Eigenschaften pseudo-regulärer Funktionen und einige Anwendungen auf Optimierungsaufgaben

Dissertation
zur Erlangung des akademischen Grades
doctor rerum naturalium (Dr. rer. nat.)
im Fach Mathematik

eingereicht an der

Mathematisch-Naturwissenschaftlichen Fakultät II
der Humboldt-Universität zu Berlin

von
Dipl.-Math. Peter Fúsek
geb. am 28.11.1970 in Kežmarok

Präsident der Humboldt-Universität zu Berlin
Prof. Dr. Dr. h. c. Hans Meyer

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät II
Prof. Dr. sc. nat. Bodo Krause

Gutachter:

1. Prof. Dr. B. Kummer, Humboldt-Universität zu Berlin

2. Prof. Dr. H. Th. Jongen, RWTH Aachen

3. Prof. Dr. D. Pallaschke, Universität Karlsruhe

Eingereicht: 8.12.1998
Tag der mündlichen Prüfung: 26.02.1999

Inhaltsverzeichnis

1 Einleitung

In der vorliegenden Arbeit beschäftigen wir uns mit der Charakterisierung der Lösungen von Gleichungen der Form $f(x) = y$, wobei f eine Lipschitzstetige, nicht notwendigerweise differenzierbare Funktion ist. Speziell sind wir an Aussagen über die topologische Struktur der Urbildmengen $f^{-1}(y)$ interessiert, wenn x und y dieselbe endliche Dimension besitzen. Dabei geht es vorrangig um die (für Verfahren und Struktur wichtige) Isoliertheit von Nullstellen - für solche Funktionen noch nicht studiert und im Gegensatz zum glatten Fall hier nichttrivial. Bisher war bekannt, daß fast alle Bildpunkte nur isolierte Urbilder besitzen können und daß dies in relevanten Spezialfällen auch für sämtliche Bildpunkte gilt.

Um die Inversen zu studieren, werden gewisse Lipschitz-Eigenschaften der Multifunktion f^{-1} vorausgesetzt. Sie bilden eine Abschwächung der bekannten strengen Regularität im Sinne von Robinson (siehe [36]) bzw. der Regularität in Clarke's Inversen-Satz ([7]). Hinreichende und notwendige Bedingungen für diese abgeschwächte Regularität kann man auf spezielle Problemklassen anwenden (z.B. auf die sog. Kojima-Funktion, deren Nullstellen die kritischen Punkte eines Optimierungsproblems charakterisieren).

Die Beschreibung der Lösungen einer Gleichung der Form $f(x) = y$ wurde von sehr vielen Autoren auf unterschiedlichsten Verallgemeinerungsebenen behandelt. Das wahrscheinlich älteste Regularitätsergebnis stammt von Graves und Lyusternik (dazu siehe [15]):

Es sei $f : X \to Y$ eine stetig differenzierbare Funktion zwischen den Banachräumen X und Y. Dann existieren Lösungen der Gleichung $f(x) = y$ für y "nahe an $f(\bar{x})$", falls $Df(\bar{x})$ surjektiv ist.

Der Beweis bildet auch heute das allgemeine Schema für Regularitätsaussagen, obwohl mit viel schwächeren Forderungen an Differenzierbarkeit und Eindeutigkeit operiert wird.

Unser wichtigster Begriff ist die Pseudo-Regularität einer Abbildung. Wir definieren ihn als die Pseudo-Lipschitz Eigenschaft des inversen Operators.
Diese Eigenschaft (die letzlich die Surjektivität einer linearen Abbildung lokal und nichtlinear verallgemeinert) wurde von Aubin in [2] und [3] eingeführt. Sie hat viele Namen in der Literatur: Metrische Regularität (in [35] von Robinson zur Beschreibung des Lipschitz-stetigen Verhaltens von zulässigen Mengen in parametrischer Optimierung benutzt); Aubin-Eigenschaft (siehe [11]); Offenheit mit linearer Rate. Die Äquivalenz zwischen der metrischen Regularität und der Offenheit mit linearer Rate wurde in [9] und [17] gezeigt. Die Äquivalenz zwischen der metrischen Regularität und der Pseudo-Regularität wurde relativ spät bekannt (wahrscheinlich zum ersten Mal in [6] und [32]).
Im Hinblick auf die Optimierung ist interessant, daß diese Regularitätsbegriffe im glatten Fall gleichbedeutend sind mit gewissen constraint qualifications für die entsprechenden Mengen der zulässigen Punkte (siehe [35]). Weitere Ergebnisse zu den Beziehungen zwischen den oben aufgeführten Begriffen findet man in [8], [10], [29], [30] und [31].

Im ersten Kapitel der Arbeit werden die grundlegenden, hier benutzen Definitionen angegeben und die wichtigsten Eigenschaften der benötigten Begriffe zusammengefaßt.

Im zweiten Abschnitt wird die Pseudo-Regularität äquivalent mit Hilfe von zwei verschiedenen Richtungsableitunskonzepten charakterisiert. Das geschieht in der Form einer Surjektivitätsbedingung für sog. Contingent-Ableitungen bzw. einer Injektivitätsbedingung für sog. Co-Ableitungen. Dabei erweist sich das bekannte Ekeland'sche Variationsprinzip als sehr hilfreich. Außerdem definieren wir mehrere "Kategorien" der Pseudo-Regularität. Die Zugehörigkeit zu einem "Typ" wird durch die Gestalt der Funktion $x(y)$ gegeben, die eine Lösung x (von mehreren möglichen Lösungen) der Inklusion $y \in F(x)$ als eine Funktion von y beschreibt. Wir geben einige Beispiele als Vertreter dieser Kategorien an.

Der Anfang des dritten Abschnitts ist einigen Eigenschaften von verallgemeinerten Richtungsableitungen (Contingent-Ableitungen, siehe [1] und Thibault-Mengen, [45]) gewidmet. Insbesondere sind wir an Kettenregeln und Formeln für partielle Ableitungen interessiert, die z.B. bei Anwendungen auf die Kojima-Funktion von Nutzen sind.
Als nächstes studieren wir die Unterschiede der beiden Regularitätskonzepte (strenge Regularität und Pseudo-Regularität) für lokale Lipschitz-Funktionen in endlichen Dimensionen. Einen wichtigen Spezialfall bilden in diesem Zusammenhang stückweise glatte Funktionen wegen ihrer leicht zu handhabenden Gestalt. Unser Hauptanliegen konzentriert sich auf die Formulierung von Bedingungen, die die lokale Endlichkeit der Urbildmengen sichern.
Für den Fall beliebiger Lipschitz Funktionen wird dabei eine Art Dimensionslemma eine wichtige Rolle spielen, das die Dimensionsgrößen des Urbildraumes und des Bildraumes in Beziehung bringt. Im Falle gleicher Dimensionen werden wir als ein Hauptergebnis erhalten, daß - bei Richtungsdifferenzierbarkeit - die Nullstellen pseudo-regulärer Lipschitz-Funktionen isoliert sind.

Die Ergebnisse aus Kapitel 3 werden im vierten Kapitel auf die Kojima-Funktion zu einem Optimierungsproblem übertragen. Sind die Gradienten der eingehenden Funktionen lokal Lipschitz und richtungsdifferenzierbar, kann man dann die Nebenbedingungen der Null-Lagrange-Multiplikatoren streichen, ohne den Zusammenhang von strenger Regularität und Pseudo-Regularität zu stören. Außerdem untersuchen wir einen Newton-ähnlichen Algorithmus für pseudo-reguläre Gleichungen.

An dieser Stelle möchte ich Prof. Dr. B. Kummer für die intensive Unterstützung und zahlreiche Diskussionen während der Erarbeitung der Dissertation herzlich danken. Für hilfreiche Bemerkungen bin ich dankbar Prof. Dr. D. Klatte (Universität Zürich) und W. Gomez Bofill (Humboldt-Universität zu Berlin). Nicht zuletzt gilt mein Dank dem Land Berlin (Promotionsstipendium gem. NaFöG) und dem Graduiertenkolleg 'Geometrie und nichtlineare Analysis' der HUB für die finanzielle Unterstützung.

1.1 Grundlegende Definitionen und Beziehungen

Der Gegenstand unserer Untersuchungen sind Multifunktionen (auch mengenwertige Abbildungen genannt). Im weiteren Verlauf dieses Abschnittes sollen X und Y Banachräume bezeichnen. Man sagt, daß die Abbildung $F : X \to Y$ eine **Multifunktion** ist, wenn für alle $x \in X$ gilt: $F(x) \subset Y$. Der Graph $\operatorname{Gph} F$ von F und die Urbildmengen $F^{-1}(y)$ zu einem $y \in Y$ seien wie folgt definiert:

$\operatorname{Gph} F := \{(x, y) \in X \times Y \mid y \in F(x)\}$,
$F^{-1}(y) := \{x \in X \mid y \in F(x)\}$.

Wir nennen eine Multifunktion F **abgeschlossen,** wenn der Graph $\operatorname{Gph} F$ abgeschlossen ist. Ein wichtiges Beispiel einer Multifunktion ist die inverse Abbildung f^{-1} zu einer Funktion f. Sie unterscheidet sich von allgemeinen Multifunktionen nur durch die Eigenschaft, daß ihre Bilder paarweise disjunkt sind.

Ähnlich wie bei Funktionen, spielen auch bei Multifunktionen ihre Stetigkeitseigenschaften eine wichtige Rolle. Die nächste Definition stellt einige der wichtigsten Stetigkeitsbegriffe für mehrwertige Abbildungen vor.

Definition 1.1 *Es sei $F : X \to Y$ eine Multifunktion.*

Man sagt, daß die Abbildung F **unterhalb stetig** *in einem Punkt x^0 ist, wenn für jede offene Menge $\Omega \subset Y$ mit $F(x^0) \cap \Omega \neq \emptyset$ eine Umgebung U von x^0 ($U \subset X$) existiert, so daß gilt:*

$$F(x) \cap \Omega \neq \emptyset \qquad \forall x \in U.$$

Man sagt, daß die Abbildung F **oberhalb stetig** *in einem Punkt x^0 ist, wenn für jede offene Menge $\Omega \subset Y$ mit $F(x^0) \subset \Omega$ eine Umgebung U von x^0 ($U \subset X$) existiert, so daß gilt:*

$$F(x) \subset \Omega \qquad \forall x \in U.$$

Eine Abbildung, die in einem Punkt x^0 unter und oberhalb stetig ist, nennen wir **stetig** *in x^0.*

Mit diesen Definitionen folgten wir C. Berge. Eine stärkere Version der Stetigkeit ist im Falle einer Funktion durch die bekannte lokale Lipschitzstetigkeit gegeben. Wir möchten diesen Begriff auf Multifunktionen übertragen. Dazu gibt es mehrere Möglichkeiten.

