $t \in R^n_+ \setminus \{0\}$ beschränkt ist und $\langle y, d\rangle < 0$ gilt, gibt es eine Konstante $c > 0$ mit $\frac{\sum_{i=1}^{n} t_i}{\|t\|}\langle y, d\rangle \leq -c \ \forall t \in R^n_+ \setminus \{0\}$. Es folgt

$$\frac{\langle y, f(\omega(t)) - f(x_0)\rangle}{\|t\|} \leq -c + \frac{o(\|t\|)}{\|t\|}.$$

Also finden wir eine Zahl $r_1 > 0$, so daß

$$\frac{\langle y, f(\omega(t)) - f(x_0)\rangle}{\|t\|} \leq -\frac{2}{3}c \qquad \forall t \in B^n_+(r_1) \setminus \{0\}. \tag{3.3}$$

Wegen der Gâteaux–Differenzierbarkeit von f gilt aber andererseits

$$f(\omega(t)) - f(x_0) = f'(x_0, \omega(t) - x_0) + \vec{o}(\|t\|)$$

oder, unter Beachtung von $\langle y, f'(x_0, \mathcal{X})\rangle \geq 0$,

$$\langle y, f(\omega(t)) - f(x_0)\rangle \geq o(\|t\|).$$

Das heißt, es gibt eine Zahl $r_2 > 0$, so daß

$$\frac{\langle y, f(\omega(t)) - f(x_0)\rangle}{\|t\|} \geq -\frac{1}{3}c \qquad \forall t \in B^n_+(r_2) \setminus \{0\}. \tag{3.4}$$

Dies ist ein Widerspruch zu Aussage (3.3). Es gibt also keinen K–Ableitungskegel für f an der Stelle x_0, der ein Element enthält, das nicht in $f'(x_0, \mathcal{X}) + K$ liegt. ■

Bemerkung 3.2 *In [5] befindet sich eine weitere Aussage über einen Ableitungstyp, der zur Konstruktion von Ableitungskegeln verwendet werden kann. Das entsprechende Resultat ist in der vorliegenden Arbeit als Theorem 2.3 aufgeführt. Es werden hierbei (rechtsseitige obere) Dini–Ableitungen betrachtet. Die Abhängigkeit der Ableitung von der Richtung wird dadurch realisiert, daß die Dini–Ableitungen von oben durch eine konvexe Funktion der Richtung beschränkt sein müssen.*

3.2.2 Geometrisch konstruierte Ableitungsmengen

Das lokale Verhalten einer Funktion $f: X \longrightarrow R^m$ $(X \subseteq \mathcal{X})$ läßt sich außer durch Ableitungsfunktionale auch durch gewisse Tangentialmengen beschreiben, die am Graphen der Funktion im Produktraum $\mathcal{X} \times R^m$ konstruiert werden. Bei der Konstruktion solcher Tangentialmengen stützte man sich zunächst auf die geometrische Vorstellung von der Tangente, die an einen Punkt der Menge, die den Graphen der Funktion repräsentiert, angelegt wird. Wir wollen untersuchen, welche Zusammenhänge zwischen derivierten Mengen und derartigen Tangentialmengen oder auch -kegeln, bestehen.

Der wohl bekannteste Tangentialkegel ist der von Bouligand um 1930 eingeführte Kontingentkegel. Wir benutzen die folgende Definition aus [41].

Definition 3.1 *Der Kontingentkegel $K_C(x_0)$ an eine nichtleere Teilmenge C eines linearen topologischen Raums $\mathcal{X}$ im Punkt $x_0 \in C$ ist definiert durch*

$$K_C(x_0) := \bigcap_{\substack{V \in \mathcal{N}(0) \\ \tau_0 > 0}} \bigcup_{\tau \in (0, \tau_0]} \left[\frac{C - x_0}{\tau} + V\right]. \tag{3.5}$$

(Hierbei beschreibt $\mathcal{N}(0)$ eine Umgebungsbasis des Punktes $0 \in \mathcal{X}$.)

Vertrauter ist die Charakterisierung der Elemente des Kontingentkegels durch Folgen. Die Äquivalenz beider Definitionen in endlichdimensionalen Räumen liefert die folgende Aussage (vgl. [42]):

Lemma 3.7 *Ein Punkt $x \in R^n$ ist genau dann ein Element des Kontingentkegels $K_C(x_0)$ an die Menge $C \subseteq R^n$ im Punkt $x_0 \in C$, wenn es Folgen $\{\tau^k\} \subset R_+ \setminus \{0\}$ und $\{x^k\} \subset R^n$ $(k = 1, 2, \ldots)$ gibt mit $\tau^k \downarrow 0$ und $x^k \to x$ $(k \to \infty)$, für die gilt*

$$x_0 + \tau^k x^k \in C \qquad (k = 1, 2, \ldots).$$

Um Aussagen über Ableitungseigenschaften von Funktionen $f : X \longrightarrow R^m$ $(X \subseteq \mathcal{X})$ in einem Punkt $x_0 \in X$ treffen zu können, betrachtet man üblicherweise den Tangentialkegel an den Graphen der Funktion

$$graph\, f := \{(x, f(x)) \in \mathcal{X} \times R^m |\ x \in X\} \tag{3.6}$$

in Punkte $(x_0, f(x_0)) \in graph\, f$. Nach dieser Vorgehensweise ist der Tangentialkegel für eine Funktion eine Teilmenge des Raumes $\mathcal{X} \times R^m$, in dem der Graph von f liegt. Derivierte Mengen gehören hingegen dem Bildraum R^m der Funktion f an. Um Vergleiche anstellen zu können, sollten wir uns daher auf den Bildraumanteil des Tangentialkegels konzentrieren.

Hierzu vereinbaren wir folgende Schreibweise für die Projektion des Kontingentkegels in den Bildraum R^m der Funktion f:

$$PK_{graph\, f}(x_0, f(x_0)) := \{y \in R^m |\ \exists x \in X : \ (x, y) \in K_{graph\, f}(x_0, f(x_0))\} \tag{3.7}$$

Bekanntlich ist in vielen Fällen der Bouligandsche Kontingentkegel $K_{graph\, f}(x_0, f(x_0))$ kein konvexer Kegel (vgl. etwa [8]). Dies trifft dann natürlich auch auf den Bildraumanteil $PK_{graph\, f}(x_0, f(x_0))$ dieses Kegels zu.

Ein K–Ableitungskegel ist nach Definition 2.5 jedoch stets konvex. Wir können daher nicht erwarten, daß ohne weitere Voraussetzungen der gesamte Kontingentkegel ein Ableitungskegel ist. Beispiel 3.3 bestätigt dies. Dieses Beispiel liefert uns allerdings eine Vermutung, die in Satz 3.8 festgehalten ist.

Beispiel 3.3 *Der gesamte Kegel $PK_{graph\ f}(x_0, f(x_0))$ ist kein Ableitungskegel, wohl aber gewisse konvexe Teilkegel davon.*
$X = R^2$, $x_0 = (0,0)$; $m_1 = 1$, $m_2 = 1$, $m_3 = 0$; $K = R^2_+$;

$$\begin{aligned} f_1(x) &:= \begin{cases} x_1^2, & \text{wenn } x_1 \leq 0; \\ -\sqrt{x_1}, & \text{wenn } x_1 > 0; \end{cases} \\ f_2(x) &:= x_1 + x_2^2 \end{aligned}$$

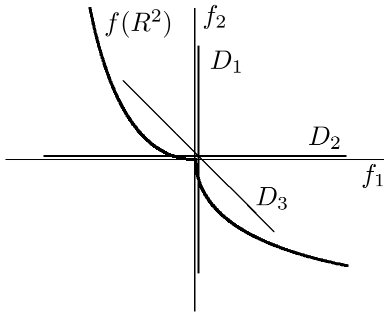


Abbildung 3.5: zu Beispiel 3.3

Man berechnet als Kontingentkegel

$$PK_{graph\ f}(x_0, f(x_0)) = \left[\left\{\mu\begin{pmatrix}0\\-1\end{pmatrix}\middle|\ \mu \geq 0\right\} \cup \left\{\mu\begin{pmatrix}-1\\0\end{pmatrix}\middle|\ \mu \geq 0\right\}\right] + R^2_+.$$

Dieser ist allerdings keine K–Ableitungsmenge für f in x_0. Ansonsten wäre wegen Lemma 3.3 der ganze Raum R^2 ein Ableitungskegel. Die Annahme, daß beispielsweise der Punkt $(-1,-1)$ Element eines Ableitungskegels für die gegebene Funktion wäre, führt zu einem Widerspruch. Hingegen erweisen sich die (konvexen) Teilkegel

$$\begin{aligned} D_1 &:= \left\{\mu\begin{pmatrix}0\\1\end{pmatrix}\middle|\ \mu \in R\right\} + R^2_+, \\ D_2 &:= \left\{\mu\begin{pmatrix}1\\0\end{pmatrix}\middle|\ \mu \in R\right\} + R^2_+, \\ D_3 &:= \left\{\mu\begin{pmatrix}1\\-1\end{pmatrix}\middle|\ \mu \in R\right\} + R^2_+ \end{aligned}$$

von $PK_{graph\ f}(x_0, f(x_0))$ sämtlich als K–Ableitungskegel für f in x_0. Wie der folgende Satz zeigt, ist dieses Phänomen von allgemeiner Natur.

Satz 3.8 *Sei $G \subseteq K_{graph\ f}(x_0, f(x_0))$ ein konvexer Teilkegel des Kontingentkegels an den Graphen einer gegebenen Funktion $f : X \longrightarrow R^m$ im Punkt $(x_0, f(x_0))$.*
Dann ist die Bildraumprojektion von G

$$D := \{y \in R^m |\ \exists x \in X : (x,y) \in G\}$$

ein K–Ableitungskegel für f in x_0.

Beweis. Offenbar ist D als Projektion eines konvexen Kegels wiederum ein konvexer Kegel. Wir betrachten ein n–Tupel $\{d^1, d^2, \ldots, d^n\} \subset D$. Dann gibt es Punkte $x^1, x^2, \ldots, x^n \in X$, so daß

$$\{(x^1, d^1), (x^2, d^2), \ldots, (x^n, d^n)\} \subset G.$$

Wegen der Konvexität von G ist dann auch

$$\frac{1}{\|t\|} \sum_{i=1}^{n} t_i (x^i, d^i) \in G \qquad t \in R^n_+ \setminus \{0\}.$$

Weil G eine Teilmenge von

$$K_{graph\, f}(x_0, f(x_0)) = \bigcap_{\substack{(U,V) \in \mathcal{N}_{\mathcal{X}}(0) \times \mathcal{N}_{R^m}(0) \\ \tau_0 > 0}} \bigcup_{\tau \in (0, \tau_0]} \left[\frac{graph\, f - (x_0, f(x_0))}{\tau} + U \times V \right]$$

ist, existiert zu allen $\tau_0 > 0$ sowie zu allen Nullumgebungen $U \subset \mathcal{X}$, $V \subset R^m$ ein $\tau \in (0, \tau_0]$ mit

$$\tau \frac{1}{\|t\|} \sum_{i=1}^{n} t_i (x^i, d^i) \in graph\, f - (x_0, f(x_0)) + \tau (U \times V).$$

Das heißt, für zu jedem τ_0, U und V gibt es ein $(\tilde{x}, f(\tilde{x})) \in graph\, f$ mit

$$\begin{aligned} \tilde{x} - x_0 &\in \tau \left(\frac{1}{\|t\|} \sum_{i=1}^{n} t_i x^i - U \right) \subseteq \tau_0 \left(\frac{1}{\|t\|} \sum_{i=1}^{n} t_i x^i - U \right), \\ f(\tilde{x}) - f(x_0) &\in \tau \left(\frac{1}{\|t\|} \sum_{i=1}^{n} t_i d^i - V \right) \subseteq \tau_0 \left(\frac{1}{\|t\|} \sum_{i=1}^{n} t_i d^i - V \right). \end{aligned}$$

Speziell mit $\tau_0 = \|t\|$ und $V = B^m(\|t\|)$ folgt für jedes $t \in R^n_+ \setminus \{0\}$ die Existenz eines $x_t \in X$ mit

$$x_t - x_0 \in \sum_{i=1}^{n} t_i x^i - \|t\| U,$$

$$\left\| f(x_t) - f(x_0) - \sum_{i=1}^{n} t_i d^i \right\| \leq \|t\|^2.$$

Setzen wir

$$\omega(t) := \begin{cases} x_t, & \text{wenn} \quad t \neq 0 \\ x_0, & \text{wenn} \quad t = 0 \end{cases} \qquad \forall t \in R^n_+$$

und

$$\varrho(t) := \begin{cases} \frac{1}{\|t\|} \left(f(x_t) - f(x_0) - \sum_{i=1}^{n} t_i d^i \right), & \text{wenn} \quad t \neq 0 \\ 0^m, & \text{wenn} \quad t = 0 \end{cases} \qquad \forall t \in R^n_+,$$

so sind ω und ϱ im Punkt $t = 0$ stetig und es gilt

$$f(\omega(t)) - f(x_0) - \sum_{i=1}^{n} t_i d^i - \|t\| \varrho(t) = 0^m \qquad \forall t \in R^n_+.$$

■

Um der oben festgestellten und oft auftretenden Nichtkonvexität des Kontingentkegels zu entgehen, wurde in den letzten vierzig Jahren viel Anstrengung in die Suche nach konvexen Ersatzkegeln des Kontingentkegels investiert. Dabei stand natürlich das Bestreben im Vordergrund, möglichst viele der Tangentialeigenschaften des Bouligandschen Kegels zu bewahren. Viele dieser neuen Konstruktionen sind konvexe Teilkegel des Kontingentkegels, wie etwa der Tangentenkegel von Clarke (1975).

Die unten angegebene Definition des Clarkeschen Kegels stammt aus [41]. Clarkes eigentliche Formulierung, die zu dieser äquivalent ist, findet man in [8] und [9].

Definition 3.2 *Der Clarkesche Tangentenkegel* $T_C(x_0)$ *an eine nichtleere Teilmenge* C *eines linearen topologischen Raums* $\mathcal{X}$ *im Punkt* $x_0 \in C$ *ist definiert durch*

$$T_C(x_0) := \bigcap_{V \in \mathcal{N}(0)} \bigcup_{\substack{U \in \mathcal{N}(x_0) \\ \tau_0 > 0}} \bigcap_{\substack{x' \in C \cap U \\ \tau \in (0, \tau_0]}} \left[\frac{C - x'}{\tau} + V \right]. \tag{3.8}$$

(Hierbei beschreibt $\mathcal{N}(0)$ *eine Umgebungsbasis des Punktes* $0 \in \mathcal{X}$ *und* $\mathcal{N}(x_0)$ *eine Umgebungsbasis des Punktes* $x_0 \in \mathcal{X}$*.)*

Auch diese Definition kann mit Hilfe von Folgen interpretiert werden (vgl. [42]).

Lemma 3.9 *Ein Punkt* $x \in R^n$ *ist genau dann ein Element des Clarkeschen Tangentenkegels* $T_C(x_0)$ *an die Menge* $C \subseteq R^n$ *im Punkt* $x_0 \in C$*, wenn es für alle Folgen* $\{\tau^k\} \subset R_+ \setminus \{0\}$ *und* $\{x'^k\} \subset C$ $(k = 1, 2, \ldots)$ *mit* $\tau^k \downarrow 0$ *und* $x'^k \to x_0$ $(k \to \infty)$ *stets Punkte* $x^k \in R^n$ *mit* $x^k \to x$ *und*

$$x'^k + \tau^k x^k \in C \qquad (k = 1, 2, \ldots)$$

gibt.

Auch beim Clarkeschen Tangentenkegel an den Graphen einer Funktion $f : X \longrightarrow R^m$ ist es sinnvoll, die Projektion in den Bildraum R^m der Funktion f zu betrachten.

$$PT_{graph\, f}(x_0, f(x_0)) := \{y \in R^m \mid \exists x \in X : \ (x, y) \in T_{graph\, f}(x_0, f(x_0))\} \tag{3.9}$$

Dann läßt sich zeigen:

Satz 3.10 *Die Bildraumprojektion* $PT_{graph\, f}(x_0, f(x_0))$ *des Clarkeschen Tangentenkegels an den Graphen einer gegebenen Funktion* $f : X \longrightarrow R^m$ *im Punkt* $(x_0, f(x_0))$ *ist ein* K*–Ableitungskegel für* f *in* x_0*.*

Beweis. Dies folgt sofort aus Satz 3.8, wenn man beachtet, daß $T_{graph\ f}(x_0, f(x_0))$ ein konvexer Teilkegel von $K_{graph\ f}(x_0, f(x_0))$ ist. ■

Bemerkung 3.3 *Ein Analogon zu Satz 3.10 gilt für jeden Tangentialkegel, der ein konvexer Teilkegel des Kontingentkegels ist.*

Da es unter den konvexen Teilkegeln eines nichtkonvexen Kegels für gewöhnlich keinen größten gibt, bestätigt sich auch hier, daß zu einer gegebenen Funktion im allgemeinen kein größter Ableitungskegel existiert. Insbesondere ist die Bildraumprojektion des Clarkeschen Kegels kein maximaler Ableitungskegel, wie Beispiel 3.4 zeigt. Eine Aussage in Form von Korallar 3.6 gilt hier also nicht.