Definition 1.2 *Es sei $F : X \to Y$ eine Multifunktion mit abgeschlossenen Bildern $F(x)$ für alle $x \in X$. Man nennt die Abbildung F* **Lipschitz** *auf einer Menge U ($U \subset X$), falls eine Konstante $L > 0$ existiert, so daß gilt:*

$$d(F(x_1), F(x_2)) \leq L\|x_1 - x_2\| \qquad \forall x_1, x_2 \in U. \tag{1.1}$$

Die Konstante L heißt dann Modul der Lipschitzstetigkeit oder die Lipschitz-Konstante von F.

Dabei bezeichnet $d(A_1, A_2)$ den **Hausdorff-Abstand** $d(A_1, A_2)$ der Mengen A_1 und A_2 $(A_1, A_2 \subset X)$:

$d(A_1, A_2) := \inf\{\varepsilon \geq 0 \,|\, A_1 \subset A_2 + \varepsilon I\!\!B_X,\ A_2 \subset A_1 + \varepsilon I\!\!B_X\}$,

wobei $I\!\!B_X$ die abgeschlossene Einheitskugel im Raum X ist und die Summe von Mengen punktweise (im Minkovski-Sinne) zu verstehen ist. Der Hausdorff-Abstand eines Punktes (aufgefaßt als einelementige Menge) und einer Menge ist der gewöhnliche Punkt-Menge Abstand.
Wenn es keine Mißverständnisse geben kann, lassen wir den Index des Raumes im Zusammenhang mit dem Abstand weg. Ferner benutzen wir die Bezeichnungen $I\!\!B^0$ für die offene Einheitskugel und S für die Sphäre (im jeweiligen Raum).
Außerdem definieren wir den **Durchmesser** diam A einer Menge $A \subset X$ wie üblich als:

$\operatorname{diam} A := \sup\{\|x_1 - x_2\| \,|\, x_1, x_2 \in A\}$.

Eine äquivalente Formulierung der Bedingung (1.1) lautet dann unter Benutzung des Punkt-Menge Abstandes:

$$F(x_1) \subset F(x_2) + L\|x_1 - x_2\| I\!\!B \qquad \forall x_1, x_2 \in U. \tag{1.2}$$

Wenn F eine Funktion ist, fällt die Lipschitz-Eigenschaft aus Definition 1.2 mit der lokalen Lipschitzstetigkeit für Funktionen zusammen.

Sind die Bildmengen $F(x)$ nicht beschränkt, birgt die Definition 1.2 eine ziemlich starke Beschränkung in sich. Man stelle sich folgende Multifunktion $F : I\!\!R \to I\!\!R^2$ vor: zu jedem $t \in I\!\!R$ sei das Bild $F(t)$ durch den Strahl gegeben, der im Nullpunkt beginnt und mit der x-Achse im Bildraum den Winkel t bildet. Die Abbildung F ist nicht mal stetig im Sinne des Hausdorff-Abstandes der Bilder, obwohl wir sie als "stetig" empfinden. Um diese Diskrepanz zu eliminieren, wurde von Aubin in [2] der Begriff der pseudo-Lipschitz-stetigen Abbildung definiert:

Definition 1.3 *Es sei $F : X \to Y$ eine Multifunktion und $(x^0, y^0) \in \operatorname{Gph} F$. Wir nennen F* **pseudo-Lipschitz in $(\mathbf{x^0}, \mathbf{y^0})$**, *falls es Umgebungen U von x^0, V von y^0 und eine Konstante $L > 0$ gibt, so daß gilt:*

$$F(x_1) \cap V \subset F(x_2) + L\|x_1 - x_2\| I\!\!B \qquad \forall x_1, x_2 \in U. \tag{1.3}$$

Ein anderer Name für solche Abbildungen wurde in [11] eingeführt. Man sagt dann, die Multifunktion F erfüllt die **Aubin-Eigenschaft.** Ein Vergleich mit (1.2) ergibt, daß es sich hier um eine Stetigkeit, lokal eingeschränkt im Bildraum handelt.

In [40] finden wir folgende Charakterisierung der pseudo-Lipschitz Eigenschaft:

Eine Multifunktion $F : I\!R^n \to I\!R^m$ mit abgeschlossenen Bildern ist pseudo-Lipschitz in einem Punkt $(x^0, y^0) \in \operatorname{Gph} F$ genau dann, wenn die Funktion $f(x, y) := d(y, F(x))$ lokal Lipschitz in dem Punkt (x^0, y^0) ist.

Diese Beschreibung ist aber nicht geeignet, um die Pseudo-Lipschitz Eigenschaft der Inversen F^{-1} von F zu untersuchen.

Bemerkung 1.4 Wir erhalten zunächst einige einfache Folgerungen aus dem Fakt, daß F in einem Punkt (x^0, y^0) pseudo-Lipschitz ist:

(i) Es gibt eine Umgebung $\tilde{U}$ von x^0, so daß $F(x) \cap V \neq \emptyset \ \forall x \in \tilde{U}$.

Dazu sei $V = y^0 + \varepsilon I\!B^0$ die Umgebung aus der Definition 1.3. Wir definieren dann $\tilde{U}$ als $\tilde{U} := \left(x^0 + \frac{\varepsilon}{L} I\!B^0\right) \cap U$. Aus der Definition 1.3 folgt:

$$y^0 \in F(x) + L\|x - x^0\| I\!B \qquad \forall x \in U.$$

Für $x \in \tilde{U}$ erhalten wir die Existenz eines $y \in F(x)$ mit

$$\|y - y^0\| \leq L\|x - x^0\| < \varepsilon$$

und es gilt $y \in F(x) \cap V$.

(ii) Es gibt eine Umgebung $W \subset X \times Y$ von (x^0, y^0), so daß F pseudo-Lipschitz in jedem Punkt $(x, y) \in W \cap \operatorname{Gph} F$ ist (mit derselben Konstante L).

(iii) Für $y^0 \in F(x^0) \cap V$ gilt:

$$d(y^0, F(x)) \leq L\|x^0 - x\| \qquad \forall x \in U.$$

Dann sagt man, daß F **Lipschitz unterhalb stetig** im Punkt (x^0, y^0) ist (vergleiche mit Definition 1.1).

(iv) In der Definition 1.3 kann man die Bedingung (1.3) durch folgende äquivalente Bedingung (mit x_1 aus X) ersetzen:

$$F(x_1) \cap V \subset F(x_2) + L\|x_1 - x_2\| I\!B \qquad \forall x_2 \in U \qquad \forall x_1 \in X. \tag{1.4}$$

Aus (1.4) folgt sofort (1.3). Umgekehrt: die Umgebungen aus (1.3) seien gegeben durch $U = x^0 + \delta I\!B^0$ und $V = y^0 + \varepsilon I\!B^0$. Wir definieren neue Umgebungen U' und V' wie folgt: $U' := x^0 + \delta' I\!B^0$, $V' := y^0 + \varepsilon' I\!B^0$ mit $0 < \delta' < \delta$, $0 < \varepsilon' < \varepsilon$ und $2L\delta' + \varepsilon' \leq L\delta$. Für $x_1 \in U$ und $x_2 \in U'$ folgt die gewünschte Inklusion aus (1.3).

Wir untersuchen den Fall $x_1 \in X \setminus U$. Es gilt: $y^0 \in F(x_2) + L\|x^0 - x_2\| I\!B$, also erhalten wir

$$F(x_1) \cap V' \subset V' = y^0 + \varepsilon' I\!B^0 \subset F(x_2) + (L\delta' + \varepsilon')I\!B \subset F(x_2) + (L\delta - L\delta')I\!B \subset$$
$$\subset F(x_2) + L\|x_1 - x_2\| I\!B,$$

weil $\|x_1 - x_2\| \geq \|x_1 - x^0\| - \|x_2 - x^0\| \geq \delta - \delta'$. Mit anderen Worten: Der Punkt x_1 ist so weit von x_2 entfernt, daß die ganze Umgebung V' in die Menge $F(x_2) + L\|x_1 - x_2\| I\!B$ hineinpaßt.

Beispiel 1.5 Wir betrachten folgendes Optimierungsproblem:

$$P(y) \qquad \min_x f(x,y) \text{ mit } x \in F(y)$$

wobei F eine mengenwertige Abbildung und y ein Parameter ist. Wenn F in einem Punkt (x^0, y^0) pseudo-Lipschitz ist, können wir Abschätzungen für Optimalwerte angeben, insbesondere besitzt die Abbildung Ψ, die die Lösungsmenge $\Psi(y)$ von $P(y)$ beschreibt, auch gewisse Lipschitz-Eigenschaften (siehe [43], [44]). Dies ist wichtig z.B. im Fall eines Zwei-Ebenen-Problems (d.h. die Lösungen von $P(y)$ bilden die Eingangsdaten für ein zweites Optimierungsproblem). Allerdings ist Ψ nur unter sehr einschränkenden Voraussetzungen (Lipschitz) unterhalb stetig.

Definition 1.6 *Es sei eine Multifunktion $F : X \to Y$ gegeben und $(x^0, y^0) \in \operatorname{Gph} F$. Dann heißt F* **pseudo-regulär in** *einem Punkt $(x^0, y^0) \in \operatorname{Gph} F$, falls F^{-1} pseudo-Lipschitz in dem Punkt (y^0, x^0) ist.*

Wenn zusätzlich die Mengen $F^{-1}(y) \cap V$ für alle $y \in U$ einelementig sind (dabei sind U und V die entsprechenden Umgebungen aus der Definition 1.3), dann nennen wir F **streng regulär in $(\mathbf{x^0}, \mathbf{y^0})$.**

Wenn F eine Funktion ist, kürzen wir die Schreibweise ab, indem wir die Abbildung F pseudo- bzw. streng regulär in x^0 nennen (anstatt in $(x^0, F(x^0))$).

Eine Charakterisierung der pseudo-Regularität für Abbildungen mit abgeschlossenem und konvexem Graphen ist durch den bekannten Robinson-Ursescu-Satz gegeben (bewiesen unabhängig von Robinson in [35] und Ursescu in [46]):

Es sei $F : X \to Y$ eine abgeschlossene Multifunktion mit konvexem Graphen und $(x^0, y^0) \in \operatorname{Gph} F$. Dann gilt: F ist pseudo-regulär in (x^0, y^0) genau dann, wenn

$$y^0 \in \text{ core Im } F.$$

Dabei bezeichnet $\operatorname{Im} F$ das Bild von F, und der core einer Menge $A \subset X$ ist die Menge aller Punkte x, so daß gilt: $\forall z \in X,\ z \neq x$ gibt es einen Punkt y, der auf der (offenen) Strecke (x, z) liegt, so daß für die (abgeschlossene) Strecke $[x, y]$ gilt: $[x, y] \subset A$. Zum Beispiel ist der core der C^1-Funktionen im Raum aller stetigen Funktionen leer.

Sehr interessant ist folgende Eigenschaft pseudo-regulärer Multifunktionen:

Satz 1.7 *([26])*
Es sei $F : X \to Y$ eine abgeschlossene Multifunktion, die pseudo-regulär im Punkt $(x^0, y^0) \in \operatorname{Gph} F$ mit der Konstante L ist. OBdA seien die Umgebungen aus (1.3) durch $U = y^0 + \delta I\!B^0$ und $V = x^0 + \delta I\!B^0$ gegeben, $\delta < 1$. Ferner sei $g : X \to Y$ eine lokale Lipschitz-Funktion mit der Lipschitz-Konstante L_g und $|g| < \frac{\delta}{3(L+1)}$ (wobei man $|g| := \max\{\sup_{x\in V} \|g(x)\|, L_g\}$ definiert). Dann gilt:

Die Abbildung $F + g$ ist pseudo-regulär mit der Konstante $2L$ im Punkt (x, y^0), falls gilt $x \in x^0 + \frac{\delta}{3} I\!B^0$ und $(x, y^0) \in \operatorname{Gph}(F + g)$. Die entsprechenden Umgebungen sind dann beschrieben durch $V_g = x + \frac{\delta}{3} I\!B^0$ und $U_g = y^0 + \varepsilon I\!B^0$ mit $0 < \varepsilon < \min\{\frac{\delta}{3}, \frac{\delta}{12(L+1)}\}$.

Hierbei handelt es sich um eine Verallgemeinerung des Satzes 2.1 aus [8]. Derselbe Fakt ist aus den Arbeiten von Robinson zur strengen Regularität wohlbekannt. Letztere bedeutet: Die Abbildung F^{-1} ist lokal eindeutig und Lipschitz. Der Begriff der strengen Regularität wurde in [36] eingeführt. Robinson zeigte in seinem Artikel, daß die Lösungen einer Variationsungleichung sich bezüglich des Parameters streng regulär verhalten, wenn dies nur für die "linearisierte" Variationsungleichung richtig ist.

Definition 1.8 *Es sei $F : X \to Y$ eine Multifunktion und $(x^0, y^0) \in \operatorname{Gph} F$. Dann heißt F* **metrisch regulär in $(\mathbf{x^0}, \mathbf{y^0})$**, *wenn es Umgebungen U von x^0, V von y^0 und eine Konstante $L > 0$ gibt, so daß gilt:*

$$d(x, F^{-1}(y)) \leq L d(y, F(x)) \qquad \forall x \in U \;\; \forall y \in V. \tag{1.5}$$

Eine einfache Interpretation ist in dem Fall möglich, wenn $F(x)$ die Menge der zulässigen Punkte zu einem Optimierungsproblem bezeichnet. Der Term $d(y, F(x))$ mißt die Unzulässigkeit des Punktes y im Bildraum von F. Der zweite Term ist dann die Größe der "Unzulässigkeit" von x im Urbildraum X.

Beispiel 1.9 In [35] beschäftigte sich Robinson mit den Mengen der Gestalt $M(a, b) := \{x \in I\!R^n \mid g(x) \leq a,\ h(x) = b\}$, wobei $g : I\!R^n \to I\!R^m, h : I\!R^n \to I\!R^k$ stetig differenzierbare Funktionen sind. Es sei ein Punkt x^0 gegeben mit $g(x^0) \leq 0$ und $h(x^0) = 0$. Die Abbildung M^{-1} ist metrisch regulär in $(x^0, (0, 0))$ genau dann, wenn die folgenden zwei Bedingungen erfüllt sind (siehe [35]):

(i) die Menge der Gradienten $\{Dh_j(x^0)\}_{j=1,\ldots,k}$ ist linear unabhängig

(ii) es gibt einen Vektor $u \in I\!R^n$, so daß gilt:

$$\begin{aligned} &Dh_j(x^0)u = 0 \qquad && \forall j = 1, \ldots, k, \\ &Dg_i(x^0)u < 0 \qquad && \forall i \in I(x^0) := \{i \in \{1, \ldots, m\} \mid g_i(x^0) = 0\}. \end{aligned} \tag{1.6}$$

Sind die Bedingungen (i) und (ii) erfüllt, sagt man, daß die **Mangasarian-Fromovitz Constraint Qualification** (kurz MFCQ) gilt.

Falls die Menge der Gradienten $\{Dh_j(x^0)\}_{j=1,\dots,k} \cup \{Dg_i(x^0)\}_{i\in I(x^0)}$ linear unabhängig ist, sagt man, daß die **Linear Independence Constraint Qualification** (LICQ) erfüllt ist.

Definition 1.10 *Es sei $F : X \to Y$ eine Multifunktion und $(x^0, y^0) \in \operatorname{Gph} F$. Dann heißt F* **offen mit linearer Rate in** *(x^0, y^0), falls Umgebungen U von x^0, V von y^0 und eine Konstante $L > 0$ existieren, so daß gilt:*

$$(F(x) + \varepsilon I\!B) \cap V \subset F(x + L\varepsilon I\!B) \qquad \forall x \in U \ \forall \varepsilon > 0. \tag{1.7}$$

Der folgende Satz summiert die Beziehungen zwischen den bis jetzt definierten Begriffen (dazu siehe auch [6], [9], [17], [32] und [41] , wo meistens noch allgemeinere Aussagen bewiesen werden):

Satz 1.11 *Es seien eine abgeschlossene Multifunktion $F : X \to Y$ und ein Punkt $(x^0, y^0) \in \operatorname{Gph} F$ gegeben. Dann sind folgende Aussagen äquivalent:*

(i) F ist pseudo-regulär in (x^0, y^0) (d.h. F^{-1} ist pseudo-Lipschitz in (y^0, x^0))

(ii) F ist metrisch regulär in (x^0, y^0)

(iii) F ist offen mit linearer Rate in (x^0, y^0)

Beweis: (i) $\Leftrightarrow$ (ii): Aus der Bemerkung 1.4 (iv) folgt, daß folgende Aussage zur Aussage (i) äquivalent ist:
Es gibt Umgebungen U von x^0, V von y^0 und eine Konstante $L > 0$ mit

$$F^{-1}(y') \cap U \subset F^{-1}(y) + L\|y' - y\| I\!B \qquad \forall y \in V \ \forall y' \in Y. \tag{1.8}$$

In anderen Worten: Es seien Punkte $y' \in Y$, $x' \in F^{-1}(y') \cap U$ und $y \in V$ gegeben. Wir finden einen Punkt $x \in F^{-1}(y)$, so daß gilt:

$$\|x' - x\| \le L\|y' - y\|. \tag{1.9}$$

Das ist äquivalent zu

$$d(x', F^{-1}(y)) \le L\|y' - y\| \qquad \forall x' \in U \ \forall y' \in F(x') \ \forall y \in V. \tag{1.10}$$

(aus (1.9) folgt $d(x', F^{-1}(y)) \le \|x' - x\| \le L\|y' - y\|$. Umgekehrt, wähle $x \in F^{-1}(y)$ so, daß $\|x' - x\| = d(x', F^{-1}(y))$ und (1.9) gilt.)
Die Minimierung der rechten Seite in (1.10) über $y' \in F(x')$ ergibt

$$d(x', F^{-1}(y)) \le L d(y, F(x')) \qquad \forall x' \in U \ \forall y \in V. \tag{1.11}$$

(ii) $\Rightarrow$ (iii): Aus (1.11) folgt:

$\forall x' \in U \ \forall y \in V$:
$d(y, F(x')) \leq \varepsilon \Rightarrow d(x', F^{-1}(y)) \leq L\varepsilon$

oder anders ausgedrückt:

$$\forall x' \in U \ \forall \varepsilon > 0 : \qquad (F(x') + \varepsilon I\!B) \cap V \subset F(x' + L\varepsilon I\!B).$$

(iii) $\Rightarrow$ (ii): Wir wählen $x' \in U,\ y \in V$ und $\varepsilon := d(y, F(x')) > 0$ (d.h. $y \in (F(x') + \varepsilon I\!B) \cap V$). Aus (iii) erhalten wir $y \in F(x' + L\varepsilon I\!B)$, also gilt: $d(x', F^{-1}(y)) \leq L\varepsilon = Ld(y, F(x'))$.
#

2 Eigenschaften pseudo-regulärer Multifunktionen

Im folgenden Abschnitt charakterisieren wir pseudo-reguläre Multifunktionen mit Hilfe von sogenannten verallgemeinerten Ableitungen für mehrwertige Abbildungen. Es wurden relativ viele Konzepte für die Differentiation von Multifunktionen entwickelt; für unsere Zwecke benutzen wir die von Aubin definierten Contingent-Ableitungen und stellen die von Mordukhovich eingeführten Co-Ableitungen vor. Im zweiten Teil des Abschnitts werden pseudo-reguläre Abbildungen in Kategorien eingeteilt; das Kriterium dafür ist die lokale Gestalt der Funktion $x(y)$, die eine Lösung x (der mehreren möglichen Lösungen) der Inklusion $y \in F(x)$ beschreibt.

2.1 Ableitungen für Multifunktionen

Mit X und Y bezeichnen wir im Folgenden Banachräume. Um die gewünschten Ableitungskonzepte definieren zu können, brauchen wir zunächst die entsprechenden Kegel:

Definition 2.1 *Es sei $C \subset X$, $x^0 \in C$. Man definiert den* **Contingent-Kegel $\mathbf{T_C(x^0)}$** *zu C in x^0 als*

$$T_C(x^0) := \left\{ u \in X \mid \exists \{x^k\},\ x^k \in C,\ x^k \to x^0,\ \exists \{t_k\},\ t_k \searrow 0,\ \frac{x^k - x^0}{t_k} \to u \right\} \tag{2.1}$$

und einen **Normalenkegel $\mathbf{N_C(x^0)}$** *zu C in x^0 als*

$$\begin{aligned} N_C(x^0) \quad := \quad & \Big\{ w \in X^* \mid w = \lim_{k\to\infty} w^k, \\ & \exists \{x^k\},\ x^k \in C,\ x^k \to x^0,\ < w^k, x - x^k > \le o(\|x - x^k\|)\ \forall x \in C \Big\} \end{aligned} \tag{2.2}$$

wobei es sich in (2.2) um die schwache Konvergenz im dualen Raum X^* handelt. Die Funktion $o(.)$ darf dabei von den Folgen der x^k abhängen.