Beispiel 3.4 *Wir betrachten noch einmal die Daten von Beispiel 3.3. In diesem Beispiel ist der Clarkesche Tangentenkegel ein echter Teilkegel des Kontingentkegels. Es gibt derivierte Kegel, die sowohl echte Teilkegel des Kontingentkegel als auch echte Oberkegel des Clarkeschen Tangentenkegels sind.*

Man berechnet als Clarkeschen Tangentenkegel

$$PT_{graph\ f}(x_0, f(x_0)) = R^2_+.$$

Als K–Ableitungskegel hatten wir in Beispiel 3.3 den Kegel

$$D_3 = \left\{ \mu \begin{pmatrix} 1 \\ -1 \end{pmatrix} \middle| \ \mu \in R \right\} + R^2_+$$

bestimmt. Es gilt

$$PK_{graph\ f}(x_0, f(x_0)) \overset{\supset}{\neq} D_3 \overset{\supset}{\neq} PT_{graph\ f}(x_0, f(x_0)).$$

Kapitel 4

Lagrangesche Multiplikatorenregeln für Näherungslösungen

4.1 Verfahren zur Herleitung von Multiplikatorenregeln für Näherungslösungen

4.1.1 Direkte Verfahren

Im Abschnitt 2.3 wurde bereits angedeutet, wie man notwendige Optimalitätsbedingungen in Form Lagrangescher Multiplikatorenregeln mit Anwendung der Bildraumtechnik erhält: Man betrachte im Bildraum zum einen die „erreichbare“ Menge, das heißt alle Bilder zulässiger Elemente; im untenstehenden Modell wäre das $f(X)$. Zum anderen konstruiere man zu allen Bildpunkten $f(x_0)$ eine (üblicherweise konvexe) Menge, die die Funktionswerte enthält, die „besser“ als $f(x_0)$ sind; diese Menge wird unten mit $W(x_0)$ bezeichnet. Die Konstruktion erfolgt so, daß beide Mengen genau dann disjunkt sind, wenn x_0 eine Optimallösung der betrachteten Aufgabe ist.

Das Ziel ist es nun, beide Mengen durch ein lineares, stetiges Funktional zu trennen. Um einen Trennungssatz anwenden zu können, ist man meist gezwungen, die Menge $f(X)$ durch eine geeignete konvexe Menge zu ersetzen, die die wichtigsten Eigenschaften der Menge beibehält. Diese Konvexifizierung stellt für gewöhnlich den aufwendigsten Teil der Arbeit mit der Bildraumtechnik dar.

Hat man ein solches trennendes Funktional gefunden, liefert dies im wesentlichen bereits den gewünschten Langrangeschen Multiplikator. Eine recht gute Darstellung dieser Vorgehensweise findet man in [39].

In einigen Fällen führt diese Methode auch bei näherungsweise optimalen Lösungen zum Erfolg. Um dies vorzuführen, wollen wir die Herangehensweise, die Pourciau in [39] zur Herleitung einer konvexen Multiplikatorenregel benutzt, auf den Fall von Näherungslösungen derselben Aufgabe übertragen.

Betrachtet wird das folgende Optimierungsproblem

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

$$\text{über} \quad x \in \{x \in R^n |\ f_{21}(x) \leq 0, \ldots, f_{2m_2}(x) \leq 0, x \in X\},$$

wobei X eine konvexe Teilmenge des R^nist und $f_1, f_{21}, \ldots, f_{2m_2}$ reellwertige konvexe Funktionen auf X sind. Dann gilt das folgende Resultat.

Theorem 4.1 (Konvexe Multiplikatorenregel) *Sei $x_0 \in X$ eine Minimallösung des obigen Problems. Dann existiert ein Multiplikator $\lambda = (\lambda_1, \lambda_{21}, \ldots, \lambda_{2m_2}) \in R^{1+m_2} \backslash \{0^{1+m_2}\}$ mit $\lambda_1 \geq 0$, $\lambda_{2k} \geq 0$ $(k = 1, \ldots, m_2)$ und*

$$\begin{aligned} \langle \lambda, f(x) \rangle &\geq \langle \lambda, f(x_0) \rangle \qquad \forall x \in X, \\ \lambda_{2k} f_{2k}(x_0) &= 0 \qquad (k = 1, \ldots, m_2). \end{aligned}$$

Zum Beweis wird in [39] für jedes $x_0 \in X$ eine konvexe Menge

$$W(x_0) := \left\{ y = (y_1, y_{21}, \ldots, y_{2m_2}) \in R^{1+m_2} \;\middle|\; \begin{array}{rcl} y_1 & < & f_1(x_0) \\ y_{21} & \leq & 0 \\ \vdots & & \\ y_{2m_2} & \leq & 0 \end{array} \right\} \tag{4.1}$$

definiert. Es ist klar, daß x_0 genau dann eine Minimallösung ist, wenn

$$f(X) \cap W(x_0) = \emptyset. \tag{4.2}$$

Sei nun x_ε lediglich eine ε–Minimallösung ($\varepsilon \geq 0$), das heißt, x_ε ist zulässig und es gilt

$$f_1(x_\varepsilon) \leq f_1(x) + \varepsilon$$

für jedes x mit $x \in X$ und $f_{2k}(x) \leq 0$ $(k = 1, \ldots, m_2)$. Offensichtlich ist x_ε eine ε–Minimallösung genau dann, wenn

$$f(X) \cap [-(\varepsilon, 0^{m_2}) + W(x_\varepsilon)] = \emptyset. \tag{4.3}$$

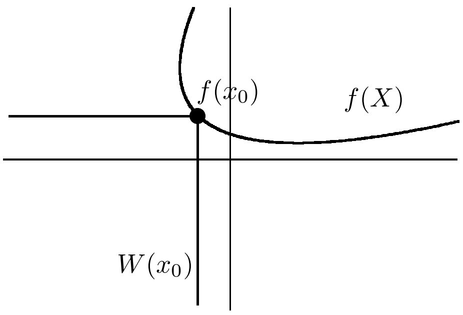


Abbildung 4.1: Trennbarkeit wegen $f(X) \cap W(x_0) = \emptyset$ im Optimalpunkt.

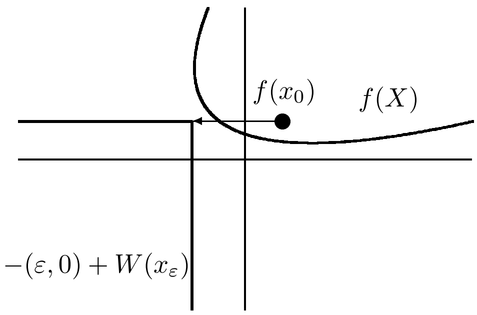


Abbildung 4.2: Trennbarkeit wegen $f(X) \cap [-(\varepsilon, 0)) + W(x_\varepsilon)] = \emptyset$ im ε–Optimalpunkt.

Unter Ausnutzung von (4.3) anstelle von (4.2) können wir den Beweis analog zu Pourciau weiterführen und erhalten schließlich eine Multiplikatorenregel für Näherungslösungen.

Zur Konvexifizierung der Menge $f(X)$ betrachten wir an ihrer Stelle die Menge $f(X) + R_+^{1+m_2}$. Offenbar ist diese Menge konvex und Gleichung (4.3) ist äquivalent mit

$$[f(X) + R_+^{1+m_2}] \cap [-(\varepsilon, 0^{m_2}) + W(x_\varepsilon)] = \emptyset. \tag{4.4}$$

Die Anwendung eines Trennungssatzes liefert die Existenz eines Multiplikators $\lambda = (\lambda_1, \lambda_{21}, \ldots, \lambda_{2m_2}) \in R^{1+m_2} \setminus \{0^{1+m_2}\}$ mit

$$\langle \lambda, f(x) + k \rangle \geq \langle \lambda, -(\varepsilon, 0^{m_2}) + y \rangle \qquad \forall x \in X \quad \forall k \in R_+^{1+m_2} \quad \forall y \in W(x_\varepsilon),$$

woraus folgt

$$\langle \lambda, f(x) \rangle \geq -\varepsilon \lambda_1 + \langle \lambda, y \rangle \qquad \forall x \in X \quad \forall y \in cl\, W(x_\varepsilon). \tag{4.5}$$

Wegen $f(x_\varepsilon) \in cl\, W(x_\varepsilon)$ folgt aus (4.5)

$$\langle \lambda, f(x) \rangle \geq \langle \lambda, f(x_\varepsilon) \rangle - \varepsilon \lambda_1 \qquad \forall x \in X. \tag{4.6}$$

Wäre $\lambda_1 < 0$ oder $\lambda_{2k} < 0$ $(k = 1, \ldots, m_2)$, so kann die rechte Seite der Ungleichung (4.5) bei geeigneter Wahl von $y \in cl\, W(x_\varepsilon)$ beliebig groß werden und jede linksseitige Konstante $\langle \lambda, f(x) \rangle$ (x fest in X) übersteigen. Mithin muß gelten

$$\lambda_1 \geq 0, \qquad \lambda_{2k} \geq 0 \quad (k = 1, \ldots, m_2). \tag{4.7}$$

Wegen

$$\begin{pmatrix} f_1(x_\varepsilon) \\ f_{21}(x_\varepsilon) \\ \vdots \\ \frac{1}{2} f_{2k}(x_\varepsilon) \\ \vdots \\ f_{2m_2}(x_\varepsilon) \end{pmatrix} \in cl\, W(x_\varepsilon) \qquad (k = 1, \ldots, m_2)$$

ergibt (4.5) mit $x = x_\varepsilon$ schließlich

$$\frac{1}{2} \lambda_{2k} f_{2k}(x_\varepsilon) \geq -\varepsilon \lambda_1 \qquad (k = 1, \ldots, m_2)$$

und wegen $\lambda_{2k} \geq 0$ und $f_{2k}(x_\varepsilon) \leq 0$

$$-2\varepsilon\lambda_1 \leq \lambda_{2k} f_{2k}(x_\varepsilon) \leq 0 \qquad (k = 1, \ldots, m_2). \tag{4.8}$$

Mit (4.6), (4.7) und (4.8) haben wir in Analogie zur konvexen Multiplikatorenregel folgende Regel für Näherungslösungen bewiesen.

Satz 4.2 *Sei $x_\varepsilon \in X$ eine ε–Minimallösung ($\varepsilon \geq 0$) des gegebenen Problems. Dann existiert ein Multiplikator $\lambda = (\lambda_1, \lambda_{21}, \ldots, \lambda_{2m_2}) \in R^{1+m_2} \setminus \{0^{1+m_2}\}$ mit $\lambda_1 \geq 0$, $\lambda_{2k} \geq 0$ ($k = 1, \ldots, m_2$) und*

$$\langle \lambda, f(x) \rangle \geq \langle \lambda, f(x_\varepsilon) \rangle - \varepsilon \lambda_1 \qquad \forall x \in X,$$

$$-2\varepsilon\lambda_1 \leq \lambda_{2k} f_{2k}(x_\varepsilon) \leq 0 \qquad (k = 1, \ldots, m_2).$$

Bemerkung 4.1 *Die in der Komplementaritätsbedingung auftretende Schranke $-2\varepsilon\lambda_1$ ist durch weitgehend analoge Übertragung der Pourciauschen Beweiskette entstanden. Die Schranke läßt sich jedoch verschärfen: Es gilt nämlich*

$$\begin{pmatrix} f_1(x_\varepsilon) \\ f_{21}(x_\varepsilon) \\ \vdots \\ f_{2,k-1}(x_\varepsilon) \\ 0 \\ f_{2,k+1}(x_\varepsilon) \\ \vdots \\ f_{2m_2}(x_\varepsilon) \end{pmatrix} \in cl\, W(x_\varepsilon) \qquad (k = 1, \ldots, m_2),$$

woraus folgt

$$\lambda_{2k} f_{2k}(x_\varepsilon) \geq -\varepsilon\lambda_1 \qquad (k = 1, \dots, m_2) \tag{4.9}$$

Dieser Schluß könnte auch im Pourciauschen Originalbeweis für $\varepsilon = 0$ zur Anwedung kommen.

4.1.2 Verfahren unter Verwendung des Variationsprinzips von Ekeland

Eine andere Variante zur Herleitung von Optimalitätsbedingungen für Näherungslösungen basiert auf der Anwendung des Ekelandschen Variationsprinzips. Dieses wurde erstmalig im Jahre 1974 veröffentlicht [11]. Eine Formulierung lautet:

Theorem 4.3 *Sei $\mathcal{X}$ ein Banachraum und sei $f : \mathcal{X} \longrightarrow R \cup \{+\infty\}$ eine eigentliche, unterhalbstetige, nach unten beschränkte Funktion.*
Dann existiert für jedes $\varepsilon > 0$, für jeden Punkt $x_0 \in \mathcal{X}$ mit

$$f(x_0) \leq \inf_{x \in \mathcal{X}} f(x) + \varepsilon$$

und jedes $\lambda > 0$ ein Punkt $x_\varepsilon \in \mathcal{X}$, so daß

(i) $f(x_\varepsilon) \leq f(x_0)$,

(ii) $\|x_\varepsilon - x_0\| \leq \lambda$,

(iii) $f(x) > f(x_\varepsilon) - \frac{\varepsilon}{\lambda}\|x - x_\varepsilon\| \ \forall x \neq x_\varepsilon$.

Für restringierte Aufgaben verwendet man die folgende Form des Satzes.

Korollar 4.4 *Sei $\mathcal{X}$ ein Banachraum und sei $f : \mathcal{X} \longrightarrow R$ eine unterhalbstetige Funktion, die auf einer nichtleeren, abgeschlossenen Teilmenge S von $\mathcal{X}$ nach unten beschränkt ist.*
Dann existiert für jedes $\varepsilon > 0$ und für jeden Punkt $x_0 \in S$ mit

$$f(x_0) \leq \inf_{x \in S} f(x) + \varepsilon$$

ein Punkt $x_\varepsilon \in S$, so daß

(i) $f(x_\varepsilon) \leq f(x_0)$,

(ii) $\|x_\varepsilon - x_0\| \leq \sqrt{\varepsilon}$,

(iii) $f(x) \geq f(x_\varepsilon) - \sqrt{\varepsilon}\|x - x_\varepsilon\| \ \forall x \in S$.

Beweis. Man ersetze f durch die eigentliche, unterhalbstetige und nach unten beschränkte Funktion

$$\bar{f}(x) := \begin{cases} f(x), & \text{wenn} \quad x \in S \\ +\infty, & \text{wenn} \quad x \in \mathcal{X} \setminus S. \end{cases}$$

und wende Theorem 4.3 mit $\lambda = \sqrt{\varepsilon}$ an. ■

Die Herleitung einer Multiplikatorenregel mit Hilfe des Variationsprinzips beruht im wesentlichen auf folgender Überlegung: Nach (iii) ist x_ε eine Minimallösung einer im Vergleich zum Ausgangsproblem etwas gestörten Aufgabe. Mithin erfüllt x_ε eine der (meist bekannten) notwendigen Optimalitätsbedingungen für das gestörte Problem. Man untersucht nun, ob sich aus diesen Bedingungen Rückschlüsse auf das ursprünglich betrachtete Problem ziehen lassen.

Diese Vorgehensweise soll beispielhaft an der konvexen Optimierungsaufgabe aus dem vorangegangenen Abschnitt erläutert werden. Zusätzlich zu den dort getroffenen Vereinbarungen sei vorausgesetzt, daß die Menge $S := \{x \in X |\ f_{21}(x) \leq 0, \ldots, f_{2m_2}(x) \leq 0\}$ abgeschlossen und die Funktion f_1 über S nach unten beschränkt ist.

Gegeben sei $\varepsilon > 0$. Sei $x_0 \in S$ eine ε–Minimallösung des Problems, das heißt

$$f_1(x_0) \leq \inf_{x \in S} f_1(x) + \varepsilon.$$

Dann gibt es ein x_ε, so daß die Aussagen (i), (ii) und (iii) von Korollar 4.4 erfüllt sind. Nach (iii) ist x_ε eine Minimalstelle der Funktion $f_1(.) + \sqrt{\varepsilon}\|. - x_\varepsilon\|$ über S. Da diese Funktion konvex ist, können wir die konvexe Multiplikatorenregel anwenden und erhalten die Existenz eines Multiplikators $\lambda = (\lambda_1, \lambda_{21}, \ldots, \lambda_{2m_2}) \in R^{1+m_2} \setminus \{0^{1+m_2}\}$ mit $\lambda_1 \geq 0$, $\lambda_{2k} \geq 0$, $\lambda_{2k} f_{2k}(x_\varepsilon) = 0$ $(k = 1, \ldots, m_2)$ und

$$\left\langle \begin{pmatrix} \lambda_1 \\ \lambda_{21} \\ \vdots \\ \lambda_{2m_2} \end{pmatrix}, \begin{pmatrix} f_1(x) \\ f_{21}(x) \\ \vdots \\ f_{2m_2}(x) \end{pmatrix} + \sqrt{\varepsilon} \begin{pmatrix} \|x - x_\varepsilon\| \\ 0 \\ \vdots \\ 0 \end{pmatrix} \right\rangle \geq \langle \lambda, f(x_\varepsilon) \rangle \qquad \forall x \in X,$$

was gleichbedeutend ist mit

$$\langle \lambda, f(x) \rangle \geq \langle \lambda, f(x_\varepsilon) \rangle - \sqrt{\varepsilon}\lambda_1 \|x - x_\varepsilon\| \qquad \forall x \in X.$$

Somit ist gezeigt:

Satz 4.5 *Sei $x_0 \in S$ eine ε–Minimallösung $(\varepsilon > 0)$ des gegebenen Problems. Dann existiert ein Punkt $x_\varepsilon \in S$, so daß* *(i)* $f_1(x_\varepsilon) \leq f_1(x_0)$,

(ii) $\|x_\varepsilon - x_0\| \leq \sqrt{\varepsilon}$,

(iii) $\exists \lambda = (\lambda_1, \lambda_{21}, \ldots, \lambda_{2m_2}) \in R^{1+m_2} \setminus \{0^{1+m_2}\}$ *mit* $\lambda_1 \geq 0$, $\lambda_{2k} \geq 0$ $(k = 1, \ldots, m_2)$ *und*

$$\begin{aligned} \langle \lambda, f(x) \rangle &\geq \langle \lambda, f(x_\varepsilon) \rangle - \sqrt{\varepsilon} \lambda_1 \|x - x_\varepsilon\| \qquad \forall x \in X, \\ \lambda_{2k} f_{2k}(x_\varepsilon) &= 0 \qquad (k = 1, \ldots, m_2). \end{aligned}$$

Optimalitätsbedingungen, die mit Hilfe des Ekelandschen Variationsprinzips erzeugt werden, finden ihre Anwendung meist in folgender Form.