Man erkennt leicht, daß $T_C(x^0)$ und $N_C(x^0)$ tatsächlich Kegel sind. Mit ihrer Hilfe lassen sich folgende verallgemeinerte Ableitungen für Multifunktionen erklären:

Definition 2.2 *Es sei $F : X \to Y$ eine Multifunktion, $(x^0, y^0) \in \operatorname{Gph} F$. Man definiert die* **Contingent-Ableitung $\mathbf{C((x^0, y^0); u)}$** *im Punkt (x^0, y^0) in Richtung $u \in X$ als*

$$CF((x^0, y^0); u) := \Big\{ v \in Y \mid (u, v) \in T_{\operatorname{Gph} F}(x^0, y^0) \Big\} \tag{2.3}$$

und die **Co-Ableitung** $\mathbf{D^*F((x^0, y^0); w)}$ *im Punkt* (x^0, y^0) *in Richtung* $w \in Y^*$ *als*

$$D^*F((x^0, y^0); w) := \left\{ z \in X^* \mid (z, -w) \in N_{\mathrm{Gph}\, F}(x^0, y^0) \right\}. \tag{2.4}$$

Wenn F *eine Funktion ist, benutzen wir eine vereinfachte Schreibweise: Statt* $CF((x^0, y^0); u)$ *schreiben wir* $CF(x^0, u)$*, und an Stelle von* $D^*F((x^0, y^0); w)$ *analog* $D^*F(x^0, w)$*.*

Die Contingent-Ableitung wurde von Aubin in [1] eingeführt. Grundlegende Untersuchungen hierzu kann man in [4] und [5] finden. Die obigen Co-Ableitungen wurden (basierend auf dem Begriff sogenannter approximate Fréchet Normalen) von Mordukhovich entwickelt und zur Analyse der Pseudo-Regularität in seinen Arbeiten [28], [29], [30],[31] angewandt. Ein wichtiger Unterschied zwischen beiden Ableitungskonzepten ist die Tatsache, daß es sich bei den Contingent-Ableitungen um eine primale und bei den Co-Ableitungen um eine duale Information handelt: Während $CF((x^0, y^0); .)$ von X nach Y abbildet, ist $D^*F((x^0, y^0); .)$ eine mehrwertige Abbildung vom Typ $Y^* \to X^*$. Außerdem existiert zu dem Kegel $N_C(x^0)$ im allgemeinen kein entsprechender Tangentenkegel.

Bemerkung 2.3

(i) Wenn man die Definition 2.1 berücksichtigt, bekommt man für die Contingent-Ableitung folgende Gestalt:

$$\begin{aligned} CF((x^0, y^0); u) \;\; = \;\; & \Big\{ v \in X \mid \exists \{t_k\},\ t_k \searrow 0,\ \exists \{u^k\},\ u^k \to u, \\ & \exists\, \{y^k\},\ y^k \in F(x^0 + t_k u^k) :\ \frac{y^k - y^0}{t_k} \to v \Big\} \end{aligned} \tag{2.5}$$

Ist F eine Funktion und ist die Menge $CF(x^0, u)$ einelementig, sagen wir, daß F **richtungsdifferenzierbar** ist. Mit $F'(x^0, u)$ bezeichnen wir die **Richtunsableitung**; dann ist $F'(x^0, u) = v \Leftrightarrow CF(x^0, u) = \{v\}$. Ein allgemeinerer Typ der Richtungsdifferenzierbarkeit ist die sog. **B-differenzierbarkeit,** die von Robinson in [37] eingeführt wurde. Für endlichdimensionale Lipschitz-Funktionen fallen diese Begriffe allerdings zusammen.

(ii) Aus der Tatsache, daß

$$(u, v) \in T_{\mathrm{Gph}\, F}(x^0, y^0) \Leftrightarrow (v, u) \in T_{\mathrm{Gph}\, F^{-1}}(y^0, x^0)$$

erhalten wir folgende **Symmetrie-Eigenschaft** für Contingent-Ableitungen:

$$v \in CF((x^0, y^0); u) \Leftrightarrow u \in CF^{-1}((y^0, x^0); v). \tag{2.6}$$

(iii) Es sei F eine lokale Lipschitz-Funktion. Dann können wir in der Darstellung (2.5) die Folge $\{u^k\}$ konstant setzen ($u^k := u$) und erhalten

$CF(x^0, u) = \Big\{ v \in X \mid \exists \{t_k\},\ t_k \searrow 0,\ \frac{1}{t_k}\Big[F(x^0 + t_k u) - F(x^0)\Big] \to v \Big\}$, weil

$\Big\| \frac{1}{t_k}\Big[F(x^0 + t_k u^k) - F(x^0 + t_k u)\Big]\Big\| \le L\|u^k - u\| \to 0$ für $u^k \to u$ gilt.

(iv) Für stetig differenzierbare Funktionen reduziert sich die Contingent-Ableitung bzw. Co-Ableitung auf die Fréchet-Ableitung bzw. die konjugierte Abbildung dazu.

Definition 2.4 *Es sei $F : X \to Y$ eine lokale Lipschitz-Funktion. Wir definieren die* **Thibault-Menge** $TF(x^0, u)$ (*auch Thibault'sche Richtungsableitung genannt) als*

$$TF(x^0, u) := \Big\{ v \in X \mid \exists \{x^k\},\ x^k \to x^0,\ \exists \{t_k\},\ t_k \searrow 0 : \frac{1}{t_k}\Big[F(x^k + t_k u) - F(x^k)\Big] \to v \Big\} \tag{2.7}$$

Das erste Mal wurden diese Mengen von Thibault in [45] benutzt. Eine Zusammenfassung der Eigenschaften der Thibault-Mengen hat Kummer in [23] gegeben. Der wesentliche Unterschied zwischen Contingent-Ableitungen und Thibault-Mengen besteht bei Lipschitz-Funktionen darin, daß nun auch der Fußpunkt x^k des Differenzenquotienten variieren darf. Folgendes Beispiel illustriert die definierten Richtungsableitungen auf einer einfachen Lipschitz-Funktion f:

Beispiel 2.5 Die Funktion $f : I\!R \to I\!R$ sei gegeben durch $f(x) = x$ für $x \ge 0$ und $f(x) = 2x$ für $x < 0$. Im Punkt $x^0 = 0$ erhalten man dann:

$$\begin{array}{llllll}
Cf(0,u) = \{u\} & \text{für} & u \ge 0, & Cf(0,u) = \{2u\} & \text{für} & u < 0 \\
D^*f(0,w) = \{w, 2w\} & \text{für} & w \ge 0, & D^*f(0,w) = [2w, w] & \text{für} & w < 0 \text{ und} \\
Tf(0,u) = [u, 2u] & \text{für} & u \ge 0, & Tf(0,u) = [2u, u] & \text{für} & u < 0.
\end{array}$$

Es ist bekannt, daß die Richtungsableitung $F'(x, .)$ einer richtungsdifferenzierbaren lokalen Lipschitz-Funktion F auch Lipschitz ist. Das nächste Lemma erweitert dieses Ergebnis auf Multifunktionen und ihre Contingent-Ableitungen (in endlichen Dimensionen).

Lemma 2.6 *Es sei $F : I\!R^n \to I\!R^m$ eine Multifunktion, $(x^0, y^0) \in \operatorname{Gph} F$, und F sei pseudo-Lipschitz in (x^0, y^0). Dann ist die Contingent-Ableitung $CF((x^0, y^0), .)$ Lipschitz mit derselben Lipschitz-Konstante.*

Beweis: Um die Schreibweise zu vereinfachen, setzen wir: $G(u) := CF((x^0, y^0), u)$. Mit U bzw. V bezeichnen wir die entsprechenden Umgebungen von x^0 bzw. y^0 wie in Definition 1.3. Nun seien folgende Punkte gegeben: $u_1, u_2 \in I\!R^n$ und $v_1 \in G(u_1)$, d.h. $v_1 = \lim\limits_{k\to\infty} v_1^k$,

$$v_1^k = \frac{1}{t_k}(y_1^k - y^0), \quad y_1^k \in F(x^0 + t_k u_1^k), \quad u_1^k \to u_1, \quad t_k \searrow 0.$$

Daraus folgt $\|y_1^k - y^0\| = t_k \|v_1^k\| \searrow 0$ (weil die Folge $\{v_1^k\}$ beschränkt ist). Für hinreichend große k erhalten wir $x^0 + t_k u_1^k \in U$ und $y_1^k \in V$.

Wir setzen $u_2^k := u_2 + (u_1^k - u_1)$ und gegebenenfalls vergrößern wir k, damit $x^0 + t_k u_2^k \in U$ gilt. Aus der pseudo-Lipschitzstetigkeit von F in (x^0, y^0) folgt die Existenz eines Punktes $y_2^k \in F(x^0 + t_k u_2^k)$ mit

$$\|y_1^k - y_2^k\| \leq L t_k \|u_1^k - u_2^k\| = L t_k \|u_1 - u_2\|. \tag{2.8}$$

Wenn wir jetzt $v_2^k := \frac{1}{t_k}(y_2^k - y^0)$ setzen, ergibt das

$$\|v_1^k - v_2^k\| = \frac{1}{t_k}\|y_1^k - y_2^k\| \leq L\|u_1 - u_2\|.$$

Wir können annehmen, daß eine Teilfolge von $\{v_2^k\}$ existiert, die gegen ein geeignetes v_2 konvergiert. Damit gilt $v_2 \in G(u_2)$ und $\|v_2 - v_1\| \leq L\|u_2 - u_1\|$. #

Korollar 2.7 *Es sei $f : I\!R^n \to I\!R^n$ lokal Lipschitz und pseudo-regulär in x^0. Dann sind die Abbildungen G und G^{-1} beide Lipschitz, wobei $G(u) := Cf(x^0, u)$ ist.*

Beweis: Die Behauptung folgt sofort aus Lemma 2.6 (angewendet auf die Funktion f und die Multifunktion f^{-1}) zusammen mit der Symmetrie-Eigenschaft (2.6) #

2.2 Pseudo-Regularität und Ableitungen von Multifunktionen

Mit Hilfe der Contingent-Ableitungen oder der Co-Ableitungen kann man die Pseudo-Regularität von Multifunktionen auf eine kompakte Art und Weise charakterisieren. Allerdings bekommt man eine äquivalente Umformulierung in beiden Fällen nur bei endlicher Dimension des Bild- und Urbildraumes. Man erhält eine Surjektivitätsbedingung für Contingent-Ableitungen und eine Injektivitätsbedingung für Co-Ableitungen.