Korollar 4.6 *Zu jedem* $\varepsilon > 0$ *gibt es eine* ε*–Minimallösung* $x_\varepsilon \in S$ *des gegebenen Problems, für die ein Multiplikator* $\lambda = (\lambda_1, \lambda_{21}, \ldots, \lambda_{2m_2}) \in R^{1+m_2} \setminus \{0^{1+m_2}\}$ *existiert mit* $\lambda_1 \geq 0$, $\lambda_{2k} \geq 0$ $(k = 1, \ldots, m_2)$ *und*

$$\begin{aligned} \langle \lambda, f(x) \rangle &\geq \langle \lambda, f(x_\varepsilon) \rangle - \sqrt{\varepsilon} \lambda_1 \|x - x_\varepsilon\| \qquad \forall x \in X, \\ \lambda_{2k} f_{2k}(x_\varepsilon) &= 0 \qquad (k = 1, \ldots, m_2). \end{aligned}$$

Beweis. Sicher gibt es zu gegebenem $\varepsilon > 0$ ein $x_0 \in S$ mit

$$f_1(x_0) \leq \inf_{x \in S} f_1(x) + \varepsilon.$$

Dann existiert es ein $x_\varepsilon \in S$, das (i), (ii) und (iii) von Satz 4.5 erfüllt. Aus (i) folgt die ε–Optimalität von x_ε:

$$f_1(x_\varepsilon) \leq f_1(x_0) \leq \inf_{x \in S} f_1(x) + \varepsilon.$$

Aus (iii) folgen die restlichen Behauptungen. ■

Bemerkung 4.2 *Die Qualität der Optimalitätsbedingung in Korollar 4.6 ist eine andere als die von Satz 4.2. Während letztere Aussage eine notwendige Bedingung ist, das heißt, jede* ε*–optimale Lösung muß die dort angegebenen Eigenschaften haben, besagt Korollar 4.6 nur:*
Unter all den Punkten, die die geforderten Eigenschaften besitzen, befindet sich wenigstens einer, der eine ε*–optimale Lösung der Aufgabe ist.*

4.2 Zur Struktur von Multiplikatorenregeln für Näherungslösungen

Obwohl die im vorangegangenen Abschnitt hergeleiteten Multiplikatorenregeln relativ rasch aus dem Ekelandschen Prinzip folgen, erlauben sie uns doch bereits einen Einblick in die fundamentale Struktur von Lagrangeschen Multiplikatorenregeln für Näherungslösungen.

Notwendige Optimalitätsbedingungen in Form von Multiplikatorenregeln bestehen, wie allgemein bekannt, im wesentlichen aus zwei Aussagen:

- einer *Stationaritätsbedingung*, die etwa das Verschwinden der Ableitung der Lagrange–Funktion im Fall differenzierbarer Funktionen unterstellt, und
- einer *Komplementaritätsbedingung*, die fordert, daß die Multiplikatoren, die mit nichtaktiven Nebenbedingungen korrespondieren, Null sein müssen.

Diese Struktur tritt im multikriteriellen Fall, aber auch im Fall skalarer Optimierungsaugaben auf.

Erwartungsgemäß stellen notwendige Optimalitätsbedingungen für Näherungslösungen schwächere Anforderungen auf als die vergleichbaren Bedingungen für exakte Lösungen. Diese Abschwächung betrifft bei Multiplikatorenregeln im allgemeinen beide Fundamentalaussagen. So wird bei Satz 4.2 im Vergleich zur konvexen Multiplikatorenregel (Theorem 4.1) die Stationaritätsbedingung

$$\langle \lambda, f(x) \rangle \geq \langle \lambda, f(x_0) \rangle \qquad \forall x \in X$$

ersetzt durch

$$\langle \lambda, f(x) \rangle \geq \langle \lambda, f(x_\varepsilon) \rangle - \varepsilon\lambda_1 \qquad \forall x \in X.$$

Die Näherungslösung x_ε erweist sich also als eine lediglich annähernde Minimalstelle der Lagrange–Funktion. Weiterhin wird die Komplementaritätsbedingung

$$\lambda_{2k} f_{2k}(x_0) = 0 \qquad (k = 1, \ldots, m_2)$$

abgeschwächt zu

$$-\varepsilon\lambda_1 \leq \lambda_{2k} f_{2k}(x_\varepsilon) \leq 0 \qquad (k = 1, \ldots, m_2),$$

was besagt, daß der komplementäre Schlupf $\lambda_{2k} f_{2k}(x_\varepsilon)$ durchaus negative Werte, die aber nicht weit von Null entfernt liegen, annehmen kann.

Multiplikatorenregeln für Näherungslösungen bestehen also aus

- einer leicht verletzten Stationaritätsbedingung und
- einer leicht verletzten Komplementaritätsbedingung.

Die „leichte Verletzung“ läßt sich durch den Toleranzparameter ε, der die Güte der Optimalität der Näherungslösung beschreibt, quantifizieren.

Die oben betrachteten Sätze haben für die Größe der Abweichung von den exakten Optimalitätsbedingungen starre Schranken vorgegeben. Es ist möglich, diese Schranken innerhalb gewisser Spielräume flexibel zu gestalten, so daß sich der Grad der Abweichung von der strengen Stationarität und der Grad der Abweichung bei der Komplementarität wechselseitig kompensieren. Ein interessantes Resultat über dieses Phänomen findet man in einer Arbeit von Strodiot, Nguyen und Heukemes [46]. Dieses besagt, daß eine ε–Näherungslösung eine Multiplikatorenregel erfüllt, bei der die Stationaritätsbedingung in der Größenordnung eines Parameters ε_1 und die Komplementaritätsbedingung in der Größenordnung von ε_2 verletzt sind, wobei $\varepsilon_1 \geq 0$, $\varepsilon_2 \geq 0$ gilt und $\varepsilon_1 + \varepsilon_2 \leq \varepsilon$ eingehalten wird.

Um dieses Resultat konkret zu formulieren, betrachten wir

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

$$\text{über} \quad x \in \{x \in R^n |\ f_{21}(x) \leq 0, \ldots, f_{2m_2}(x) \leq 0\},$$

wobei $f_1, f_{21}, \ldots, f_{2m_2}$ reellwertige konvexe Funktionen auf R^n sind. Nach [40] bezeichnen wir für eine konvexe Funktion $g : R^n \longrightarrow R$ die Menge

$$\partial_\varepsilon g(x_0) := \{x^* \in R^n |\ \langle x^*, x - x_0 \rangle \leq g(y) - g(x) + \varepsilon\}$$

als ε–Subdifferential von g im Punkt $x_0 \in R^n$.

Für diese Aufgabe wird in [46] gezeigt:

Theorem 4.7 *Für das obige Problem sei die Regularitätsbedingung*

$$\exists \bar{x} \in R^n : \qquad f_{2k}(\bar{x}) < 0 \quad (k = 1, \ldots, m_2)$$

erfüllt. Der Punkt $x_\varepsilon \in R^n$ ist eine ε–Minimallösung ($\varepsilon \geq 0$) des Problems genau dann, wenn Konstanten $\varepsilon_1 \geq 0$, $\varepsilon_{2k} \geq 0$ ($k = 1, \ldots, m_2$) mit

$$\varepsilon_1 + \sum_{k=1}^{m_2} \varepsilon_{2k} \leq \varepsilon \tag{4.10}$$

und Multiplikatoren $\varepsilon_{2k} \geq 0$ $(k = 1, \ldots, m_2)$ *existieren, so daß*

$$0 \in \partial_{\varepsilon_1} f_1(x_\varepsilon) + \sum_{k=1}^{m_2} \partial_{\varepsilon_2 k}(\lambda_{2k} f_{2k})(x_\varepsilon), \tag{4.11}$$

$$\varepsilon_1 + \sum_{k=1}^{m_2} \varepsilon_{2k} - \varepsilon \leq \sum_{k=1}^{m_2} \lambda_{2k} f_{2k}(x_\varepsilon) \leq 0. \tag{4.12}$$

Bemerkung 4.3 *Bei der Wahl von* $\varepsilon_1 = \varepsilon_{21} = \cdots = \varepsilon_{2m_2} = 0$ *ergibt sich aus (4.11), (4.12) eine strenge Erfüllung der Stationaritätsbedingung und eine maximale Verletzung der Komplementaritätsbedingung:*

$$0 \in \partial f_1(x_\varepsilon) + \sum_{k=1}^{m_2} \partial(\lambda_{2k} f_{2k})(x_\varepsilon),$$

$$-\varepsilon \leq \sum_{k=1}^{m_2} \lambda_{2k} f_{2k}(x_\varepsilon) \leq 0.$$

Im anderen Extremfall ergibt sich bei Wahl von ε_1 *und* ε_{2k} $(k = 1, \ldots, m_2)$ *mit* $\varepsilon_1 + \sum_{k=1}^{m_2} \varepsilon_{2k} = \varepsilon$ *die größtmögliche Störung bei der Stationarität und die strenge Komplementarität:*

$$0 \in \partial_{\varepsilon_1} f_1(x_\varepsilon) + \sum_{k=1}^{m_2} \partial_{\varepsilon_2 k}(\lambda_{2k} f_{2k})(x_\varepsilon),$$

$$\lambda_{2k} f_{2k}(x_\varepsilon) = 0 \qquad (k = 1, \ldots, m_2).$$

In der jüngeren Literatur findet man in einer Arbeit von Yokoyama [55] eine Anwendung dieses Theorems auf multikriterielle Aufgaben.

Bei Multiplikatorenregeln für Näherungslösungen, die mit Hilfe des Ekelandschen Variationsprinzips hergeleitet wurden, zeigt sich die Veränderung im Vergleich zur Regel für exakte Lösungen nur in einer Abschwächung der Stationaritätsbedingung, während der komplementäre Schlupf weiterhin verschwindet. Dies ist typisch für derartige Aussagen, da hier nicht Näherungslösungen der ursprünglichen Aufgabe, sondern exakte Lösungen einer gestörten Aufgabe betrachtet werden.

4.3 Multiplikatorenregeln für Näherungslösungen unter Benutzung derivierter Mengen

Nach einer allgemeinen Betrachtung der Struktur näherungsweiser Optimalitätsbedingungen wollen wir uns nun auf solche Multiplikatorenregeln konzentrieren, bei denen derivierte Mengen Verwendung finden. Es ist mir gelungen, eine mit der Multiplikatorenregel von Breckner (Theorem 2.2) zusammenhängende Optimalitätsbedingung zu beweisen, die für Näherungslösungen der betrachteten Aufgabe gilt. Diese Herleitung erfolgte unter Verwendung des Variationprinzips von Ekeland.

Eine Anwendung des Ekelandschen Prinzips auf derivierte Mengen allgemeiner Natur erwies sich zunächst als undurchführbar. Es mußten zusätzliche Anforderungen an die Ableitungmengen gestellt werden, um überhaupt Ergebnisse zu erzielen. Einige dieser Zusatzbedingungen werden im folgenden diskutiert. Es zeigt sich weiterhin, daß diese Anforderungen die Allgemeinheit in nur geringem Maße einschränken; inbesondere werden wir nachweisen, daß bereits die in Theorem 2.3 benutzte derivierte Menge einer der Zusatzbedingungen genügt.

Da die Zielfunktion der zugrunde liegenden Aufgabe vektorwertig ist, benötigen wir eine Erweiterung des Ekelandschen Prinzips für solche Funktionen. Eine derartige Erweiterung stammt von Tammer [47] aus dem Jahre 1992.

Hierbei betrachten wir annähernd effiziente Punkte der Menge $f(S)$, wobei S eine nichtleere, abgeschlossene Teilmenge eines Banachraumes $\mathcal{X}$ und $f : \mathcal{X} \longrightarrow R^p$ eine vektorwertige Funktion sind.

Gegeben seien ein konvexer Kegel $K \in R^p$ mit nichtleerem Inneren sowie ein Element $k^0 \in int\, K$. Sei weiterhin $\tilde{K} \supseteq K$ ein konvexer Oberkegel von K mit den Eigenschaften

$$cl\, \tilde{K} + (K \setminus \{0\}) \subseteq int\, \tilde{K}, \tag{4.13}$$

$$bd\, \tilde{K} + bd\, \tilde{K} \subseteq cl\, \tilde{K}. \tag{4.14}$$

Wir setzen voraus, daß die Funktion f auf $\mathcal{X}$ nach unten halbstetig bezüglich k^0 und $\tilde{K}$ ist, das heißt, die Menge

$$M_r := \{x \in \mathcal{X};\ f(x) \in rk^0 - cl\, \tilde{K}\} \tag{4.15}$$

ist abgeschlossen für jedes $r \in R$. Schließlich sei f auf S nach unten beschränkt, das heißt, es existiert ein $y \in R^p$ mit

$$f(S) \subset y + K. \tag{4.16}$$

Bemerkung 4.4 *Jeder konvexe, abgeschlossene Kegel $\tilde{K} \in R^p$ mit*

$$K \subseteq int\, \tilde{K}$$

erfüllt die Bedingungen (4.13) und (4.14). Diese Bedingungen erfassen jedoch eine größere Menge zulässiger Oberkegel. Im skalaren Fall ($p = 1$) kann beispielsweise $\tilde{K} = K = R_+$ gesetzt werden, so daß das gewöhnliche Ekelandsche Variationsprinzip ein Spezialfall der Erweiterung von Tammer ist.

Bemerkung 4.5 *Für $p = 1$ und $\tilde{K} = K = R_+$ sind die Forderungen (4.15) und (4.16) identisch mit der unteren Halbstetigkeit*

$$\text{Die Mengen } M_r := \{x \in \mathcal{X};\ f(x) \le r\} \text{ sind abgeschlossen für jedes } r.$$

beziehungsweise der Beschränktheit

$$f(x) \ge y \qquad \forall x \in S$$

der reellwertigen Funktion f.

Wir benutzen die folgende Modifikation des vektorwertigen Variationsprinzips. Sie ergibt sich als Folgerung von Theorem 4.1 in [47].

Theorem 4.8 *Sei $\mathcal{X}$ ein Banachraum und $f : \mathcal{X} \longrightarrow R^p$ eine vektorwertige Funktion definiert auf $\mathcal{X}$. Sei S eine nichtleere, abgeschlossene Teilmenge von $\mathcal{X}$. Im Raum R^p seien ein konvexer Kegel K mit nichtleerem Inneren sowie ein Element $k^0 \in int\, K$ gegeben. Weiterhin sei $\tilde{K} \supseteq K$ ein konvexer Oberkegel von K mit den Eigenschaften (4.13), (4.14). Die Funktion f sei auf $\mathcal{X}$ nach unten halbstetig bezüglich k^0 und $\tilde{K}$ und auf S nach unten beschränkt.*
Dann existiert für jedes $\varepsilon > 0$ und jedes $x_0 \in S$ mit

$$(f(x_0) - \varepsilon k^0 - \tilde{K} \setminus \{0\}) \cap f(S) = \emptyset$$

ein $x_\varepsilon \in S$ mit

(i) $(f(x_\varepsilon) - \varepsilon k^0 - K \setminus \{0\}) \cap f(S) = \emptyset$,

(ii) $\|x_\varepsilon - x_0\| \le \sqrt{\varepsilon}$,

(iii) $(f_{\varepsilon k^0}(x_\varepsilon) - K \setminus \{0\}) \cap f_{\varepsilon k^0}(S) = \emptyset$,

wobei $f_{\varepsilon k^0}(x) := f(x) + \sqrt{\varepsilon}\|x - x_\epsilon\| k^0$.