Satz 2.8 *([3], [25], [29], [30], [31])*
Es sei $F : I\!R^n \to I\!R^m$ eine abgeschlossene Multifunktion, $(x^0, y^0) \in \operatorname{Gph} F$. Dann sind folgende Aussagen äquivalent:

(i) F ist pseudo-regulär in (x^0, y^0)

(ii) es gibt eine Umgebung W von (x^0, y^0) und eine Konstante $\alpha > 0$ mit

$$\mathbb{B} \subset \bigcup_{\|u\| \leq \alpha} CF((x,y);u) \qquad \forall (x,y) \in W \cap \operatorname{Gph} F$$

(iii) $0 \in D^*F((x^0, y^0); w) \;\Rightarrow\; w = 0$

Wir zeigen nur die Äquivalenz zwischen (i) und (ii), da wir mit den Co-Ableitungen nicht mehr arbeiten werden.

Die Bedingung (ii) (siehe [25]) ist eine vereinfachte Version der Bedingung, die Aubin/Ekeland in [3] benutzt haben. Die Injektivitätsbedingung (iii) stammt von Mordukhovich. Ursprünglich hat er die Äquivalenz zwischen (iii) und der Offenheit mit linearer Rate gezeigt.

Auf den ersten Blick scheint die Bedingung (iii) geeigneter zu sein, um die Pseudo-Regularität von F zu überprüfen (sie sieht wie eine "Punkt-Bedingung" aus, dagegen muß (ii) auf einer ganzen Umgebung von (x^0, y^0) erfüllt sein). Andererseits, wenn man die Definition des Normalkegels bedenkt, ist dies auch eine Bedingung auf einer ganzen Umgebung.

Die Bedingung (iii) bildet eine Brücke zu vielen anderen Gebieten. Andererseits ist sie ziemlich schwer zu handhaben. Eine Hilfe dabei sind einige Kettenregeln und Rechenregeln für Co-Ableitungen (siehe z.B. [29], [30], [31]).

Bevor wir den Satz 2.8 beweisen, formulieren wir ein wichtiges Hilfsmittel für Untersuchungen in diesem Zusammenhang: Das Ekeland'sche Variationsprinzip ([12]). Die Idee, es in diesem Kontext zu nutzen, stammt von Aubin und Ekeland (siehe [3]). Ein typisches Beispiel ist der Beweis von Satz 4 in Abschnitt 7.5 in [3]. Den Beweis von Satz 2.9 führen wir nicht auf; man findet ihn z.B. in [3] oder [7].

Satz 2.9 *(Das Ekeland'sche Variationsprinzip)*
Es sei X ein vollständiger metrischer Raum und $f : X \to \mathbb{R} \cup \{+\infty\}$ ein unterhalb stetiges Funktional (d.h. die Niveau-Mengen $\{x \in X \,|\, f(x) \leq a\}$ sind abgeschlossen), das von unten beschränkt ist. Weiter seien $\varepsilon, \delta > 0$ und ein $\bar{x} \in X$ gegeben, so daß $f(\bar{x}) \leq \inf_x f(x) + \varepsilon$ gilt.

Dann existiert ein $x^* \in X$ mit folgenden Eigenschaften:

$$\begin{aligned} d(x^*, \bar{x}) &\leq \delta \\ f(x^*) &\leq f(\bar{x}) \\ f(x) + \frac{\varepsilon}{\delta} d(x, x^*) &\geq f(x^*) \qquad \forall x \in X \end{aligned}$$

Beweis von Satz 2.8 (i) $\Leftrightarrow$ (ii):
(i) $\Rightarrow$ (ii): Die Umgebungen U, V und die Konstante L seien wie in der Definition 1.3, d.h.

$$F^{-1}(y_1) \cap V \subset F^{-1}(y_2) + L\|y_1 - y_2\| I\!\!B \qquad \forall y_1, y_2 \in U.$$

Wir definieren $W := V \times \tilde{U}$ und $\alpha := L$ wobei $\tilde{U}$ die Umgebung aus der Bemerkung 1.4(i) ist. Nun sei $(x, y) \in W \cap \operatorname{Gph} F$ und ein $v \in I\!\!B$ gegeben. Für t hinreichend klein ist $y + tv \in \tilde{U}$ und $F^{-1}(y + tv) \cap V \neq \emptyset$, d.h. es existiert ein $x(t) \in F^{-1}(y + tv) \cap V$ und es gilt $\|x(t) - x\| \leq \alpha t\|v\|$. Wir setzen

$$u(t) := \frac{x(t) - x}{t}$$

und erhalten dadurch $\|u(t)\| \leq \alpha v$. Für eine beliebige Folge $\{t_k\},\ t_k \searrow 0$ bekommen wir einen Häufungspunkt u der Folge $\{u(t_k)\}$. Nun sieht man leicht, daß $v \in CF((x, y); u)$ gilt (da $y + tv \in F(x + tu(t))$).

(ii) $\Rightarrow$ (i): Es seien W und α wie in (ii) gegeben. Wir wählen ein $r > 0$ mit $r < \frac{1}{4(\alpha+1)}$. Der Beweis erfolgt indirekt. Angenommen, es existieren Folgen $\{x^k\}$, $\{y^k\}$, $\{z^k\}$, $x^k \to x^0$, $y^k, z^k \to y^0$ mit der Eigenschaft, daß $y^k \in F(x^k)$ und

$$d(x^k, F^{-1}(z^k)) > k\|y^k - z^k\| \qquad \forall k \in I\!\!N. \tag{2.9}$$

Nun sei k so gewählt, daß $k > \frac{1}{r}$ ist. Für dieses k definieren wir ein ε^k als $\varepsilon^k := \|y^k - z^k\|$ und betrachten ein Funktional $f : (I\!\!R^n \times I\!\!R^m) \cap \operatorname{Gph} F \to I\!\!R$, definiert durch

$$f(x, y) := r\|x - x^k\| + \|y - z^k\|.$$

Damit ergibt sich $f(x^k, y^k) = \varepsilon^k$, so daß gilt:

$$f(x^k, y^k) \leq \inf_{(x,y) \in \operatorname{Gph} F} f(x, y) + \varepsilon^k.$$

Wir setzen $\delta^k := \frac{\varepsilon^k}{r}$ und wenden das Ekeland'sche Variationsprinzip auf die Funktion f und den vollständigen metrischen Raum $\operatorname{Gph} F$ an. Dadurch erhalten wir einen Punkt $(x^*, y^*) = (x^*(k), y^*(k)) \in \operatorname{Gph} F$ mit folgenden Eigenschaften:

$$\begin{aligned}
\|x^* - x^k\| + \|y^* - y^k\| &\leq \delta^k && (2.10)\\
f(x^*, y^*) = r\|x^* - x^k\| + \|y^* - z^k\| &\leq f(x^k, y^k) = \varepsilon^k && (2.11)\\
f(x, y) + r(\|x - x^*\| + \|y - y^*\|) &\geq r\|x^* - x^k\| + \|y^* - z^k\| \qquad \forall (x, y) \in \operatorname{Gph} F && (2.12)
\end{aligned}$$

Wir setzen $v := z^k - y^*$. Wäre jetzt $v = 0$, würde (2.10) und $z^k = y^* \in F(x^*)$ implizieren:

$$\frac{\varepsilon^k}{r} = \delta^k \geq \|x^* - x^k\| \geq d(x^k, F^{-1}(z^k)) > k\varepsilon^k,$$

was aber im Widerspruch zur Wahl von k steht. Also ist $v \neq 0$.

OBdA sei $(x^*, y^*) \in W$ (wir können k so groß machen, daß (x^*, y^*) beliebig nahe an (x^0, y^0) rückt, weil $\|x^* - x^0\| + \|y^* - y^0\| \leq \delta^k + \|x^k - x^0\| + \|y^k - y^0\| \to 0$).

Aus (ii) erhalten wir für eine bestimmte Folge $\{t_\ell\}$, $t_\ell \searrow 0$ Punkte $(x(t_\ell), y(t_\ell)) \in \operatorname{Gph} F$ und eine Richtung u mit

$$\begin{aligned} x(t_\ell) &= x^* + t_\ell u + o_x(t_\ell) \\ y(t_\ell) &= y^* + t_\ell v + o_y(t_\ell) \qquad \text{und} \qquad \|u\| \leq \alpha\|v\|. \end{aligned}$$

Daraus folgt:

$$\begin{aligned} \|x(t_\ell) - x^*\| &= \|t_\ell u + o_x(t_\ell)\| \leq 2t_\ell\|u\| \leq 2t_\ell\alpha\|v\| \qquad \text{und} \\ \|y(t_\ell) - z^k\| &= \|y^* + t_\ell v + o_y(t_\ell) - z^k\| \leq (1 - t_\ell)\|v\| + \|o_y(t_\ell)\|. \end{aligned}$$

Zusammen mit (2.12) ergibt das

$$r\|x(t_\ell) - x^k\| - t_\ell\|v\| + \|o_y(t_\ell)\| + r(2\alpha t_\ell\|v\| + t_\ell\|v\| + \|o_y(t_\ell)\|) \geq r\|x^* - x^k\|.$$

Mit Hilfe von

$$r\|x(t_\ell) - x^k\| \leq r\|x(t_\ell) - x^*\| + r\|x^* - x^k\| \leq 2rt_\ell\alpha\|v\| + r\|x^* - x^k\|$$

bekommen wir

$$2rt_\ell\alpha\|v\| - t_\ell\|v\| + \|o_y(t_\ell)\| + r(2\alpha t_\ell\|v\| + t_\ell\|v\| + \|o_y(t_\ell)\|) \geq 0.$$

Division durch $t_\ell\|v\|$ ergibt

$$2r\alpha - 1 + 2r\alpha + r + (1 + r)\frac{\|o_y(t_\ell)\|}{t_\ell\|v\|} \geq 0.$$

Für $t \searrow 0$ ergibt das einen Widerspruch, weil $4r(\alpha + 1) < 1$ angenommen wurde. #

Wenn wir die Forderung nach endlicher Dimension der Räume fallen lassen, erhalten wir nur partielle Aussagen, z.B. reicht für die Richtigkeit der Implikation (i) $\Rightarrow$ (ii), daß der Urbildraum X endlichdimensional ist; die Implikation (ii) $\Rightarrow$ (i) bekommen wir sogar für beliebige Banachräume X, Y.

Die Charakterisierung der Pseudo-Regularität via Satz 2.8 (ii) ergibt folgende Aussage:

Korollar 2.10 *Es sei $F : \mathbb{R}^n \to \mathbb{R}^m$ eine abgeschlossene Multifunktion, $(x^0, y^0) \in \operatorname{Gph} F$. Dann sind folgende Aussagen äquivalent:*

(i) F ist pseudo-Lipschitz in (x^0, y^0)

(ii) es existiert eine Umgebung W von (x^0, y^0) und eine uniforme Konstante $L > 0$, so daß F Lipschitz unterhalb stetig in jedem Punkt $(x, y) \in W \cap \operatorname{Gph} F$ mit der Lipschitz-Konstante L ist.

Beweis: (i) $\Rightarrow$ (ii): siehe Bemerkung 1.4 (ii).
(ii) $\Rightarrow$ (i): Die Multifunktion F sei Lipschitz unterhalb stetig in einem Punkt $(\bar{x}, \bar{y}) \in W \cap \operatorname{Gph} F$, d.h. für x hinreichend nahe an $\bar{x}$ erhalten wir $d(\bar{y}, F(x)) \leq L\|\bar{x} - x\|$. Nun definieren wir Punkte $x(t) := \bar{x} + tv$ mit $\|v\| \leq 1$. Wir bekommen Lösungen $y(t)$ von $y(t) \in F(x(t))$, so daß gilt:

$$\frac{\|y(t) - \bar{y}\|}{t} \leq \frac{L\|x(t) - \bar{x}\|}{t} = L\|v\|$$

also besitzt die Folge $\left\{\frac{y(t) - \bar{y}}{t}\right\}$ einen Häufungspunkt u mit $\|u\| \leq L\|v\|$. Daraus folgt: $v \in CF^{-1}((\bar{y}, \bar{x}); u)$ und F^{-1} ist pseudo-regulär in (y^0, x^0) (nach Satz 2.8). #

2.3 "Arten" der Pseudo-Regularität

Bisher haben wir uns mit theoretischen Aussagen über die Pseudo-Regularität beschäftigt. Nun möchten wir uns einem praktischen Gesichtspunkt dieser Eigenschaft zuwenden. Es sei $F : X \to Y$ eine in dem Punkt (x^0, y^0) pseudo-reguläre Multifunktion. Wie findet man Lösungen $x = x(y)$ von $y \in F(x)$ unter der Voraussetzung, daß y hinreichend nahe an y^0 liegt? Dabei soll x als eine Funktion von y eine der möglicherweise mehreren Lösungen beschreiben. Je nach dem, welche Möglichkeiten wir für die (einfachste) Gestalt von $x(y)$ erhalten, unterscheiden wir zwischen mehreren "Typen "der Pseudo-Regularität.

Wichtig bei diesen Überlegungen ist Folgendes: Obwohl wir die Gestalt der Lösungen in der Form $x = x(y, x^0, y^0)$ schreiben, kann man diesen Ansatz in jedem Punkt $(\bar{x}, \bar{y})$ machen, der sich hinreichend nahe an (x^0, y^0) befindet. Dann bekommen wir zwar eine andere Funktion $x = x(y, \bar{x}, \bar{y})$, jedoch die Gestalt soll die gleiche bleiben. In diesem Fall liefert die Pseudo-Regularität sogar eine uniforme Beschränktheit der Parameter, die in die Formel für $x(y)$ einfließen.

Typ 1:

Es existiert eine Richtung u, so daß

$$x(y) = x^0 + \|y - y^0\|u + o(\|y - y^0\|) \tag{2.13}$$

eine Lösung von $y \in F(x)$ ist. Es liegt also eine (bis auf $o(\|y - y^0\|)$) eindimensionale Variation der Lösungen vor.

Beispiel 2.11 Es sei $F : \mathbb{R}^n \to \mathbb{R}^m$ eine Multifunktion, $K \subset \mathbb{R}^n$ ein konvexer Kegel mit nichtleerem Innern. Weiter erfülle F folgende Monotoniebedingung:

$y \in F(x)$ und $k \in K \Rightarrow y \in F(x + k)$ für alle (x, y) nahe an (x^0, y^0).

Dann ist F pseudo-regulär in $(x^0, y^0) \Leftrightarrow$ es gibt eine Richtung $u \in \operatorname{int} K$ und eine Umgebung W von (x^0, y^0) mit:

$\forall (x, y) \in W \cap \operatorname{Gph} F \quad \forall y'$ nahe an $y : y' \in F(x + \|y' - y\|u)$.

Beweis: $(\Leftarrow)$ ist klar.
$(\Rightarrow)$: Wir nehmen ein $k \in \operatorname{int} K$ und finden ein $\lambda > 0$, so daß $\lambda k + \mathbb{B} \subset K$ ist. Die Abbildung F ist pseudo-regulär in (x^0, y^0), d.h. für alle (x, y) nahe an (x^0, y^0) und für alle y' nahe an y^0 gibt es ein $x' \in F^{-1}(y')$ mit

$$\|x' - x\| \leq L\|y' - y\|$$

für ein geeignetes $L > 0$. Wir definieren $u := L\lambda k \in \operatorname{int} K$ und

$$\ell := \|y' - y\|u + x - x'.$$

Das bedeutet:

$$\ell \in L\|y' - y\|\lambda k + L\|y' - y\|\mathbb{B} = L\|y' - y\|(\lambda k + \mathbb{B}) \subset K$$

und wir erhalten

$$y' \in F(x') \Rightarrow y' \in F(x' + \ell) = F(x + \|y' - y\|u).$$

Beispiel 2.12 Es sei $f : \mathbb{R}^n \to \mathbb{R}^m$ eine Funktion, $K := \operatorname{cl} \mathbb{R}^m_+$ und f erfülle die Bedingung

$$y \leq f(x), \; k \in K \; \Rightarrow y \leq f(x + k).$$

Die Multifunktion $F : \mathbb{R}^n \to \mathbb{R}^m$ sei definiert durch

$$y \in F(x) \; \Leftrightarrow y \leq f(x).$$

Dann ist F pseudo-regulär in $(x^0, y^0) \Leftrightarrow \exists\, \varepsilon > 0 \quad \forall t \in (0, \varepsilon) \quad \forall (x, y) \in \operatorname{Gph} F$ nahe an (x^0, y^0) gilt:

$$y_i + t\varepsilon \leq f_i(x + tu^*) \qquad \forall i = 1, \ldots, m \qquad \text{wobei } u^* = (1, \ldots, 1) \text{ ist.}$$

Beweis: ($\Rightarrow$): Es ist $u^* + I\!B \subset K$ und mit Hilfe der vorherigen Aussage erhalten wir

$$y' \le f(x + \|y' - y\| L u^*).$$

Nun setzt man $y' := y + L^{-1} t v$ mit $\|v\| = 1$. In der $i-$ten Gleichung sei $v := e^i$ (der $i-$te Einheitsvektor), so daß wir für t hinreichend klein bekommen

$$y_i + L^{-1} t \le f_i(x + t u^*) \quad \forall i = 1, \dots, m.$$

Die Definition $\varepsilon := L^{-1}$ liefert die Behauptung.

($\Leftarrow$): Es sei ein Punkt y' gegeben durch $y' = y + t\varepsilon v$, $\|v\| = 1$. Das ergibt:

$$y'_i = y_i + t\varepsilon v_i \le y_i + t\varepsilon \le f_i(x + t u^*) = f_i\left(x + \|y' - y\| \frac{u^*}{\varepsilon}\right) \qquad \forall i = 1, \dots, m. \qquad \#$$

Beispiel 2.13 Wir betrachten die Multifunktion $M : I\!R^m \to I\!R^n$, gegeben durch $M(a) := \{x \in I\!R^n \mid g(x) \le a\}$, wobei $g : I\!R^n \to I\!R^m$ eine stetig differenzierbare Funktion ist, $g(x^0) \le 0$. Wir haben bereits erwähnt, daß unter MFCQ die Abbildung M^{-1} pseudo-regulär in $(x^0, 0)$ ist (siehe [35]). Hierbei handelt es sich ebenfalls um Typ 1.

Beweis: Es sei ein $u \in I\!R^n$ mit $Dg_i(x^0)u < 0$ für $i \in I(x^0) := \{i \in \{1, \dots, m\} \mid g_i(x^0) = 0\}$ und ein $a \in I\!R^m$ nahe am Nullpunkt gegeben, $a \ne 0$. Unser $x(a)$ definieren wir als

$$x(a) := x^0 + \|a\| u.$$

Wir sehen:

$$g_i(x(a)) = g_i(x^0) + Dg_i(x^0)(x - x^0) + o(\|x - x^0\|) = Dg_i(x^0)u\|a\| + o(\|a\|) \quad \forall i \in I(x^0)$$

und erhalten dadurch

$$\frac{g_i(x(a))}{\|a\|} = Dg_i(x^0)u + \frac{o(\|a\|)}{\|a\|} < 0$$

für a nahe am Nullpunkt und $i \in I(x^0)$.
Für $i \notin I(x^0)$ (d.h. $g_i(x^0) < 0$) können wir gegebenenfalls $\|a\|$ verkleinern, so daß $g_i(x) < 0$ erfüllt bleibt. #

Typ 2: "Lineare" Abhängigkeit

es existiert eine lineare Abbildung $A : Y \to X$, so daß

$$x(y) = x^0 + A(y - y^0) + o(\|y - y^0\|) \tag{2.14}$$

eine Lösung der Inklusion $y \in F(x)$ beschreibt. Als einen Vertreter dieses Typs können wir die Situation erwähnen, in der die Voraussetzungen des klassischen Satzes über die implizite Funktion erfüllt sind:

Es sei $f : I\!R^n \to I\!R^m$ eine stetig differenzierbare Funktion, $f(x^0) = y^0$, die Matrix $Df(x^0)$ habe vollen Rang. Nach Einfügen zusätzlicher Gleichungen, die das System zu einer regulären $(I\!R^n, I\!R^n)$-Gleichung machen, beschrieben als $F(x) = z$, folgt die Behauptung aus

$$x(z) = x^0 + [DF(x^0)]^{-1}(z - z^0) + o(\|z - z^0\|).$$

Jetzt geben wir ein Beispiel an, wo eine Mischung von Typ 1 und Typ 2 vorliegt:

Beispiel 2.14 Wir betrachten eine Multifunktion $M : I\!R^m \times I\!R^k \to I\!R^n$ der Form $M(a,b) := \{x \in I\!R^n \,|\, g(x) \leq a,\ h(x) = b\}$, wobei $g \in C^1(I\!R^n, I\!R^m)$, $h \in C^1(I\!R^n, I\!R^k)$, $x^0 \in M(0,0)$. Wir wissen, daß unter MFCQ die Abbildung M^{-1} pseudo-regulär in $(x^0, (0,0))$ ist.