Bemerkung 4.6 *Bemerkenswert ist, daß im vektorwertigen Fall eine zusätzliche Bedingung an den Punkt x_0 auftritt, die im skalaren Fall keine Entsprechung hat. Es genügt also nicht, daß $f(x_0)$ ein effizienter Punkt der Menge $f(S)$ bezüglich K ist, vielmehr wird die Effizienz von $f(x_0)$ bezüglich eines Oberkegels $\tilde{K}$ gefordert.*
Eine solche Eigenschaft wird in der Literatur als „eigentliche“ Effizienz von x_0 bezeichnet (vgl. [31], Definition 2.2.1).

Dieses Ekelandsche Varationsprinzip für vektorwertige Funktionen soll im folgenden zur Herleitung einer Multiplikatorenregel für Näherunglösungen, die sich auf die Verwendung derivierter Mengen stützt, dienen. Wir betrachten erneut das Optimierungsproblem, welches von Theorem 2.2 behandelt wird. Für die Anwendbarkeit von Theorem 4.8 müssen die Anforderungen an die Aufgabe allerdings etwas verschärft werden.
Betrachtet wird das Problem

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

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

unter den Voraussetzungen:

(V1) $\mathcal{X}$ ist ein reeller Banachraum.

(V2) $X \subseteq \mathcal{X}$ ist eine nichtleere Teilmenge von X. $K_1 \subseteq R^{m_1}$, $K_2 \subseteq R^{m_2}$ und $K_3 \subseteq R^{m_3}$ sind konvexe Kegel, K_1 und K_2 haben ein nichtleeres Inneres, K_2 und K_3 sind abgeschlossen.

(V3) Die zulässige Menge S ist abgeschlossen.

(V4) Ein Element $k^0 \in int\, K_1$ sei gegeben.

(V5) Gegeben sei weiterhin ein konvexer Oberkegel $\tilde{K} \supseteq K_1$ mit den Eigenschaften (4.13) und (4.14).

(V6) $f_1 : \mathcal{X} \longrightarrow R^{m_1}$, $f_2 : \mathcal{X} \longrightarrow R^{m_2}$, $f_3 : \mathcal{X} \longrightarrow R^{m_3}$.

(V7) f_1 ist auf $\mathcal{X}$ nach unten halbstetig bezüglich k^0 und $\tilde{K}$.

(V8) f_1 ist auf S nach unten beschränkt.

Die Aussage (iii) von Theorem 4.8 liefert einen effizienten Punkt der Menge $f_{\varepsilon k^0,1}(S)$ bezüglich K_1, wobei $f_{\varepsilon k^0} : X \longrightarrow R^m$ definiert ist durch

$$f_{\varepsilon k^0}(x) := f(x) + \sqrt{\epsilon}\|x - x_\epsilon\| \begin{pmatrix} k^0 \\ 0^{m_2} \\ 0^{m_3} \end{pmatrix}. \tag{4.17}$$

Um Theorem 2.2 anwenden zu können, müssen wir zunächst einen K–Ableitungskegel $D_{\varepsilon k^0}$ für $f_{\varepsilon k^0}$ an der Stelle x_ε konstruieren.

Dazu sei angenommen, es wäre eine K–Ableitungsmenge D für f an der Stelle x_ε gegeben. Wir nehmen weiterhin an, daß D die folgende Spezialbedingung erfüllt:

(S1) *Derivierte Menge mit linearer ω–Funktion.*

Jedem $d \in D$ kann ein $h \in \mathcal{X}$ zugeordnet werden, so daß gilt: Für jedes n–Tupel $\{d^1, \dots, d^n\} \subset D$ werden die Anforderungen von Definition 2.4 durch die Funktion

$$\omega(t) := x_\varepsilon + \sum_{i=1}^{n} t_i h^i \qquad \forall (t_1, \dots, t_n) \in B_+^n(r)$$

erfüllt, wobei h^i der zu d^i gehörige Vektor ist ($i = 1, \dots, n$).

Gegeben sei ein n–Tupel $\{d^1, \dots, d^n\} \subset D$. Für dieses existieren nach Definition 2.4 gewisse Größen $r > 0$, $\omega :\ B_+^n(r) \longrightarrow X$ und $\varrho :\ B_+^n(r) \longrightarrow R^m$ mit

$$f(\omega(t)) - f(x_\varepsilon) - \sum_{i=1}^{n} t_i d^i - \|t\| \varrho(t) \in -K \qquad \forall (t_1, \dots, t_n) \in B_+^n(r). \tag{4.18}$$

Gesucht ist nun ein n–Tupel $\{d_\varepsilon^1, \dots, d_\varepsilon^n\}$, so daß

$$f_{\varepsilon k^0}(\omega(t)) - f_{\varepsilon k^0}(x_\varepsilon) - \sum_{i=1}^{n} t_i d_\varepsilon^i - \|t\| \varrho(t) \in -K \qquad \forall (t_1, \dots, t_n) \in B_+^n(r). \tag{4.19}$$

Unter Beachtung von (S1) ergibt sich

$$\begin{aligned}
& f_{\varepsilon k^0}(\omega(t)) - f_{\varepsilon k^0}(x_\varepsilon) - \sum_{i=1}^{n} t_i d_\varepsilon^i - \|t\| \varrho(t) \\
= \;& f(\omega(t)) - f(x_\varepsilon) - \sum_{i=1}^{n} t_i d^i - \|t\| \varrho(t) + \sqrt{\varepsilon} \|\omega(t) - x_\varepsilon\| (k^0, 0^{m_2}, 0^{m_3}) + \sum_{i=1}^{n} t_i (d^i - d_\varepsilon^i) \\
= \;& f(\omega(t)) - f(x_\varepsilon) - \sum_{i=1}^{n} t_i d^i - \|t\| \varrho(t) + \sqrt{\varepsilon} \left\| \sum_{i=1}^{n} t_i h^i \right\| (k^0, 0^{m_2}, 0^{m_3}) + \sum_{i=1}^{n} t_i (d^i - d_\varepsilon^i) \\
& \forall (t_1, \dots, t_n) \in B_+^n(r).
\end{aligned}$$

Wegen $\left\| \sum_{i=1}^{n} t_i h^i \right\| \leq \sum_{i=1}^{n} t_i \|h^i\|$ für $(t_1, \dots, t_n) \in R_+^n$ folgt

$$\left\| \sum_{i=1}^{n} t_i h^i \right\| (k^0, 0^{m_2}, 0^{m_3}) \in \sum_{i=1}^{n} t_i \|h^i\| (k^0, 0^{m_2}, 0^{m_3}) - K \qquad \forall (t_1, \dots, t_n) \in R_+^n.$$

Hieraus und aus (4.18) folgt dann

$$\begin{aligned}
& f_{\varepsilon k^0}(\omega(t)) - f_{\varepsilon k^0}(x_\varepsilon) - \sum_{i=1}^{n} t_i d_\varepsilon^i - \|t\| \varrho(t) \\
\in \;& \sum_{i=1}^{n} t_i (d^i + \sqrt{\varepsilon} \|h^i\| (k^0, 0^{m_2}, 0^{m_3}) - d_\varepsilon^i) - K \qquad \forall (t_1, \dots, t_n) \in B_+^n(r).
\end{aligned}$$

Setzen wir also

$$d^i_\varepsilon := d^i + \sqrt{\varepsilon}\|h^i\|(k^0, 0^{m_2}, 0^{m_3}) \qquad (i = 1, \ldots, n),$$

so ist (4.19) erfüllt.

Mithin ist die Menge $D_{\varepsilon k^0} \subseteq R^m$, die definiert ist durch

$$d + \sqrt{\varepsilon}\|h\|(k^0, 0^{m_2}, 0^{m_3}) \in D_{\varepsilon k^0} \Longleftrightarrow d \in D,$$

eine K–Ableitungmenge für $f_{\varepsilon k^0}$ an der Stelle x_ε und nach Lemma 2.1 der Kegel *convcone* $(D_{\varepsilon k^0})$ ein K–Ableitungskegel für $f_{\varepsilon k^0}$ an der Stelle x_ε, sofern nur $D \in R^m$ eine K–Ableitungsmenge für f an der Stelle x_ε mit der Eigenschaft (S1) ist.

Auch in einigen anderen Spezialfällen können solche Transformationen angegeben werden, daß $\{d^1_\varepsilon, \ldots, d^n_\varepsilon\}$ die Bedingung (4.19) erfüllt.

(S2) *Derivierte Menge mit Lipschitz–stetiger ω–Funktion.*

Jedem $d \in D$ kann eine Konstante $L \geq 0$ zugeordnet werden, so daß gilt: Für jedes n–Tupel $\{d^1, \ldots, d^n\} \subset D$ gibt es eine Funktion ω nach Definition 2.4, die zusätzlich

$$\|\omega(t) - x_\varepsilon\| \leq \sum_{i=1}^{n} L^i t_i \qquad \forall (t_1, \ldots, t_n) \in B^n_+(r)$$

erfüllt, wobei L^i die zu d^i gehörige Konstante ist. Dann wird durch

$$d^i_\varepsilon := d^i + \sqrt{\varepsilon} L^i (k^0, 0^{m_2}, 0^{m_3}) \qquad (i = 1, \ldots, n)$$

die Bedingung (4.19) erfüllt.

(S3) *Derivierte Menge mit differenzierbarer ω–Funktion.*

Jedem $d \in D$ kann ein $\omega' \in \mathcal{X}$ zugeordnet werden, so daß gilt: Für jedes n–Tupel $\{d^1, \ldots, d^n\} \subset D$ gibt es eine Funktion ω nach Definition 2.4, die zusätzlich

$$\omega(t) = x_\varepsilon + \sum_{i=1}^{n} t_i \omega'^i + \|t\| s(t) \qquad \forall (t_1, \ldots, t_n) \in B^n_+(r)$$

mit $\lim_{t \downarrow 0} s(t) = 0$ erfüllt, wobei ω'^i der zu d^i gehörige Vektor ist. Dann wird durch

$$d^i_\varepsilon := d^i + \sqrt{\varepsilon}\|\omega'^i\|(k^0, 0^{m_2}, 0^{m_3}) \qquad (i = 1, \ldots, n)$$

zwar nicht die Bedingung (4.19) erfüllt, aber es gilt an deren Stelle eine Aussage der Form

$$f_{\varepsilon k^0}(\omega(t)) - f_{\varepsilon k^0}(x_\varepsilon) - \sum_{i=1}^{n} t_i d^i_\varepsilon - \|t\| \varrho_\varepsilon(t) \in -K \qquad \forall (t_1, \ldots, t_n) \in B^n_+(r),$$

wenn wir setzen

$$\varrho_\varepsilon(t) := \varrho(t) + \sqrt{\varepsilon}\|s(t)\|(k^0, 0^{m_2}, 0^{m_3}).$$

Es ist einzusehen, daß wegen $\lim_{t\downarrow 0} s(t) = 0$ mit ϱ auch ϱ_ε die Bedingung (C) von Definition 2.4 erfüllt.

Ein Resultat, das man unter Verwendung von (S1) erhält, ist das folgende.

Satz 4.9 *Für jedes $\varepsilon > 0$ und jedes $x_0 \in S$ mit*

$$(f_1(x_0) - \varepsilon k^0 - \tilde{K} \setminus \{0^{m_1}\}) \cap f_1(S) = \emptyset$$

existiert ein $x_\varepsilon \in S$ mit

(i) $(f_1(x_\varepsilon) - \varepsilon k^0 - K_1 \setminus \{0^{m_1}\}) \cap f_1(S) = \emptyset$,

(ii) $\|x_\varepsilon - x_0\| \leq \sqrt{\varepsilon}$,

(iii) Für jede K–Ableitungsmenge $D \subseteq R^m$ für f an der Stelle x_ε, die (S1) erfüllt, existiert ein Multiplikator $\lambda \in K_1^* \times K_2^* \times K_3^* \setminus \{0^m\}$ mit

$$\begin{aligned} \langle \lambda, d\rangle &\geq -\sqrt{\varepsilon}\|h\|\langle \lambda_1, k^0\rangle \qquad \forall d \in D, \\ \langle \lambda_2, f_2(x_\varepsilon)\rangle &= 0. \end{aligned}$$

Beweis. Nach Theorem 4.8 existiert ein $x_\varepsilon \in \mathcal{X}$ mit den Eigenschaften (i) und (ii) sowie

$$(f_{\varepsilon k^0,1}(x_\varepsilon) - K_1 \setminus \{0^{m_1}\}) \cap f_{\varepsilon k^0,1}(S) = \emptyset.$$

Daher gilt erst recht

$$(f_{\varepsilon k^0,1}(x_\varepsilon) - int\, K_1) \cap f_{\varepsilon k^0,1}(S) = \emptyset.$$

Nach den vorangegangenen Überlegungen ist die Menge $convcone(D_{\varepsilon k^0}) \subseteq R^m$ ein K–Ableitungskegel für $f_{\varepsilon k^0}$ an der Stelle x_ε, weil $D \in R^m$ eine K–Ableitungsmenge für f an der Stelle x_ε ist. Also folgt aus Theorem 2.2 die Existenz eines Multiplikators $\lambda \in K_1^* \times K_2^* \times K_3^* \setminus \{0^m\}$ mit

$$\langle \lambda_2, f_2(x_\varepsilon)\rangle = 0$$

und

$$\langle \lambda, d_\varepsilon\rangle \geq 0 \qquad \forall d_\varepsilon \in convcone\,(D_{\varepsilon k^0}),$$

$$\begin{aligned}\langle \lambda, d_\varepsilon \rangle &\geq 0 \qquad \forall d_\varepsilon \in D_{\varepsilon k^0},\\ \left\langle \lambda, (d + \sqrt{\varepsilon}\|h\|(k^0, 0^{m_2}, 0^{m_3}))\right\rangle &\geq 0 \qquad \forall d \in D,\\ \langle \lambda, d \rangle &\geq -\sqrt{\varepsilon}\|h\|\langle \lambda_1, k^0\rangle \qquad \forall d \in D.\end{aligned}$$

■

Ähnliche Sätze lassen sich auch für die Spezialfälle (S2) und (S3) angeben.

Es zeigt sich, daß in vielen praktischen Fällen derivierte Mengen auftreten, die eine solche spezielle Struktur haben. Dies trifft bereits unter den Voraussetzungen von Theorem 2.3 zu. Um dies zu zeigen, setzen wir $K_1 := R_+^{m_1}$, $K_2 := R_+^{m_2}$ und $K_3 := \{0^{m_3}\}$. In diesem Fall sei $\tilde{K}$ ein entsprechender Oberkegel von $R_+^{m_1}$.

Dann ergibt sich das folgende Resultat als Folgerung von Satz 4.9.

Satz 4.10 *Gegeben seien $\varepsilon > 0$ und ein Punkt $x_0 \in S$ mit*

$$(f_1(x_0) - \varepsilon k^0 - \tilde{K} \setminus \{0^{m_1}\}) \cap f_1(S) = \emptyset.$$

Für jedes $\tilde{x} \in X$ mit $\|\tilde{x} - x_0\| \leq \sqrt{\varepsilon}$ gebe es eine Menge $X_0(\tilde{x}) \subseteq \mathcal{X}$ sowie eine Funktion $F_{\tilde{x}} : X_0(\tilde{x}) \longrightarrow R^m$, die die folgenden Bedingungen erfüllen:

($\tilde{a}$) $X_0(\tilde{x})$ ist nichtleer und konvex.

($\tilde{b}$) $F_{\tilde{x},1}$ und $F_{\tilde{x},2}$ sind konvex, $F_{\tilde{x},3}$ ist affin.

($\tilde{c}$) Für jede Zahl $n \in N$ und jedes n–Tupel $\{x^1, \ldots, x^n\} \subset X_0(\tilde{x})$ gibt es ein $r > 0$, so daß

$$\tilde{x} + \sum_{i=1}^{n} t_i x^i \in X \qquad \forall (t_1, \ldots, t_n) \in B_+^n(r)$$

gilt und die Funktion

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

stetig ist.

($\tilde{d}$) Für alle $x \in X_0(\tilde{x})$ gilt

$$\limsup_{a \downarrow 0} \frac{f_1(\tilde{x} + ax) - f_1(\tilde{x})}{a} \quad \leq \quad F_{\tilde{x},1}(x),$$

$$\begin{aligned}\limsup_{a\downarrow 0}\frac{f_2(\tilde{x}+ax)-f_2(\tilde{x})}{a} &\leq F_{\tilde{x},2}(x),\\ \lim_{a\downarrow 0}\frac{f_3(\tilde{x}+ax)-f_3(\tilde{x})}{a} &= F_{\tilde{x},3}(x),\end{aligned}$$

und für jedes konvexe Polytop $P \subseteq X_0(\tilde{x})$ ist die Konvergenz gleichmäßig für alle $x \in P$.

Dann existiert ein $x_\varepsilon \in S$ mit

(i) $(f_1(x_\varepsilon) - \varepsilon k^0 - R_+^{m_1} \setminus \{0^{m_1}\}) \cap f_1(S) = \emptyset$,

(ii) $\|x_\varepsilon - x_0\| \leq \sqrt{\varepsilon}$,

(iii) Es existiert ein Multiplikator $\lambda \in R^m \setminus \{0^m\}$ mit $\lambda_1 \geq 0^{m_1}$, $\lambda_2 \geq 0^{m_2}$,

$$\begin{aligned}\langle \lambda, F_{x_\varepsilon}(x)\rangle &\geq -\sqrt{\varepsilon}\|x\|\langle \lambda_1, k^0\rangle \qquad \forall x \in X_0(x_\varepsilon),\\ \langle \lambda_2, f_2(x_\varepsilon)\rangle &= 0.\end{aligned}$$

Beweis. Zunächst zeigen wir, daß für jedes $\tilde{x} \in X$ mit $\|\tilde{x} - x_0\| \leq \sqrt{\varepsilon}$ die Menge $F_{\tilde{x}}(X_0)$ eine K–Ableitungsmenge für f an der Stelle $\tilde{x}$ ist, die (S1) erfüllt.
Wie im im Originalbeweis von Theorem 2.3 (vgl. [5], Theorem 4.1) gezeigt wird, ist unter den dort angegebenen Voraussetzungen (a) – (d) die Menge $F(X_0)$ eine K–Ableitungsmenge für f an der Stelle x_0, wobei für jede Zahl $n \in N$ und jedes n–Tupel $\{F(x^1), \ldots, F(x^n)\} \subset F(X_0)$ die Anforderungen in Definition 2.4 von der Größe r gemäß (c), von der Größe ω definiert durch

$$\omega(t) := x_0 + \sum_{i=1}^{n} t_i x^n \qquad \forall (t_1, \ldots, t_n) \in B_+^n(r)$$

sowie von der Größe ϱ definiert durch

$$\varrho(t) := \begin{cases} \frac{1}{\|t\|}\left(f(\omega(t)) - f(x_0) - \sum_{i=1}^{n} F(x^n)\right), & \text{wenn} \quad (t_1, \ldots, t_n) \in B_+^n(r) \setminus \{0\} \\ 0, & \text{wenn} \quad (t_1, \ldots, t_n) = 0 \end{cases}$$

erfüllt werden.
Analog ergibt sich, daß wegen ($\tilde{\text{a}}$) – ($\tilde{\text{d}}$) für jedes $\tilde{x} \in X$ mit $\|\tilde{x} - x_0\| \leq \sqrt{\varepsilon}$ die Menge $F_{\tilde{x}}(X_0(\tilde{x}))$ eine K–Ableitungsmenge für f an der Stelle $\tilde{x}$ ist, deren ω–Funktion für jedes n–Tupel $\{F(x^1), \ldots, F(x^n)\} \subset F_{\tilde{x}}(X_0(\tilde{x}))$ die Gestalt

$$\omega(t) = \tilde{x} + \sum_{i=1}^{n} t_i x^n \qquad \forall (t_1, \ldots, t_n) \in B_+^n(r)$$

besitzt. Mithin erfüllt $F_{\tilde{x}}(X_0(\tilde{x}))$ die Spezialbedingung (S1), indem wir jedem $d \in F_{\tilde{x}}(X_0(\tilde{x}))$ einen Punkt $h \in X_0(\tilde{x})$ zuordnen, für den $F_{\tilde{x}}(x) = d$ gilt.

Nach Satz 4.9 existiert ein Punkt $x_\varepsilon \in S$, welcher (i) und (ii) erfüllt. Nach der vorangegangenen Überlegung ist wegen (ii) die Menge $D = F_{x_\varepsilon}(X_0(x_\varepsilon))$ eine K–Ableitungsmenge für f an der Stelle x_ε, die (S1) erfüllt. Somit existiert nach Satz 4.9 weiterhin ein Multiplikator $\lambda \in R_+^{m_1} \times R_+^{m_2} \times R^{m_3} \setminus \{0^m\}$ mit

$$
\begin{aligned}
\langle \lambda_2, f_2(x_\varepsilon)\rangle &= 0, \\
\langle \lambda, F(x)\rangle &\geq -\sqrt{\varepsilon}\|x\|\langle \lambda_1, k^0\rangle \qquad \forall x \in X_0(x_\varepsilon)
\end{aligned}
$$

■

Die Struktur der Voraussetzungen für Satz 4.10 lehnen sich an die Voraussetzungen von Theorem 2.3 an. Wie bereits erwähnt, erlaubt Theorem 2.3 weitere Spezialisierungen für den konvexen (vgl. Satz 2.4) und den differenzierbaren Fall (vgl. Satz 2.5). Ebenso folgen aus Satz 4.10 die Sätze 4.11 (im konvexen Fall) und 4.12 (im differenzierbaren Fall).

Satz 4.11 *Gegeben seien $\varepsilon > 0$ und ein Punkt $x_0 \in S$ mit*

$$(f_1(x_0) - \varepsilon k^0 - \tilde{K} \setminus \{0^{m_1}\}) \cap f_1(S) = \emptyset.$$

Sei die Menge X konvex, seien die Funktionen f_1 und f_2 konvex und sei f_3 affin.
Dann existiert ein $x_\varepsilon \in S$ mit

(i) $(f_1(x_\varepsilon) - \varepsilon k^0 - R_+^{m_1} \setminus \{0^{m_1}\}) \cap f_1(S) = \emptyset$,

(ii) $\|x_\varepsilon - x_0\| \leq \sqrt{\varepsilon}$,

(iii) Es existiert ein Multiplikator $\lambda \in R^m \setminus \{0^m\}$ mit $\lambda_1 \geq 0^{m_1}$, $\lambda_2 \geq 0^{m_2}$,

$$
\begin{aligned}
\langle \lambda, f(x)\rangle - \langle \lambda_1, f_1(x_\varepsilon)\rangle &\geq -\sqrt{\varepsilon}\|x - x_\varepsilon\|\langle \lambda_1, k^0\rangle \qquad \forall x \in X, \\
\langle \lambda_2, f_2(x_\varepsilon)\rangle &= 0.
\end{aligned}
$$

Beweis. Wie im im Beweis von Satz 2.4 (vgl. [5], Corollary 4.2) gezeigt wird, erfüllen unter den dort angegebenen Voraussetzungen die Menge $X_0 := X - x_0$ und die Funktion $F(x) := f(x + x_0) - f(x_0)$ die Voraussetzungen (a) – (d) von Theorem 2.3.
Aus analogen Gründen erfüllen unter den Voraussetzungen des zu beweisenden Satzes die Menge $X_0(\tilde{x}) := X - \tilde{x}$ und die Funktion $F_{\tilde{x}}(x) := f(x + \tilde{x}) - f(\tilde{x})$ die Voraussetzungen (ã) – ($\tilde{\text{d}}$) von Satz 4.10 und zwar für jedes beliebige $\tilde{x} \in X$.

Dann folgt aus Satz 4.10 die Existenz eines $x_\varepsilon \in S$ mit den Eigenschaften (i) und (ii), für das darüberhinaus gilt:
Es existiert ein $\lambda \in R^m \setminus \{0^m\}$ mit $\lambda_1 \geq 0^{m_1}$, $\lambda_2 \geq 0^{m_2}$,

$$\langle \lambda_2, f_2(x_\epsilon) \rangle = 0$$

und

$$\begin{aligned} \langle \lambda, f(x+x_\varepsilon) - f(x_\varepsilon) \rangle &\geq -\sqrt{\varepsilon}\|x\|\langle \lambda_1, k^0 \rangle \qquad \forall x \in X - x_\varepsilon, \\ \langle \lambda, f(x+x_\varepsilon) \rangle - \langle \lambda_1, f_1(x_\varepsilon) \rangle &\geq -\sqrt{\varepsilon}\|x\|\langle \lambda_1, k^0 \rangle \qquad \forall x \in X - x_\epsilon, \\ \langle \lambda, f(x) \rangle - \langle \lambda_1, f_1(x_\varepsilon) \rangle &\geq -\sqrt{\varepsilon}\|x - x_\varepsilon\|\langle \lambda_1, k^0 \rangle \qquad \forall x \in X. \end{aligned}$$

■

Satz 4.12 *Gegeben seien $\varepsilon > 0$ und ein Punkt $x_0 \in S$ mit*

$$(f_1(x_0) - \varepsilon k^0 - \tilde{K} \setminus \{0^{m_1}\}) \cap f_1(S) = \emptyset.$$

Die Funktion f sei auf der Menge $\{\tilde{x} \in \mathcal{X};\ \|\tilde{x} - x_0\| \leq \sqrt{\varepsilon}\}$ Fréchet–differenzierbar.
Dann existiert ein $x_\varepsilon \in S$ mit

(i) $(f_1(x_\varepsilon) - \varepsilon k^0 - R_+^{m_1} \setminus \{0^{m_1}\}) \cap f_1(S) = \emptyset$,

(ii) $\|x_\varepsilon - x_0\| \leq \sqrt{\varepsilon}$,

(iii) Es existiert ein Multiplikator $\lambda \in R^m \setminus \{0^m\}$ mit $\lambda_1 \geq 0^{m_1}$, $\lambda_2 \geq 0^{m_2}$,

$$\left\| \sum_{i=1}^{m_1} \lambda_{1i}, f'_{1i}(x_\varepsilon) + \sum_{k=1}^{m_2} \lambda_{2k}, f'_{2k}(x_\varepsilon) + \sum_{l=1}^{m_3} \lambda_{3l}, f'_{3l}(x_\varepsilon) \right\|_* \leq \sqrt{\varepsilon}\langle \lambda_1, k^0 \rangle,$$

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

Beweis. Wie im im Beweis von Satz 2.5 (vgl. [5], Corollary 5.2) gezeigt wird, erfüllen unter den dort angegebenen Voraussetzungen die Menge $X_0 := \mathcal{X}$ und die Funktion $F(x) := f'(x_0)(x)$ die Voraussetzungen (a) – (d) von Theorem 2.3.
Aus analogen Gründen erfüllen unter den Voraussetzungen des zu beweisenden Satzes die Menge $X_0(\tilde{x}) := \mathcal{X}$ und die Funktion $F_{\tilde{x}}(x) := f'(\tilde{x})(x)$ die Voraussetzungen (ã) – ($\tilde{\text{d}}$) von Satz 4.10 für jedes beliebige $\tilde{x} \in \mathcal{X}$ mit $\|\tilde{x} - x_0\| \leq \sqrt{\varepsilon}$.

Dann folgt aus Satz 4.10 die Existenz eines $x_\varepsilon \in S$ mit den Eigenschaften (i) und (ii), für das darüberhinaus gilt:
Es existiert ein $\lambda \in R^m \setminus \{0^m\}$ mit $\lambda_1 \geq 0^{m_1}$, $\lambda_2 \geq 0^{m_2}$,

$$\langle \lambda_2, f_2(x_\epsilon) \rangle = 0$$

und

$$\langle \lambda, f'(x_\varepsilon)(x) \rangle \geq -\sqrt{\varepsilon}\|x\|\langle \lambda_1, k^0 \rangle \qquad \forall x \in \mathcal{X}.$$

Aus Symmetriegründen gilt daher auch

$$|\langle \lambda, f'(x_\varepsilon)(x) \rangle| \leq \sqrt{\varepsilon}\|x\|\langle \lambda_1, k^0 \rangle \qquad \forall x \in \mathcal{X}.$$

Daraus folgt schließlich

$$\begin{aligned}
&\left\| \sum_{i=1}^{m_1} \lambda_{1i}, f'_{1i}(x_\varepsilon) + \sum_{k=1}^{m_2} \lambda_{2k}, f'_{2k}(x_\varepsilon) + \sum_{l=1}^{m_3} \lambda_{3l}, f'_{3l}(x_\varepsilon) \right\|_* \\
= &\sup_{\|x\|=1} \left| \sum_{i=1}^{m_1} \lambda_{1i}, f'_{1i}(x_\varepsilon)(x) + \sum_{k=1}^{m_2} \lambda_{2k}, f'_{2k}(x_\varepsilon)(x) + \sum_{l=1}^{m_3} \lambda_{3l}, f'_{3l}(x_\varepsilon)(x) \right| \\
= &\sup_{\|x\|=1} |\langle \lambda, f'(x_\varepsilon)(x) \rangle| \\
\leq &\sqrt{\varepsilon} \sup_{\|x\|=1} \|x\|\langle \lambda_1, k^0 \rangle \\
= &\sqrt{\varepsilon}\langle \lambda_1, k^0 \rangle.
\end{aligned}$$

■

Bemerkung 4.7 *Satz 4.12 ist die Erweiterung eines bekannten Resultats von Ekeland (vgl. [11], Theorem 3.1) für Probleme mit vektorwertigen Zielfunktionen. Die dort verwendete Regularitätsbedingung*

> Die Fréchet–Ableitungen $f'_{2k}(x)$ $(k \in I(x))$ und $f'_{3l}(x)$ $(l \in \{1, \ldots, m_3\})$ sind für jedes x mit $\|x - x_0\| \leq \sqrt{\varepsilon}$ linear unabhängig.
> (Hierbei sei $I(x) := \{k \in \{1, \ldots, m_2\} |\ f_{2k}(x) = 0\}$.)

sichert auch im Fall von Satz 4.12 in (iii) die Existenz eines Multiplikators λ mit $\lambda_1 \neq 0^{m_1}$.

Beweis. Wäre in (iii) $\lambda_1 = 0^{m_1}$, so würde folgen

$$\sum_{k=1}^{m_2} \lambda_{2k}, f'_{2k}(x_\varepsilon) + \sum_{l=1}^{m_3} \lambda_{3l}, f'_{3l}(x_\varepsilon) = 0.$$

Aus $\langle\lambda_2, f_2(x_\varepsilon)\rangle = 0$ ergibt sich $\lambda_{2k} = 0$ $(k \in \{1, \ldots, m_2\} \setminus I(x_\varepsilon))$. Also wäre

$$\sum_{k\in I(x_\varepsilon)} \lambda_{2k}, f'_{2k}(x_\varepsilon) + \sum_{l=1}^{m_3} \lambda_{3l}, f'_{3l}(x_\varepsilon) = 0.$$

Wegen der linearen Unabhängigkeit der Funktionale $f'_{2k}(x_\varepsilon)$ $(k \in I(x_\varepsilon))$ und $f'_{3l}(x_\varepsilon)$ $(l \in \{1, \ldots, m_3\})$ müßte dann gelten

$$\begin{aligned} \lambda_{2k} &= 0 \qquad (k \in I(x_\varepsilon)), \\ \lambda_{3l} &= 0 \qquad (l \in \{1, \ldots, m_3\}). \end{aligned}$$

Dies widerspricht aber $\lambda = (\lambda_1, \lambda_2, \lambda_3) \neq 0^m$. ∎

Kapitel 5

Regularitätsbedingungen

5.1 Eine notwendige und hinreichende Regularitätsbedingung

Beim Studium Lagrangescher Multiplikatorenregeln in ihrer Grundform, bestehend aus Stationaritäts– und Komplementaritätsbedingung, — betrachten wir etwa Theorem 4.2 — bemerkt man, daß es zunächst keine Aussage gibt, die sichert, daß in der Lagrange–Funktion

$$L(\lambda, x) = \langle \lambda_1, f_1(x) \rangle + \langle \lambda_2, f_2(x) \rangle + \langle \lambda_3, f_3(x) \rangle \tag{5.1}$$

der zur Zielfunktion gehörige Multiplikator λ_1 nicht verschwindet. Falls dieser gleich Null wäre, bliebe die Aussage der Multiplikatorenregel zwar theoretisch richtig; praktisch würde sie allerdings kaum noch anwendbar sein, weil die Lagrange–Funktion keine Beziehung mehr zur Zielfunktion besitzt. Die entstehende notwendige Optimalitätsbedingung wäre somit viel zu schwach, um die Menge der effizienten Punkte hinreichend gründlich zu beschreiben.

Eine Bedingung, mit deren Hilfe das Verschwinden von λ_1 verhindert werden kann, heißt oft *Regularitätsbedingung*. Für die in Kapitel 1 vorgestellte klassische Lagrange–Regel ist beispielsweise vorauszusetzen, daß die Ableitungen der Funktionen ψ_i $(i = 1, \ldots, k)$, die in (1.20) auftauchen, voneinander linear unabhänig sind, um in (1.21) den Multiplikator vor der Zielfunktion ϕ von Null verschieden (in diesem Fall gleich Eins) wählen zu können.

Für die Brecknersche Multiplikatorenregel (Theorem 2.2) existierte bislang keine solche Regularitätsbedingung. Für den von Hestenes eingeführten Begriff derivierter Mengen (vgl. Definition 2.1) und die damit in [19] hergeleitete Multiplikatorenregel wird allerdings eine derartige Bedingung angegeben (vgl. [19], Corollary im Anschluß an Theorem 10.1).

Basierend auf der Hestenesschen Bedingung habe ich eine Regularitätsaussage für die Multiplikatorenregel von Breckner aufgestellt.

Wir bezeichnen die Eingangsdaten der von Theorem 2.2 betrachteten Aufgabe

- als *regulär*, wenn es einen Multiplikator λ gibt, für den $\lambda_1 \neq 0^{m_1}$ gilt;
- als *irregulär*, wenn für jeden möglichen Multiplikator λ stets $\lambda_1 = 0^{m_1}$ gilt.