Behauptung: Es existieren eine Matrix A und eine Richtung u, so daß (eine Inverse) $x(a,b)$ folgende Gestalt hat

$$x(a,b) = x^0 + Ab + (\|a\| + \|b\|)u + o(\|a\| + \|b\|). \tag{2.15}$$

Beweis: Durch $I(x^0) := \{i \in \{1,\ldots,m\} \,|\, g_i(x^0) = 0\}$ definieren wir die aktive Indexmenge. Aus MFCQ folgt, daß $k \leq n$ sein muß. Im Falle der Gleichheit definieren wir die Matrix $D := Dh(x^0)$. Falls $k < n$ gilt, finden wir noch $n - k$ Vektoren $d_{k+1}, \ldots, d_n$ aus $I\!R^n$, so daß die Matrix

$$D := \begin{pmatrix} Dh_i(x^0) \\ d_j \end{pmatrix} \begin{array}{l} \\ i = 1,\ldots,k \\ j = k+1,\ldots,n \end{array}$$

regulär ist und D^{-1} existiert. Jetzt streichen wir in D^{-1} die Spalten $k+1,\ldots,n$ und bezeichnen die auf diese Weise entstandene Matrix mit A. Nun betrachten wir folgende Funktion: $F : I\!R \times I\!R^k \times I\!R^n \to I\!R^n$

$$\begin{array}{lcll} F_i(t,b,r) & := & h_i(x^0 + Ab + tu + r) - b_i & i = 1,\ldots,k \\ F_j(t,b,r) & := & d_j^T(Ab + r) & j = k+1,\ldots,n \end{array}$$

wobei u die MFCQ-Richtung sei. Zusätzlich erfülle u die folgende Bedingung:

$$\frac{1}{2}\|Dg_i(x^0)u\| > \|A\|\|Dg_i(x^0)\| \qquad \forall i \in I(x^0)$$

(wenn u sie nicht erfüllt, finde ein $\lambda > 0$, so daß λu es tut).

Wegen $F(0,0,0) = 0$ und $D_rF(0,0,0) = D$ kann der Satz über implizite Funktionen angewendet werden und ergibt: Es existiert eine lokal eindeutige Funktion $r(t,b)$ mit $F(t,b,r(t,b)) = 0$. Außerdem erhält man

$$Dr(0,0) = -D^{-1}(D_tF(0,0,0), D_bF(0,0,0)).$$

Wenn wir noch in Betracht ziehen, daß

$$D_t F(0,0,0) = \begin{pmatrix} Dh_i(x^0)u \\ 0 \end{pmatrix} = 0$$

und

$$D^{-1} D_b F(0,0,0) = D^{-1} \left(DA - \begin{pmatrix} E_k \\ 0 \end{pmatrix} \right) = A - D^{-1} \begin{pmatrix} E_k \\ 0 \end{pmatrix} = A - A = 0,$$

(wobei E_k die k-dimensionale Einheitsmatrix bezeichnet), so bekommen wir $Dr(0,0) = 0$, d.h. $r(t,b) = o(t + \|b\|)$, falls die betrachtete Norm die Maximumnorm ist.

Setzt man nun $t := \|a\| + \|b\|$, erhält man Punkte

$$x(a,b) := x^0 + Ab + (\|a\| + \|b\|)u + r(\|a\| + \|b\|, b)$$

und eine Gestalt wie in (2.15). Daraus folgt:
$h(x(a,b)) = b$ und
$g_i(x(a,b)) = Dg_i(x^0)[Ab + (\|a\| + \|b\|)u + o(\|a\| + \|b\|)] + o(\|x(a,b) - x^0\|), \quad i \in I(x^0).$

Mit Hilfe der Gestalt von $x(a,b)$ ergibt das:

$$g_i(x(a,b)) = Dg_i(x^0)[Ab + (\|a\| + \|b\|)u] + o(\|a\| + \|b\|).$$

Wir teilen durch $\|a\| + \|b\|$ und erhalten:

$$\frac{g_i(x(a,b))}{\|a\| + \|b\|} = Dg_i(x^0)A\frac{b}{\|a\| + \|b\|} + Dg_i(x^0)u + \frac{o(\|a\| + \|b\|)}{\|a\| + \|b\|}.$$

Nun ist aber

$$\left\| Dg_i(x^0)A\frac{b}{\|a\| + \|b\|} \right\| \leq \|Dg_i(x^0)\| \|A\| < \frac{1}{2}\|Dg_i(x^0)u\|$$

und damit folgt $g_i(x(a,b)) \leq 0$ für (a,b) nahe Null und $i \in I(x^0)$.

Für Indizes $i \notin I(x^0)$ ergibt sich $g_i(x(a,b)) \leq 0$ aus Stetigkeitsgründen. Deshalb liegt $x(a,b)$ in $M(a,b)$ und hat die verlangte Form (2.15). #

Typ3: "Stückweise lineare " Abhängigkeit

es existieren endlich viele lineare Abbildungen $A_1, \ldots, A_k : Y \to X$, so daß

$x(y)$ eine (nicht notwendigerweise stetige) Auswahlfunktion der Funktionen

$$x_i(y) := x^0 + A_i(y - y^0) + o_i(\|y - y^0\|) \tag{2.16}$$

ist und $x(y) \in F^{-1}(y)$ gilt. Im dritten Abschnitt werden wir zeigen: Ist $f : I\!\!R^n \to I\!\!R^n$ eine stückweise glatte pseudo-reguläre Funktion, erhalten wir eine Gestalt vom Typ 3.

Typ 4

es existieren endlich viele positiv homogene (aber nicht notwendigerweise stetige) Abbildungen $A_1, \ldots, A_k : Y \to X$, so daß

(eine Inverse) $x(y)$ eine (nicht notwendigerweise stetige) Auswahlfunktion der Funktionen

$$x_i(y) := x^0 + A_i(y - y^0) + o_i(\|y - y^0\|) \tag{2.17}$$

ist.

Im Abschnitt 3 werden wir zeigen, daß eine große Klasse von lokal Lipschitz-stetigen Funktionen $f : I\!\!R^n \to I\!\!R^n$, die pseudo-regulär sind, diese Eigenschaft besitzen.

Zuletzt möchten wir ein Beispiel von B. Kummer angeben, das in keine der bis jetzt beschriebenen Gruppen einzuordnen ist. Gleichzeitig illustriert es ein kompliziertes Verhalten, das durch die unendliche Dimension des Urbildraumes zustande kommt.

Beispiel 2.15 Es sei $f : \ell_2 \to I\!\!R$ eine Funktion, definiert durch $f(x) = \sup x_k$, wobei ℓ_2 den Hilbert'schen Folgenraum bezeichnet. Unser x^0 sei als $x^0_k = -\frac{1}{k}$ gegeben, und es sei $f(x^0) = 0$. Wir definieren die zu untersuchende Multifunktion $F : \ell_2 \to I\!\!R$ auf folgende Weise: $y \in F(x) \Leftrightarrow y \leq f(x)$. Die Abbildung F ist pseudo-regulär in $(x^0, 0)$, was man wie folgt sieht:
Seien $s, t \in I\!\!R$ hinreichend nahe am Nullpunkt und $x(t) \in F^{-1}(t)$. Ist $s \leq t$, so können wir $x(s) = x(t)$ setzen und erfüllen alle Bedingungen.
Andernfalls nutzt man Folgendes:
Für jedes $\varepsilon > 0$ muß es ein $k = k(\varepsilon)$ geben mit $x_{k(\varepsilon)}(t) \geq t - \varepsilon$. Wir wählen jetzt ein spezielles $\varepsilon < s - t$ und finden das entsprechende $k(\varepsilon)$. Unser $x(s)$ können wir dann definieren als:

$$\begin{aligned} x_{k(\varepsilon)}(s) &:= x_{k(\varepsilon)}(t) + 2(s - t) \\ x_k(s) &:= x_k(t) \quad \forall k \neq k(\varepsilon). \end{aligned}$$

Man sieht:

$$x_{k(\varepsilon)}(s) \geq t - \varepsilon + 2(s - t) = s + (s - t - \varepsilon) \geq s,$$

also ist tatsächlich $f(x(s)) \geq s$ und $x(s) \in F^{-1}(s)$. Außerdem erhält man
$\|x(t) - x(s)\| \leq 2|s - t|$.

Als nächstes behaupten wir, daß es keine feste Richtung u gibt mit $x(t) = x^0 + tu + o(t)$, $x(t) \in F^{-1}(t)$. Angenommen, so eine Richtung existiert; dann müßte gelten:

$$\|tu + o(t)\| = \|x(t) - x^0\| \leq Lt$$

für eine geeignete Konstante $L > 0$. Durch den Grenzübergang $t \to 0$ liefert dies $\|u\| \leq L$.

Jetzt sei $t > 0$ gegeben. Für $x(t)$ soll $f(x(t)) \geq t$ gelten, also gibt es ein $k = k(t)$ mit $x_{k(t)}(t) > \frac{t}{2}$. Das würde ergeben:

$$\begin{aligned} -\frac{1}{k(t)} + tu_{k(t)} + o(t) &> \frac{t}{2} \qquad \text{und} \\ L \geq u_{k(t)} &> \frac{1}{2} - \frac{o(t)}{t} + \frac{1}{tk(t)}. \end{aligned}$$

Wir erhalten $u_{k(t)} \geq \frac{1}{3}$ (für t hinreichend klein) einerseits und $k(t) \to \infty$ für $t \to 0$ andererseits. In anderen Worten, je mehr sich t der Null nähert, um so mehr Komponenten $k(t)$ von u müssen größer als $\frac{1}{3}$ sein und es folgt $\|u\| = \infty$. Eine weitere Analyse des Beispiels zeigt, daß dessen Pseudo-Regularität weder mittels Contingent-Ableitungen noch mittels Co-Ableitungen erkannt werden kann. #

3 Lipschitz-Funktionen $f : I\!R^n \to I\!R^n$

Im ersten Unterabschnitt untersuchen wir einige nützliche Eigenschaften von verschiedenen Arten von Richtungsableitungen (B-Ableitungen, Contingent-Ableitungen, Thibault-Mengen). Von besonderem Interesse für uns sind Kettenregeln und Formeln für partielle Ableitungen zu diesen Ableitungskonzepten.

Der wichtigste Teil dieses Kapitels ist den Untersuchungen der topologischen Struktur der Urbildmengen $f^{-1}(y)$ gewidmet, wobei f eine lokale Lipschitz-Funktion ist. Insbesondere sind wir an Aussagen über die Isoliertheit der Urbilder interessiert. Zu diesem Zweck untersuchen wir die Regularitätseigenschaften derartiger Funktionen. Als ein wichtiger Spezialfall werden stückweise glatte Funktionen behandelt.