Es sei an die Festlegung in (2.21) erinnert. Dann gilt die folgende Aussage:

Theorem 5.1 *Es seien die Voraussetzungen von Theorem 2.2 erfüllt.*

a) Existiert ein Punkt $a \in R^{m_1}$*, so daß*

$$\begin{pmatrix} a \\ 0^{m_2} \\ 0^{m_3} \end{pmatrix} \notin D + K_1 \times L^* \times K_3, \tag{5.2}$$

so existiert ein Multiplikator $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ *gemäß Theorem 2.2 mit* $\lambda_1 \neq 0^{m_1}$.

b) Gilt hingegen für alle $a \in R^{m_1}$

$$\begin{pmatrix} a \\ 0^{m_2} \\ 0^{m_3} \end{pmatrix} \in D + K_1 \times L^* \times K_3, \tag{5.3}$$

so gilt für alle Multiplikatoren $\lambda = (\lambda_1, \lambda_2, \lambda_3)$*, die die in Theorem 2.2 genannten Eigenschaften besitzen, stets* $\lambda_1 = 0^{m_1}$.

Beweis. a) Falls ein solcher Punkt existiert, so besitzt er eine Darstellung

$$(a, 0^{m_2}, 0^{m_3}) = x - \lambda$$

mit

$$\begin{aligned} x &\in D + K_1 \times L^* \times K_3, \\ -\lambda &\in -[D + K_1 \times L^* \times K_3]^* \end{aligned}$$

und $\langle \lambda, x \rangle = 0$. Es folgt $\lambda \neq 0^m$, denn sonst wäre $(a, 0^{m_2}, 0^{m_3}) = x \in D + K_1 \times L^* \times K_3$, was (5.2) widerspricht.

Wegen $D \subseteq D + K_1 \times L^* \times K_3$ erhalten wir $[D + K_1 \times L^* \times K_3]^* \subseteq D^*$, mithin gilt $\lambda \in D^*$, das heißt

$$\langle \lambda, d \rangle \geq 0 \qquad \forall d \in D.$$

Wegen $K_1 \times L^* \times K_3 \subseteq D + K_1 \times L^* \times K_3$ erhalten wir $[D + K_1 \times L^* \times K_3]^* \subseteq K_1^* \times L \times K_3^*$, mithin gilt

$$\lambda \in K_1^* \times L \times K_3^*.$$

Schließlich haben wir

$$\langle \lambda_1, a \rangle = \langle \lambda, (a, 0^{m_2}, 0^{m_3}) \rangle = \langle \lambda, x \rangle - \langle \lambda, \lambda \rangle = -\|\lambda\|^2 < 0.$$

Also muß gelten $\lambda_1 \neq 0^{m_1}$.

b) Zu beliebigem $a \in R^{m_1}$ gibt es wegen (5.3) einen Vektor $(x_1, x_2, x_3) \in K_1 \times L^* \times K_3$, so daß

$$(a, 0^{m_2}, 0^{m_3}) - (x_1, x_2, x_3) \in D.$$

Sei $\lambda = (\lambda_1, \lambda_2, \lambda_3) \in K_1^* \times L \times K_3^* \setminus \{0^m\}$ ein Multiplikator mit den in Theorem 2.2 genannten Eigenschaften. Dann folgt:

$$\begin{aligned}
\langle \lambda, (a, 0^{m_2}, 0^{m_3}) - (x_1, x_2, x_3) \rangle &\geq 0, \\
\langle \lambda_1, a \rangle &\geq \langle \lambda_1, x_1 \rangle + \langle \lambda_2, x_2 \rangle + \langle \lambda_3, x_3 \rangle, \\
\langle \lambda_1, a \rangle &\geq 0.
\end{aligned}$$

Man wähle ein spezielles $a \in -int\, K_1$. Für dieses gilt obige Aussage und wegen

$$\langle \lambda_1, y \rangle \geq 0 \qquad \forall y \in K_1$$

sogar $\langle \lambda_1, a \rangle = 0$.

Wir betrachten nun die Abbildung $\langle \lambda_1, . \rangle : R^{m_1} \longrightarrow R$. Diese ist nichtnegativ auf dem Bereich K_1 und verschwindet in einem inneren Punkt dieses Bereiches, nämlich in $-a \in int\, K_1$. Eine lineare Funktion mit einem solchen Verhalten ist zwangsläufig identisch Null. Also folgt $\lambda_1 = 0^{m_1}$. ■

Bemerkung 5.1 *Beim Beweisteil b) wird nur ausgenutzt, daß es ein spezielles $a \in -int\, K_1$ gibt, welches (5.3) erfüllt. Die Forderung, daß (5.3) für alle $a \in R^{m_1}$ gelten soll, erscheint daher unnötig*

scharf zu sein. In der Tat sind jedoch beide Bedingungen äquivalent:
Sei $\tilde{a} \in -int\, K_1$ ein Punkt, für den gilt

$$(\tilde{a}, 0^{m_2}, 0^{m_3}) \in D + K_1 \times L^* \times K_3.$$

Dann gibt es für jeden Punkt $a \in R^{m_1}$ eine Zahl $\kappa > 0$, so daß $\tilde{a} + \kappa a \in -K_1$. Mithin gilt

$$\begin{aligned} \kappa(a, 0^{m_2}, 0^{m_3}) &= (\tilde{a}, 0^{m_2}, 0^{m_3}) - (\tilde{a} + \kappa a, 0^{m_2}, 0^{m_3}) \\ &\in (D + K_1 \times L^* \times K_3) + (K_1 \times L^* \times K_3) \\ &\subseteq D + K_1 \times L^* \times K_3. \end{aligned}$$

Da $D + K_1 \times L^* \times K_3$ ein Kegel ist, folgt hieraus (5.3) für beliebiges $a \in R^{m_1}$.
Die andere Richtung der Äquivalenz ist trivial.

Bemerkung 5.2 *Die Bedingung in a) ist notwendig und hinreichend für die Regulärität der Daten des betrachteten Problems. Einerseits sichert das Erfülltsein von a) die Existenz eines Multiplikators für die Brecknersche Regel, dessen erste Komponente nicht verschwindet. Andererseits wird im Falle der Verletzung von a) sofort die Bedingung in b) erfüllt, so daß Irregularität folgt.*

5.2 Diskussion in Spezialfällen

Wir wollen in diesem und im folgenden Abschnitt die Regularitätsbedingung von Theorem 5.1 auf die in Abschnitt 2.2 vorgestellten Anwendungen der Brecknerschen Multiplikatorenregel (Theorem 2.3, Sätze 2.4 und 2.5) übertragen.

Bei diesen Sätzen besitzen die auftretenden derivierten Mengen eine spezielle Struktur. Für solche Strukturen sind allerdings bereits einige andere Regularitätsbedingungen bekannt. Von besonderem Interesse wird daher sein, wie sich die Anwendungen von Theorem 5.1 in Bezug auf die bekannten Bedingungen verhalten.

Auffällig ist zunächst, daß die vorhandenen Regularitätsbedingungen lediglich auf Eigenschaften der Restriktionsmenge S zurückgreifen, während Theorem 5.1 auch die Zielfunktion f_1 mit einbezieht. Zur Abkürzung werden Regularitätsbedingungen, die ausschließlich Anforderungen an S stellen, im folgenden als *constraint qualifications* bezeichnet. (Es gibt leider keine deutsche Entsprechung für diesen Begriff.) Vom informationstheoretischen Standpunkt aus ist zu erwarten, daß die Bedingungen, die auch Eigenschaften der Zielfunktion betrachten, eine gründlichere Charakterisierung der Regulärität erlauben.

Wir beginnen mit einer Anwendung der Regularitätsbedingung in Satz 5.1 auf die Verhältnisse der Multiplikatorenregel von Theorem 2.3.

Satz 5.2 *Es seien die Voraussetzungen von Theorem 2.3 erfüllt.*

a) Existiert ein Punkt $a \in R^{m_1}$, *so daß*

$$\begin{pmatrix} a \\ 0^{m_2} \\ 0^{m_3} \end{pmatrix} \notin \mathit{cone}\,(F(X_0) + R_+^{m_1} \times L^* \times \{0^{m_3}\}), \tag{5.4}$$

so existiert ein Multiplikator $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ *gemäß Theorem 2.3 mit* $\lambda_1 \neq 0^{m_1}$.

b) Gilt hingegen für alle $a \in R^{m_1}$

$$\begin{pmatrix} a \\ 0^{m_2} \\ 0^{m_3} \end{pmatrix} \in \mathit{cone}\,(F(X_0) + R_+^{m_1} \times L^* \times \{0^{m_3}\}), \tag{5.5}$$

so gilt für alle Multiplikatoren $\lambda = (\lambda_1, \lambda_2, \lambda_3)$, *die die in Theorem 2.3 genannten Eigenschaften besitzen, stets* $\lambda_1 = 0^{m_1}$.

Beweis. Theorem 2.3 folgt aus Theorem 2.2 unter Verwendung von $D := \mathit{convcone}\,(F(X_0))$, $K_1 := R_+^{m_1}$, $K_2 := R_+^{m_2}$ und $K_3 := 0^{m_3}$. Man verfahre analog in Theorem 5.1.
Zu zeigen bleibt

$$\mathit{convcone}\,(F(X_0)) + R_+^{m_1} \times L^* \times \{0^{m_3}\} = \mathit{cone}\,(F(X_0) + R_+^{m_1} \times L^* \times \{0^{m_3}\}).$$

Wegen $\mathit{cone}\,(F(X_0)) \subseteq \mathit{convcone}\,(F(X_0))$ folgt sofort

$$\begin{aligned} \mathit{convcone}\,(F(X_0)) + R_+^{m_1} \times L^* \times \{0^{m_3}\} &\supseteq \mathit{cone}\,(F(X_0)) + R_+^{m_1} \times L^* \times \{0^{m_3}\} \\ &= \mathit{cone}\,(F(X_0) + R_+^{m_1} \times L^* \times \{0^{m_3}\}). \end{aligned}$$

Andererseits ist die Menge $F(X_0) + R_+^{m_1} \times L^* \times \{0^{m_3}\}$ wegen der Konvexität der Funktion F und der Menge X_0 konvex und mithin $\mathit{cone}\,(F(X_0) + R_+^{m_1} \times L^* \times \{0^{m_3}\})$ ein konvexer Kegel. Dieser Kegel enthält sicher die Menge $F(X_0)$. Also folgt

$$\mathit{convcone}\,(F(X_0)) \subseteq \mathit{cone}\,(F(X_0) + R_+^{m_1} \times L^* \times \{0^{m_3}\})$$

und somit

$$convcone\,(F(X_0)) + R_+^{m_1} \times L^* \times \{0^{m_3}\} \subseteq cone\,(F(X_0) + R_+^{m_1} \times L^* \times \{0^{m_3}\}).$$

■

Im Fall $K_2 = R_+^{m_2}$ hat der Kegel L^* die Gestalt

$$L^* = \{y = (y_1, \ldots, y_{m_2}) \in R^{m_2} |\; y_k \geq 0 \,\forall k \in I_0\}, \tag{5.6}$$

wobei $I_0 := \{k \in \{1, \ldots, m_2\}|\; f_{2k}(x_0) = 0\}$ die Indexmenge der in x_0 aktiven Ungleichungsnebenbedingungen bezeichne. Dann lassen sich (5.4), (5.5) folgendermaßen interpretieren.

Korollar 5.3 *Es seien die Voraussetzungen von Theorem 2.3 erfüllt.*

a) Existiert zu einem Punkt $a = (a_1, \ldots, a_{m_1}) \in R^{m_1}$ kein Paar $(\alpha, x) \in R_+ \times X_0$, welches das System

$$\begin{array}{rcll} \alpha F_{1i}(x) & \leq & a_i & (i = 1, \ldots, m_1) \\ \alpha F_{2k}(x) & \leq & 0 & (k \in I_0) \\ \alpha F_{3l}(x) & = & 0 & (l = 1, \ldots, m_3) \end{array} \tag{5.7}$$

löst, so existiert ein Multiplikator $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ gemäß Theorem 2.3 mit $\lambda_1 \neq 0^{m_1}$.

b) Existiert hingegen für alle $a \in R^{m_1}$ eine Lösung $(\alpha, x) \in R_+ \times X_0$ des obigen Systems, so gilt für alle Multiplikatoren $\lambda = (\lambda_1, \lambda_2, \lambda_3)$, die die in Theorem 2.3 genannten Eigenschaften besitzen, stets $\lambda_1 = 0^{m_1}$.

Bemerkung 5.3 *Wie in Bemerkung 5.1 vorgeführt wird, läßt sich auch hier zeigen, daß die Voraussetzungen von b) genau dann erfüllt sind, wenn das System (5.7) für ein spezielles $a = \tilde{a} \in -int\, R_+^{m_1}$, das heißt $\tilde{a}_i < 0$ $(i = 1, \ldots, m_1)$, eine Lösung $(\alpha, x) \in R_+ \times X_0$ hat.*
Darüberhinaus reicht es aus, anstelle der Lösbarkeit von (5.7) lediglich die Lösbarkeit des Systems

$$\begin{array}{rcll} F_{1i}(x) & \leq & a_i & (i = 1, \ldots, m_1) \\ F_{2k}(x) & \leq & 0 & (k \in I_0) \\ F_{3l}(x) & = & 0 & (l = 1, \ldots, m_3) \end{array} \tag{5.8}$$

für ein $a = \tilde{a} \in -int\, R_+^{m_1}$ *zu verlangen. Es gilt nämlich: Das System (5.7) hat genau dann für alle* $a \in R^{m_1}$ *eine Lösung, wenn (5.8) für wenigstens ein* $a = \tilde{a} \in -int\, R_+^{m_1}$ *eine Lösung besitzt. Dies sieht man wie folgt:*

Sei $\tilde{a} \in -int\, R_+^{m_1}$ und $\tilde{x} \in X_0$ eine Lösung von (5.8) mit $a = \tilde{a}$, das heißt

$$\begin{array}{rcll} F_{1i}(\tilde{x}) & \leq & \tilde{a}_i & (i = 1, \ldots, m_1) \\ F_{2k}(\tilde{x}) & \leq & 0 & (k \in I_0) \\ F_{3l}(\tilde{x}) & = & 0 & (l = 1, \ldots, m_3) \end{array}$$

Für jedes $a \in R^{m_1}$ gibt es eine Zahl $\kappa > 0$, so daß $\tilde{a} + \kappa a \in -R_+^{m_1}$. Mithin gilt

$$F_{1i}(\tilde{x}) \leq \tilde{a}_i \leq \kappa a_i \qquad (i = 1, \ldots, m_1).$$

Somit ist $(\frac{1}{\kappa}, \tilde{x}) \in R_+ \times X_0$ eine Lösung von (5.7), und eine solche Lösung existert für jedes $a \in R^{m_1}$. Nun nehmen wir an, (5.7) sei für jedes $a \in R^{m_1}$ lösbar. Dann gibt es insbesondere für $a = \tilde{a} \in -int R_+^{m_1}$ eine Lösung $(\alpha, x) \in R_+ \times X_0$. Es folgt sofort $\alpha > 0$. Damit ist $x \in X_0$ eine Lösung von (5.8) für $a = \frac{1}{\alpha}\tilde{a} \in -int\, R_+^{m_1}$.

Der Regularitätsbedingung von Korollar 5.3 wollen wir die folgende Aussage in Form von *constraint qualifications* gegenüberstellen. Es ist eine Bedingung über innere Punkte, wie sie üblicherweise verwendet wird, wenn die Eingangsdaten, wie hier, konvex sind. Wir werden zeigen, daß diese *constraint qualifications* lediglich sicherstellen, daß die Bedingung a) in Korollar 5.3 erfüllt wird, aber im Fall ihrer Verletztheit keineswegs Irregularität vorliegen muß.

Satz 5.4 *Zusätzlich zu den Voraussetzungen von Theorem 2.3 gelte:*

- *Es gibt ein* $\bar{x} \in int\, X_0$ *mit*

$$\begin{array}{rcll} F_{2k}(\bar{x}) & < & 0 & (k \in I_0), \\ F_{3l}(\bar{x}) & = & 0 & (l = 1, \ldots, m_3). \end{array}$$

- *Die (affinen) Funktionen* F_{3l} $(l = 1, \ldots, m_3)$ *sind linear unabhängig.*

Dann existiert ein Multiplikator $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ *gemäß Theorem 2.3 mit* $\lambda_1 \neq 0^{m_1}$.

Beweis. Angenommen, in Korollar 5.3 wäre a) verletzt. Dann gibt es für alle $a \in R^{m_1}$ eine Lösung von (5.7). Also existiert insbesondere für ein $\tilde{a} = (\tilde{a}_1, \ldots, \tilde{a}_{m_1}) \in R^{m_1}$ mit $\tilde{a}_i < 0$ $(i = 1, \ldots, m_1)$ ein

Paar $(\tilde{\alpha}, \tilde{x}) \in R_+ \times X_0$ mit

$$\begin{aligned} \tilde{\alpha} F_{1i}(\tilde{x}) &\leq \tilde{a}_i \qquad (i = 1, \ldots, m_1), \\ \tilde{\alpha} F_{2k}(\tilde{x}) &\leq 0 \qquad (k \in I_0), \\ \tilde{\alpha} F_{3l}(\tilde{x}) &= 0 \qquad (l = 1, \ldots, m_3). \end{aligned}$$

Nach Theorem 2.4 gibt es einen Multiplikator $\lambda = (\lambda_1, \lambda_2, \lambda_3)$, $\lambda \neq 0^m$ mit $\lambda_{1i} \geq 0$ $(i = 1, \ldots, m_1)$ und $\lambda_{2k} \geq 0$ $(k = 1, \ldots, m_2)$, so daß gilt

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

Hieraus folgt einerseits $\lambda_{2k} = 0$ $(k \in \{1, \ldots, m_2\} \setminus I_0)$ und andererseits

$$\sum_{i=1}^{m_1} \lambda_{1i} F_{1i}(x) + \sum_{k \in I_0} \lambda_{2k} F_{2k}(x) \sum_{l=1}^{m_3} \lambda_{3l} F_{3l}(x) \geq 0 \qquad \forall x \in X_0. \tag{5.9}$$

Für $x = \tilde{x}$ folgt aus (5.9)

$$\sum_{i=1}^{m_1} \lambda_{1i} F_{1i}(\tilde{x}) \geq 0$$

und somit

$$\sum_{i=1}^{m_1} \lambda_{1i} \tilde{a}_i \geq \tilde{\alpha} \sum_{i=1}^{m_1} \lambda_{1i} F_{1i}(\tilde{x}) \geq 0,$$

woraus sich mit $\tilde{a}_i < 0$ $(i = 1, \ldots, m_1)$ sofort $\lambda_{1i} = 0$ $(i = 1, \ldots, m_1)$ ergibt.

Weiterhin liefert $x = \bar{x}$ in (5.9) die Aussage

$$\sum_{k \in I_0} \lambda_{2k} F_{2k}(\bar{x}) \geq 0,$$

was $\lambda_{2k} = 0$ $(k \in I_0)$ bedeutet.

So reduziert sich (5.9) zu

$$\sum_{l=1}^{m_3} \lambda_{3l} F_{3l}(x) \geq 0 \qquad \forall x \in X_0. \tag{5.10}$$

Da die affine Abbildung $\sum_{l=1}^{m_3} \lambda_{3l} F_{3l}(.)$ auf dem Gebiet X_0 nichtnegativ ist und in einem inneren Punkt dieser Menge, nämlich in $\bar{x}$ verschwindet, ist sie identisch Null. Wegen der linearen Unabhängigkeit der Funktionen F_{3l} $(l = 1, \ldots, m_3)$ folgt hieraus $\lambda_{3l} = 0$ $(l = 1, \ldots, m_3)$.

Mithin ergibt sich der Widerspruch $\lambda = (\lambda_1, \lambda_2, \lambda_3) = 0^m$. ■

Die Anforderungen von Satz 5.4 sind hinreichend, aber nicht notwendig für die Regularität der Daten des betrachteten Problems. Auch wenn diese Voraussetzungen an X_0, F_2 und F_3 nicht erfüllt sind, kann es durchaus vorkommen, daß das gesamte Problem regulär ist, wenn auch nur für bestimmte Zielfunktionen. Man beachte das folgende Beispiel:

Beispiel 5.1 *$\mathcal{X} = R^2$; $m_1 = 1$, $m_2 = 3$, $m_3 = 0$*

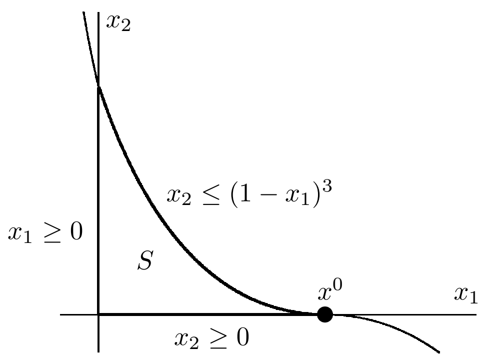
$$f_1(x) := x_2 \longrightarrow \min_{x \in S}$$

$$S = \left\{ x = (x_1, x_2) \in R^2 \;\middle|\; \begin{array}{rcll} f_{21}(x) & := & -x_1 & \leq 0 \\ f_{22}(x) & := & -x_2 & \leq 0 \\ f_{23}(x) & := & -(1-x_1)^3 + x_2 & \leq 0 \end{array} \right\}$$

Wir betrachten als Optimalpunkt $x^0 = (1, 0)$. Für diesem Punkt ist $I_0 = \{2, 3\}$.

Die auftretenden Funktionen sind sämtlich differenzierbar. Für solche Funktionen gilt Theorem 2.3 mit $X_0 = \mathcal{X}$ und $F(x) = f'(x_0)(x)$ (vgl. Beweis von Satz 2.5). Also sei:

$$\begin{array}{rcl} X_0 & := & R^2; \\ F_1(x) & := & x_2, \\ F_{21}(x) & := & -x_1, \\ F_{22}(x) & := & -x_2, \\ F_{23}(x) & := & x_2. \end{array}$$



Offenbar sind die constraint qualifications *von Satz 5.4 hier nicht erfüllt:*

$$\text{Es gibt kein } \bar{x} = (\bar{x}_1, \bar{x}_2) \in R^2 \text{ mit } \begin{array}{rcll} F_{22}(\bar{x}) & = & -\bar{x}_2 & < 0, \\ F_{23}(\bar{x}) & = & \bar{x}_2 & < 0. \end{array}$$

Wohl aber sind die Voraussetzungen für a) in Korollar 5.3 gegeben, wenn $a = -1$ gewählt wird:

$$\text{Es gibt kein } (\alpha, x) \in R_+ \times R^2 \text{ mit } \begin{array}{rcll} \alpha F_1(x) & = & \alpha x_2 & \leq -1, \\ \alpha F_{22}(x) & = & -\alpha x_2 & \leq 0, \\ \alpha F_{23}(x) & = & \alpha x_2 & \leq 0. \end{array}$$

Das hat zur Folge, daß das betrachtete Beispiel tatsächlich regulär ist. Beispielsweise erfüllt der Multiplikator $\lambda = (1,0,1,0)$ *(beachte:* $\lambda_1 \neq 0$*) die Anforderungen von Theorem 2.3.*

Es gibt jedoch Zielfunktionen über S*, so daß das Beispiel irregulär ist. Man betrachte etwa*

$$\dot{f}_1(x) := x_1 + x_2 \text{ bzw. } \dot{F}_1(x) := x_1 + x_2.$$

Dann tritt der Fall b) von Korollar 5.3 ein:

Für jedes $a \in R^{m_1}$ löst das Paar (α, x) mit $\alpha = 1$ und $x = (a,0)$ das System

$$\begin{array}{rclcl} \alpha\dot{F}_1(x) & = & \alpha(x_1+x_2) & \leq & a, \\ \alpha F_{22}(x) & = & -\alpha x_2 & \leq & 0, \\ \alpha F_{23}(x) & = & \alpha x_2 & \leq & 0. \end{array}$$

Hier zeigt sich auch der generelle Unterschied zwischen Regularitätsbedingungen nach der Hestenesschen Idee, wie der in Theorem 5.1 verwendeten, und den so häufig vorkommenden *constraint qualifications*. Letztere sind stets so angelegt, daß sie die Regularität der Aufgabe sichern, welche Gestalt die Zielfunktion auch immer besitzt. Eine tiefgründigere Aussage ist auch nicht zu erwarten, da in die *constraint qualifications* keine Information über die Zielfunktion einfließt. Selbstverständlich ist es möglich, die in Satz 5.4 gestellten Bedingungen weiter abzuschwächen und dennoch die Regularität zu sichern. Welche Bedingungen wir aber auch an die Größen X_0, F_2 und F_3 stellen, es wird permanent Fälle geben, in denen diese *constraint qualifications* verletzt sind, obwohl das Problem für die vorliegende Zielfunktion regulär ist.

Beispiel 5.2 *Gegeben sei eine — wie immer geartete — Regularitätsbedingung für das Problem vom Theorem 2.3, bei der nur die Eigenschaften von* X_0*,* F_2 *und* F_3 *Verwendung finden. Sei* S *eine zulässige Menge entsprechend dieser Aufgabenstellung, zu der man Größen* X_0*,* F_2 *und* F_3 *finden kann, die die gegebene Regularitätsbedingung verletzen. Auf dieser Menge* S *betrachte man eine triviale Zielfunktion* f_1 *mit* $F_1 \equiv 0^{m_1}$*. Dann ist jedes* $\lambda = (\lambda_1, 0^{m_2}, 0^{m_3})$ *mit* $\lambda_1 \in R_+^{m_1} \setminus \{0^{m_1}\}$ *ein Multiplikator, der die in Theorem 2.3 genannten Eigenschaften besitzt; die Gesamtdaten sind mithin regulär.*

Aus diesen Überlegungen kann man den folgenden Schluß ziehen:
Ein grundlegendes Merkmal aller *constraint qualifications* ist, daß sie regelmäßig zu harte Anforderungen an die Nebenbedingungen stellen. Dies ist typisch für solche Regularitätsbedingungen, die die Gestalt der Zielfunktion ignorieren.
Derartige Bedingungen sichern hinlänglich die Regularität der Aufgabe. Im Falle ihrer Verletzheit sind die Daten keineswegs notwendig irregulär.

Eine möglichst schwache *constraint qualification* liegt dann vor, wenn man zeigen kann, daß es bei Verletzung dieser mindestens eine Zielfunktion gibt, für die auch die Gesamtdaten irregulär sind. Eine weitere Abschwächung der Voraussetzungen ist bei Regularitätsbedingungen dieses Typs generell nicht möglich.

5.3 Weitere Spezialisierungen

Abschließend werden wir beide Typen von Regularitätsbedingungen auf die Verhältnisse im konvexen und im differenzierbaren Fall übertragen.

Wir beginnen mit dem konvexen Fall. Das Analogon zu Korollar 5.3 lautet hier:

Satz 5.5 *Es seien die Voraussetzungen von Satz 2.4 erfüllt.*

a) Existiert zu einem Punkt $a = (a_1, \dots, a_{m_1}) \in R^{m_1}$ kein Paar $(\alpha, x) \in R_+ \times X$, welches das System

$$\begin{aligned} \alpha(f_{1i}(x) - f_{1i}(x_0)) &\leq a_i \qquad (i = 1, \dots, m_1) \\ \alpha f_{2k}(x) &\leq 0 \qquad (k \in I_0) \\ \alpha f_{3l}(x) &= 0 \qquad (l = 1, \dots, m_3) \end{aligned} \tag{5.11}$$

löst, so existiert ein Multiplikator $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ gemäß Satz 2.4 mit $\lambda_1 \neq 0^{m_1}$.

b) Existiert hingegen für alle $a \in R^{m_1}$ eine Lösung $(\alpha, x) \in R_+ \times X$ des obigen Systems, so gilt für alle Multiplikatoren $\lambda = (\lambda_1, \lambda_2, \lambda_3)$, die die in Satz 2.4 genannten Eigenschaften besitzen, stets $\lambda_1 = 0^{m_1}$.

Beweis. Satz 2.4 folgt aus Theorem 2.3 unter Verwendung von $X_0 := X - x_0$ und $F(x) := f(x + x_0) - f(x_0)$. Damit spezialisiert sich (5.7) zu (5.11), wenn man beachtet, daß $f_{2k}(x_0) = 0$ $(k \in I_0)$ und $f_{3l}(x_0) = 0$ $(l = 1, \dots, m_3)$. ■

Bemerkung 5.4 *Als Spezialisierung von Bemerkung 5.3 ergibt sich hier:*
Das System (5.11) hat genau dann für alle $a \in R^{m_1}$ eine Lösung $(\alpha, x) \in R_+ \times X$, wenn das System

$$\begin{aligned} f_{1i}(x) &\leq a_i \qquad (i = 1, \dots, m_1) \\ f_{2k}(x) &\leq 0 \qquad (k \in I_0) \\ f_{3l}(x) &= 0 \qquad (l = 1, \dots, m_3) \end{aligned} \tag{5.12}$$

für wenigstens ein $a = \tilde{a} \in f_1(x_0) - int\ R_+^{m_1}$*, das heißt* $\tilde{a}_i < f_{1i}(x_0)$ $(i = 1, \ldots, m_1)$*, eine Lösung* $x \in X$ *besitzt.*

Die entsprechenden *constraint qualifications* lauten im konvexen Fall:

Satz 5.6 *Zusätzlich zu den Voraussetzungen von Satz 2.4 gelte:*

- *Es gibt ein* $\bar{x} \in int\ X$ *mit*

$$\begin{aligned} f_{2k}(\bar{x}) &< 0 \qquad (k \in I_0), \\ f_{3l}(\bar{x}) &= 0 \qquad (l = 1, \ldots, m_3). \end{aligned}$$

- *Die Funktionen* f_{3l} $(l = 1, \ldots, m_3)$ *sind linear unabhängig.*

Dann existiert ein Multiplikator $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ *gemäß Satz 2.4 mit* $\lambda_1 \neq 0^{m_1}$*.*

Beweis. Die Behauptung folgt unter Verwendung von $X_0 := X - x_0$ und $F(x) := f(x + x_0) - f(x_0)$ sofort aus Satz 5.4. ■

Im Fall differenzierbarer Funktionen ergibt sich aus Korollar 5.3:

Satz 5.7 *Es seien die Voraussetzungen von Satz 2.5 erfüllt.*

a) Existiert zu einem Punkt $a = (a_1, \ldots, a_{m_1}) \in R^{m_1}$ *kein* $x \in \mathcal{X}$*, welches das System*

$$\begin{aligned} f'_{1i}(x_0)(x) &\leq a_i \qquad (i = 1, \ldots, m_1) \\ f'_{2k}(x_0)(x) &\leq 0 \qquad (k \in I_0) \\ f'_{3l}(x_0)(x) &= 0 \qquad (l = 1, \ldots, m_3) \end{aligned} \tag{5.13}$$

löst, so existiert ein Multiplikator $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ *gemäß Satz 2.5 mit* $\lambda_1 \neq 0^{m_1}$*.*

b) Existiert hingegen für alle $a \in R^{m_1}$ *eine Lösung* $x \in \mathcal{X}$ *des obigen Systems, so gilt für alle Multiplikatoren* $\lambda = (\lambda_1, \lambda_2, \lambda_3)$*, die die in Satz 2.5 genannten Eigenschaften besitzen, stets* $\lambda_1 = 0^{m_1}$*.*

Beweis. Satz 2.5 folgt aus Theorem 2.3 unter Verwendung von $X_0 := \mathcal{X}$ und $F(x) := f'(x_0)(x)$. Wegen der Linearität von $f'(x_0)$ spezialisiert sich (5.7) zu

$$\begin{aligned} f'_{1i}(x_0)(\alpha x) &\leq a_i \qquad (i = 1, \ldots, m_1) \\ f'_{2k}(x_0)(\alpha x) &\leq 0 \qquad (k \in I_0) \\ f'_{3l}(x_0)(\alpha x) &= 0 \qquad (l = 1, \ldots, m_3). \end{aligned}$$

Wir können das Variablenpaar $(\alpha, x) \in R_+ \times \mathcal{X}$ durch eine einzige Variable $y \in \mathcal{X}$ mit $y = \alpha x$ substituieren. ■

Bemerkung 5.5 *Aus Bemerkung 5.3 folgt:*
Das System (5.11) hat bereits dann für alle $a \in R^{m_1}$ eine Lösung, wenn es für wenigstens ein $a \in -int\, R_+^{m_1}$ eine Lösung besitzt.

Es ergeben sich die folgenden *constraint qualifications*:

Satz 5.8 *Zusätzlich zu den Voraussetzungen von Satz 2.5 gelte:*

- *Es gibt ein $\bar{x} \in \mathcal{X}$ mit*

$$\begin{array}{rcll} f'_{2k}(x_0)(\bar{x}) & < & 0 & (k \in I_0), \\ f'_{3l}(x_0)(\bar{x}) & = & 0 & (l = 1, \ldots, m_3). \end{array}$$

- *Die Ableitungen $f'_{3l}(x_0)$ $(l = 1, \ldots, m_3)$ sind linear unabhängig.*

Dann existiert ein Multiplikator $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ gemäß Satz 2.4 mit $\lambda_1 \neq 0^{m_1}$.

Beweis. Man verwende in Satz 5.4 $X_0 = \mathcal{X}$ und $F(x) = f'(x_0)(x)$. ■

Nachwort

Die vorliegende Dissertation entstand in den Jahren 1996 bis 1999 während meines Promotionsstudiums am Institut für Optimierung und Stochastik im Fachbereich Mathematik und Informatik der Martin–Luther–Universität Halle–Wittenberg.

Mit der Thematik der dervierten Mengen bei multikriteriellen Optimierungsaufgaben habe ich mich bereits seit dem Jahr 1995 befaßt. Während der Untersuchungen zur Anfertigung der Diplomarbeit stand die Arbeit mit den im Jahre 1994 von Breckner [3] eingeführten K–Ableitungskegeln im Vordergrund.

Damals gelang es mir, ein Resultat, welches in [3] (Proposition 2.2) sowie in [37] (Lemma 3.4) jeweils als Folgerung eines Satzes von H. Tuy ([52], Theorem 1) verwendet wurde, auf einem direkten Weg zu beweisen, so daß der Inhalt dieses Resultats auch ohne Rückgriff auf die von Tuy betrachtete Theorie der Ungleichungssysteme dargestellt werden konnte (vgl. [43]).

Ein weiteres Ergebnis dieser Arbeiten war die Herleitung einer allgemeinen Transformation der zulässigen Menge, die als

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

($K_2 \in R^{m_2}$, $K_3 \in R^{m_3}$, $int\, K_2 \neq \emptyset$) gegeben ist, in eine Menge der Gestalt

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

($k \in [0, m_3]$, $\bar{K}_2 \subseteq R^{m_2+k}$, $int\, \bar{K}_2 \neq \emptyset$). Diese Transformation erlaubte unter anderem die Angabe der zweiten Komplementaritätsbedingung für die Brecknersche Multiplikatorenregel. Dargestellt ist dies in [6] und [44]. Andererseits erwies sich die Umwandlung der Relation $f_3(x) \in -K_3$ in einfache Gleichungen $\bar{f}_{3i}(x) = 0$ ($i = 1, \ldots, k$) als hilfreich beim Umgang mit dem Beweis der Multiplikatorenregel. Dadurch konnte das oben beschriebene Resultat ([3], Proposition 2.2) über nichtlineare Ungleichungssysteme durch einen Auflösungssatz für Gleichungssysteme ersetzt werden, was wir auch in dieser Arbeit (siehe Abschnitt 2.3) ausgenutzt haben.

Die im Verlauf des Diplomstudiums gewonnenen Ergebnisse fanden auch Eingang in das am Institut betriebene und von der DFG geförderte Forschungsprojekt „Ekelandsches Variationsprinzip und Optimalitätsbedingungen in der Optimierung / Optimalen Steuerung“ unter Leitung von Herrn Prof. Göpfert.

Zu Beginn meines Promotionsstudiums hatte ich mich eingehender mit notwendigen Optimalitätsbedingungen für Näherungslösungen von multikriteriellen Aufgaben beschäftigt. Eine gute Übersicht über die unterschiedlichen Begriffe von näherungsweiser Effizienz findet man in [54]. Einblicke in die Struktur solcher Optimalitätsbedingungen gewährten Arbeiten wie [30], [38], [46], [53] und [55]. Einige dieser Arbeiten beschäftigen sich lediglich mit skalaren Aufgabenstellungen. Dennoch erwiesen die resultierenden Multiplikatorenregeln als hilfreich, da wesentliche Eigenschaften solcher Aussagen bereits hier auftraten. Die daraus gewonnenen Erkenntnisse habe ich in Abschnitt 4.2 dargestellt.

Einige Autoren, etwa [17], [30] oder [33], benutzten das Variationsprinzip von Ekeland zur Herleitung näherungsweiser Optimalitätsbedingungen. Insbesondere findet man bei Loridan [30] eine Formulierung des Variationsprinzips für vektorwertige Funktionen. Mit Hilfe dieser Methode gelangte ich zu den in Abschnitt 4.3 aufgeführten Multiplikatorenregeln.

Die Theorie der derivierten Mengen erlaubt auch Anwendungen bei Problemen der optimalen Steuerung (siehe [4]). Erwägenswert wäre für die Zukunft, unter Anlehnung an die Resultate von Abschnitt 4.3, nach Optimalitätsbedingungen für Näherungslösungen solcher Steuerprobleme zu suchen.

Die Untersuchungen zur Struktur näherungsweiser Optimalitätsbedingungen führten mich auch zu Erkenntnissen, die in dieser Arbeit keine Berücksichtigung fanden, weil zunächst keine Beziehungen zu den derivierten Mengen erkennbar war: In einer Arbeit von Vályi [53] werden die ε–Optimalitätsbedingungen für eine vektorwertige Aufgabe in Form von ε–Sattelpunktaussagen angegeben. Weiterhin wird dort eine ε–Dualitätstheorie für diese Aufgabe entwickelt. In Analogie zu den Resultaten von Vályi lassen sich näherungsweise Sattelpunktaussagen auch für das von Tammer/Tammer ([48], [49]) betrachtete vektorwertige Approximationsproblem nachweisen (vgl. [50]). Gemeinsam mit Frau Prof. Tammer und Herrn Prof. Breckner ist hierzu eine Veröffentlichung geplant (siehe [7]).

Eine andere Anwendung des Variationsprinzips im Zusammenhang mit näherungsweisen Trennungssätzen (siehe auch [28]) habe ich in [16] vorgestellt.

Bei den in Kapitel 3 festgehaltenen Ableitungseigenschaften des Brecknerschen K–Ableitungskegels handelt es sich um solche, die zum Teil bekannt oder in mündlichen Erörterungen und Diskussionen zumindest vermutet worden waren. Jedoch waren dieser Eigenschaften bislang nicht schriftlich festgehalten worden. Aus diesem Grunde hielt ich es für angebracht, die entsprechenden Aussagen hier aufzuführen und zu beweisen. Für Vermutungen, die sich als unzutreffend herausstellten, habe ich Gegenbeispiele angegeben.

Besonders bemerkenswert fand ich die Tatsache, daß zwar jede einzelne Richtungsableitung einen derivierten Kegel erzeugt, nicht aber eine Kollektion aus mehreren Richtungsableitungen. Denkbar wäre eine weitere Untersuchung der Zusammenhänge der derivierten Kegel mit modernen Ableitungs– oder Subdifferentialbegriffen und den damit formulierten Optimalitätsbedingungen (vgl. [10], [12], [13], [20], [21], [22]).

Im Kapitel 5 habe ich, durchaus einmal abseits der Theorie derivierter Mengen, Gedanken über Regularitätsbedingungen, die sowohl notwendige als auch hinreichende Aussagen liefern können, geäußert. Es fiel mir auf, daß die Hestenessche Regularitätsbedingung in [19] seither nur unwesentliche Beachtung gefunden hat. Andererseits wird die Suche nach möglichst schwachen Regularitätsbedingungen verstärkt fortgesetzt. In [25] beispielsweise wird für die (hinreichende) Reguläritätsbedingung von Kurcyusz, Robinson und Zowe gezeigt, daß sie derart schwache Anforderungen stellt, daß bereits eine leichte Abwandlung dieser Bedingung im Fall der Regularität notwendigerweise erfüllt sein muß. Ich habe dargestellt, warum bei Regularitätsbedingungen ohne Einbeziehung von Aussagen über die Zielfunktion stets eine Diskrepanz zwischen notwendiger und hinreichender Bedingung besteht.

Regularitätsaussagen tauchen in der jüngeren Literatur ([34], [26], [27], [51]) immer häufiger vereint mit dem Begriff der metrischen Regularität auf. Es stellt sich die Frage, ob sich auch mit den Bedingungen von Hestenesschen Typ eine solche Verbindung herstellen läßt; insbesondere, ob diese dann metrische Aussagen nicht nur über die Funktionen in den Restriktionen, sondern auch über die Zielfunktion liefern. Dieses soll demnächst unteruncht werden. Besonders die von Thibault [51] vorgestellten Zusammenhänge geben hierfür einen Ausgangspunkt.

Abschließend möchte ich die Gelegenheit nutzen, mich bei Herrn Prof. A. Göpfert, der mich während meines Promotionsstudiums als Betreuer unterstützt hat, auf das herzlichste zu bedanken. Es waren nicht nur seine engagierte fachliche Anregung und Hilfe, sondern auch sein warmherziger Beistand in persönlichen Dingen, die mir den Abschluß dieser Arbeit ermöglichten. Weiterhin gilt mein Dank Frau Prof. Chr. Tammer und Herrn Prof. W. W. Breckner, die mir die Gelegenheit gaben, in zahlreichen interessanten Diskussionen neue Anregungen und Ideen für die Fortführung meiner Untersuchungen zu gewinnen.

Literaturverzeichnis

[1] Aubin, J.-P.: Optima and equilibria: An introduction to nonlinear analysis. Berlin Heidelberg, 1993.

[2] Aubin, J.-P.; Ekeland, I.: Applied nonlinear analysis. New York, 1984.

[3] Breckner, W. W.: Derived sets in multiobjective optimization. Z. Anal. Anw. **13** (1994), 725–738.

[4] Breckner, W. W.: Derived sets for weak multiobjective optimization problems with state and control variables. J. Optim. Theory Appl. **93** No. 1 (1997), 73–102.

[5] Breckner, W. W.; Göpfert, A.: Multiplier rules for weak Pareto optimization problems. Optimization **38** No. 1 (1996), 23–37.

[6] Breckner, W. W.; Göpfert, A.; Sekatzek, M.: Lagrange multipliers in vector optimization. Zeitschr. Angew. Math. Mech. **77** Suppl. 2 (1997), 525–526.

[7] Breckner, W. W.; Sekatzek, M.; Tammer, Chr.: Approximate saddle point theorems for a general class of approximation problems. (in Vorbereitung)

[8] Clarke, F. H.: Generalized gradients and applications. Trans. Amer. Math. Soc. **205** (1975), 247–262.

[9] Clarke, F. H.: Optimization and nonsmooth analysis. New York, 1983.

[10] Craven, B. D.; Ralph, D.; Glover, B. M.: Small convex–valued subdifferentials in mathematical programming. Optimization **32** (1995), 1–21.

[11] Ekeland, I.: On the variational principle. J. Math. Analysis Appl. **47** (1974), 324–353.

[12] El Abdouni, B.; Thibault, L.: Lagrange multipliers for Pareto nonsmooth programming problems in Banach spaces. Optimization **26** (1992), 277–285.

[13] Gianessi, F.: Semidifferentiable functions and necessary optimality consitions. J. Optim. Theory Appl. **60** (1989), 191–241.

[14] Gittleman, A.: A general multiplier rule. J. Optim. Theory Appl. **7** No. 1 (1971), 29–38.

[15] Göpfert., A; Nehse, R.: Vektoroptimierung. Leipzig, 1990.

[16] Göpfert, A.; Sekatzek, M.; Tammer, Chr.: Novelties around Ekeland's variational principle. *Reports of the Institute of Optimization and Stochastics.* Martin–Luther–Universität Halle–Wittenberg, Report No. 23 (1998). (erscheint in: Proc. of the Intern. Conf. of Operat. Research at Liberec 1998)

[17] Hamel, A.: A generalized Lagrange multiplier rule. (erscheint in Optimization)

[18] Hestenes, M. R.: On variational theory and optimal control theory. SIAM J. Control **3** No. 1 (1965), 23–48.

[19] Hestenes, M. R.: Optimization theory: The finite dimensional case. New York, 1975.

[20] Ioffe, A. D.: Nonsmooth analysis: Differential calculus in nonsmooth optimization. Trans. Amer. Math. Soc. **266** (1981), 1–56.

[21] Ioffe, A. D.: Necessary conditions in nonsmooth optimization. Math. Operat. Research **9** (1984), 159–189.

[22] Ioffe, A. D.: A Lagrange multiplier rule with small convex–valued subdifferentials for problems of mathematical programming involving equaltity and nonfunctional constraints. Math. Programming **58** (1993), 137–145.

[23] Ioffe, A. D.; Tichomirov, V. M.: Theorie der Extremalaufgaben. Berlin, 1979.

[24] Jahn, J.: Mathematical vector optimization in partially ordered linear spaces. Frankfurt Bern New York, 1986.

[25] Jahn, J.: Introduction to the theory of nonlinear optimization. 2nd rev. ed. Berlin Heidelberg, 1996.

[26] Jourani, A.: Constraint qualifications an Lagrange multipliers in nondifferentiable programming problems. J. Optim. Theory Appl. **81** No. 3 (1994), 533–548.

[27] Jourani, A.; Thibault, L.: Verifiable conditions for openness and regularity of multivalued mappings in Banach spaces. Trans. Amer. Math. Soc. **347** No. 4 (1985), 1255–1268.

[28] Jourani, A.: Necessary conditions for extremality and separation theorems with applications to multiobjective optimization. (erscheint in Optimization)

[29] Kruger, A. Y.; Mordukhovich, B. S.: Extremal points and Euler equations in nonsmooth optimization. (russisch) Dokl. Akad. Nauk BSSR **24** (1980), 684–687.

[30] Loridan, P.: ϵ–solutions in vector minimization problems. J. Optim. Theory Appl. **43** No. 2 (1984), 265–276.

[31] Luc, D. T.: Theory of vector optimization. Berlin, 1989.

[32] Mangasarian, O. L.: Nonlinear Programming. New York, 1969.

[33] McLinden, L.: An application of Ekeland's theorem to minimax problems. Nonlinear analysis. **6** No. 2 (1982), 189–196.

[34] Mordukhovich, B. S.: Complete characterization of openness, metric regularity, and Lipschitzian properties of multifunctions. Trans. Amer. Math. Soc. **340** No. 1 (1993), 1–35.

[35] Mordukhovich, B. S.; Shao, Y.: Extremal characterizations of Asplund spaces. Proc. Amer. Math. Soc. **124** No. 1 (1996), 197–205.

[36] Neustadt, L. W.: An abstract variational theory with applications to a broad class of optimization problems. I. General theory. SIAM J. Control **4** No. 3 (1966), 505–527.

[37] Nieuwenhuis, J. W.: A general multiplier rule. J. Optim. Theory Appl. **31** (1980), 167–176.

[38] Oettli, W.: Epsilon–solutions and epsilon–supports. Optimization **16** (1985), 491–496.

[39] Pourciau, B. H.: Modern multiplier rules. Amer. Math. Monthly **87** (1980), 433–452.

[40] Rockafellar, R. T.: Convex analysis. Princeton, 1970.

[41] Rockafellar, R. T.: The theory of subgradients and its applications to problems of optimization: convex and nonconvex functions. Berlin 1981.

[42] Rockafellar, R. T.; Wets, R. J.-B.: Variational analysis. Berlin Heidelberg, 1998

[43] Sekatzek, M.: Multiplikatorenregeln für die multikriterielle Optimierung. Diplomarbeit. Halle, 1996.

[44] Sekatzek, M.: A normal form for the feasible set of optimization problems. *Reports of the Institute of Optimization and Stochastics.* Martin–Luther–Universität Halle–Wittenberg, Report No. 26 (1997).

[45] Sekatzek, M.: Contributions to a multiplier rule for multiobjective optimization problems. *Reports of the Institute of Optimization and Stochastics.* Martin–Luther–Universität Halle–Wittenberg, Report No. 35 (1998). (erscheint in: Proc. of the 9th Belgian–French–German Conf. on Optimization. Namur, 1998)

[46] Strodiot, J.-J.; Nguyen, V. H.; Heukemes, N.: ϵ–optimal solutions in nondifferentiable convex programming and some related questions. Math. Programming **25** (1983), 307–328.

[47] Tammer, Chr.: A generalization of Ekeland's variational principle. Optimization **25** (1992), 129–141.

[48] Tammer, Chr.; Tammer, K.: Generalization and sharpening of some duality relations for a class of vector optimization problems. Zeitschr. Operat. Research **35** (1991), 249–265.

[49] Tammer, Chr.; Tammer, K.: Duality results for convex vector optimization problems with linear restrictions. In: Kall, P. [Ed.]: System modelling and optimization. Lect. Notes Control Inf. Sci. 180. Berlin, 1992.

[50] Tammer, Chr..: Lagrange–Kuhn–Tucker multipliers for general mathematical programming problems. In: Göpfert, A.; Seeländer, J.; Tammer, Chr. [Eds.]: Methods of multicriteria decision theory. Deutsche Hochschulschriften 2398. Egelsbach, 1997.

[51] Thibault, L.: Lagrange–Kuhn–Tucker multipliers for general mathematical programming problems. In: Ioffe, A.; Marcus, M.; Reich, S. [Eds.]: Optimization and nonlinear analysis. Pitman Research Notes in Mathematics. Series 244. Harlow New York, 1992.

[52] Tuy, H.: On the convex approximation of nonlinear inequalities. Math. Operationsforsch. Stat. **5** (1974), 451–466.

[53] Vályi, I.: Approximate saddle–point theorems in vector optimization. J. Optim. Theory Appl. **55** No. 3 (1987), 435–448.

[54] White, D. J.: Epsilon efficiency. J. Optim. Theory Appl. **49** No. 2 (1986), 319–337.

[55] Yokoyama, K.: Epsilon approximate solutions for multiobjective programming problems. J. Math. Analysis Appl. **203** (1996), 142–149.

Lebenslauf

1. Angaben zur Person

Name:	Sekatzek
Vorname:	Matthias
akad. Grad:	Diplom-Mathematiker
Geburtsdatum:	05. 01. 1972
Geburtsort:	Merseburg
Staatsangehörigkeit:	deutsch
Anschrift:	Unteraltenburg 41 06217 Merseburg

2. Angaben zum Bildungsgang

1978 - 1988	Besuch der zehnklassigen allgemeinbildenden polytechnischen Oberschule in Merseburg
01. 07. 1988	Erwerb der mittleren Reife
1988 - 1990	Besuch der Spezialklassen für Chemie an der Technischen Hochschule Leuna-Merseburg
29. 06. 1990	Abitur
1990 - 1991	Zivildienst
1991 - 1993	Studium der Mathematik an der Technischen Hochschule Leuna-Merseburg
1993 - 1996	Studium der Mathematik an der Martin-Luther-Universität Halle-Wittenberg
02. 07. 1996	Hochschulabschluß in der Fachrichtung Mathematik, Verleihung des akad. Grades Diplom-Mathematiker
1996 - 1999	Promotionsstudium im Fachbereich Mathematik und Informatik an der Martin-Luther-Universität Halle-Wittenberg als Stipendiat der Graduiertenförderung des Landes Sachsen-Anhalt

Erklärung nach § 5 Abs. 2 lit. d der Promotionsordnung

Hiermit versichere ich, daß ich die vorliegende Arbeit selbständig und ohne fremde Hilfe verfaßt, andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt und die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Merseburg, den 22. 07. 1999 gez. Matthias Sekatzek