Unter zusätzlichen Bedingungen an die Richtungsdifferenzierbarkeit von f und $f'(x,.)$ können wir zeigen, daß die Pseudo-Regularität von f zur Gestalt vom Typ 4 führt. Als ein wichtiges Hilfsmittel wird sich eine Art Dimensionslemma erweisen. Es schließt für Lipschitz-Funktionen aus, daß die Dimension der Bildmenge höher als die der Urbildmenge ist.

3.1 Kettenregeln für Richtungsableitungen

Im Folgenden erwähnen wir ein paar nützliche Eigenschaften von (verallgemeinerten) Richtungsableitungen, die wir für unsere späteren Untersuchungen brauchen werden. Das nächste Lemma erhalten wir sofort aus der Definition der Contingent-Ableitung für Funktionen:

Lemma 3.1 *Ist $g : I\!R^m \to I\!R^n$ eine positiv homogene Funktion, so gilt:*

$$Cg(\alpha u, w) = Cg(u, w) \quad \forall \alpha > 0.$$

Nun untersuchen wir die Frage, inwiefern sich die Differenzierbarkeitseigenschaften einer Funktion auf ihre Richtungsableitungen übertragen lassen. Dazu sei eine lokale Lipschitz-Funktion f gegeben, die richtungsdifferenzierbar ist. Unter welchen Bedingungen ist dann die Richtungsableitung $f'(x,.)$ selber richtungsdifferenzierbar?

Das folgende Beispiel zeigt, daß dies im allgemeinen nicht der Fall ist.

Beispiel 3.2 Zunächst definieren wir eine Funktion $h : [0,1] \to [0,1]$ durch

$$h(x) := \begin{cases} \frac{1}{2}x + \frac{1}{2^n} & \text{für} \quad x \in \left[\frac{3}{2^{n+1}}, \frac{1}{2^{n-1}}\right], \; n \in I\!N \\ \frac{3}{2}x - \frac{1}{2^{n+1}} & \text{für} \quad x \in \left[\frac{1}{2^n}, \frac{3}{2^{n+1}}\right], \; n \in I\!N \\ 0 & \text{für} \quad x = 0. \end{cases}$$

Für $x \in [-1, 0]$ setzen wir $h(-x) := h(x)$. Wir sehen: h ist im Nullpunkt nicht richtungsdifferenzierbar. Man kann leicht die Contingent-Ableitung von h berechnen: $Ch(0,1) = [1, \frac{7}{6}]$.

Nun definieren wir eine Funktion $g(x,y)$ auf der Menge $A := \{(x,y) \in I\!R^2 \mid |y| \leq x\}$ wie folgt:

$$g(1,y) := h(y)$$

und

$$g(x,y) := \begin{cases} xh\left(1, \frac{y}{x}\right) & \text{falls} \quad x \neq 0 \\ 0 & \text{falls} \quad x = 0. \end{cases}$$

Ähnlich wie die Funktion h, ist g auch nicht richtungsdifferenzierbar in allen Punkten $(x, 0)$ mit $x > 0$. Mit Hilfe von g erklären wir jetzt die Funktion f zunächst auf der Menge A:

$$f(x,y) := \begin{cases} 2y & \text{für} \quad (x,y) \in A, \; -x^2 \leq y \leq x^2 \\ 2x^2 + \left(1 - \frac{x^2}{y}\right) g(x,y) & \text{für} \quad (x,y) \in A, \; y \geq x^2 \text{ oder } y \leq -x^2. \end{cases}$$

Diese Funktion setzen wir auf den ganzen Raum $I\!R^2$ durch die Vorschriften $f(-x,y) := f(x,y)$ und $f(y,x) := f(x,y)$ fort.

Wir untersuchen f auf Richtungsdifferenzierbarkeit. Probleme könnte es nur im Nullpunkt geben, aber dort gilt (für eine Richtung $u = (u_x, u_y)$ mit $\|u\| = 1$ und $u \neq (1,0)$):

$$\lim_{t \searrow 0} \frac{1}{t} f(tu) = \lim_{t \searrow 0} \frac{1}{t} \left[2t^2 u_x^2 + \left(1 - \frac{t^2 u_x^2}{t u_y}\right) t g(u_x, u_y) \right] = g(u_x, u_y) = g(u)$$

und $f'((0,0);(1,0)) = 0$. D.h., f ist richtungsdifferenzierbar, $f'(0,u) = g(u)$ und die Richtungsableitung ist selber nicht richtungsdifferenzierbar. #

Wir erhalten aufgrund des Beispiels eine positive Aussage nur unter einer Zusatzbedingung. Die nachstehende ist aber ziemlich einschränkend. Es handelt sich hierbei um eine hinreichende, aber nicht notwendige Bedingung.

Definition 3.3 *Es sei $f : I\!R^n \to I\!R^n$ lokal Lipschitz und richtungsdifferenzierbar. Wir sagen, f ist* **streng B-differenzierbar** *in x^0, falls es für jedes $\varepsilon > 0$ eine Umgebung U vom Nullpunkt gibt, so daß die Funktion*

$$r_{x^0}(u) := f(x^0 + u) - f(x^0) - f'(x^0, u)$$

auf U Lipschitz mit der Konstante ε ist.

Bemerkung: Der Buchstabe B bezieht sich auf den Namen Bouligand-Ableitungen, (eine andere Bezeichnung für Richtungsableitungen) der von Robinson in [37] eingeführt wurde. Die strenge B-differenzierbatkeit definierte er in [39], um eine "strenge" Approximation der zugrundeliegenden Funktion zu erhalten.

Lemma 3.4 *Es sei $f : I\!R^n \to I\!R^n$ lokal Lipschitz und richtungsdifferenzierbar. Weiterhin sei f streng B-differenzierbar in x^0. Dann ist auch die Funktion $g(.) := f'(x^0, .)$ richtungsdifferenzierbar.*

Beweis: Der Einfachheit halber schreiben wir im Folgenden r anstatt r_{x^0}. Es sei ein $v \in Cg(u, w)$ gegeben, d.h. es existiert eine Folge $\{t_k\}$, $t_k \searrow 0$ mit

$$v = \lim_{t_k \searrow 0} \frac{1}{t_k} \left[g(u + t_k w) - g(u) \right].$$

Den Term in der eckigen Klammer können wir schreiben als

$$f(x^0 + u + t_k w) - r(u + t_k w) - f(x^0 + u) + r(u).$$

Das heißt, es existiert der Limes

$$v' = \lim_{t_k \searrow 0} \frac{1}{t_k} \left[r(u + t_k w) - r(u) \right]$$

und es ist $v = f'(x^0 + u, w) - v'$. Wir erhalten dadurch

$$Cg(u, w) = f'(x^0 + u, w) - Cr(u, w) = \{ f'(x^0 + u, w) - v' \mid v' \in Cr(u, w) \}.$$

Insbesondere gilt $\operatorname{diam} Cg(u, w) = \operatorname{diam} Cr(u, w)\ \forall u\ \forall w$. Angenommen, $Cg(u, w)$ ist nicht einelementig, d.h. $\operatorname{diam} Cg(u, w) > 0$. (OBdA sei $\|w\| = 1$.) Wir definieren $\varepsilon := \frac{1}{3} \operatorname{diam} Cg(u, w)$. Für dieses ε bekommen wir eine Umgebung $U(\varepsilon)$ des Nullpunktes, so daß r auf $U(\varepsilon)$ Lipschitz mit der Konstante ε ist. Man kann daher ein $\alpha > 0$ mit $\alpha u \in \operatorname{int} U(\varepsilon)$ finden.

Nun betrachten wir die Menge $Cg(\alpha u, \alpha w)$. Die Funktion g ist positiv homogen, deshalb gilt:

$$\operatorname{diam} Cg(\alpha u, \alpha w) = \operatorname{diam} \alpha Cg(u, w) = \alpha \operatorname{diam} Cg(u, w).$$

Weiterhin nehmen wir uns ein beliebiges $v' \in Cr(\alpha u, \alpha w)$ her. Für hinreichend große Indizes k gilt $\alpha u + t_k \alpha w \in U(\varepsilon)$ und wir erhalten $\|v'\| \leq \varepsilon \alpha \|w\|$.

Wir dürfen daher schreiben:

$$\begin{aligned} \alpha \operatorname{diam} Cg(u, w) &= \operatorname{diam} Cg(\alpha u, \alpha w) = \operatorname{diam} Cr(\alpha u, \alpha w) \leq 2\varepsilon\alpha\|w\| = \\ &= \frac{2}{3} \alpha \operatorname{diam} Cg(u, w). \end{aligned}$$

Das bedeutet aber: diam $Cg(u,w) =$ diam $Cr(u,w) = 0$, und g ist richtungsdifferenzierbar. #

In [23] hat Kummer einige Kettenregeln für Thibault-Mengen bewiesen. Die Ergebnisse sind in folgender Aussage zusammengefaßt:

Proposition 3.5 *Es seien $f : \mathbb{R}^m \to \mathbb{R}^p$, $g : \mathbb{R}^n \to \mathbb{R}^m$ lokal Lipschitz und $h := f(g(.))$. Dann gilt:*

(i)
$$Th(x,u) \subset \bigcup_{v \in Tg(x,u)} Tf(g(x),v)$$
$$Ch(x,u) \subset \bigcup_{v \in Cg(x,u)} Cf(g(x),v)$$

(ii) Ist f oder g richtungsdifferenzierbar, gilt
$$Ch(x,u) = \bigcup_{v \in Cg(x,u)} Cf(g(x),v)$$

(iii) Ist f stetig differenzierbar, gilt
$$Th(x,u) = \bigcup_{v \in Tg(x,u)} Df(g(x))v$$

(iv) Ist g stetig differenzierbar und g^{-1} unterhalb stetig in $(g(x),x)$, gilt
$$Th(x,u) = Tf(g(x), Dg(x)u)$$

Beweis: (i): Es sei $w \in Th(x,u)$, d.h. $w = \lim w^k$,
$$w^k = \frac{1}{t_k}\left[f(g(x^k + t_k u)) - f(g(x^k))\right]$$
für geeignete Folgen $t_k \searrow 0$ und $x^k \to x$. Die Folge
$$v^k := \frac{1}{t_k}\left[g(x^k + t_k u) - g(x^k)\right] \tag{3.1}$$
ist beschränkt (da g Lipschitz ist) und wir können annehmen, daß v^k gegen ein $v \in Tg(x,u)$ konvergiert. Mit (3.1) haben wir
$$w^k = \frac{1}{t_k}\left[f(g(x^k) + t_k v^k) - f(g(x^k))\right]$$

und $w \in Tf(g(x), v)$
Wenn wir $x^k := x$ festhalten, bleiben unsere Argumente richtig und liefern die Behauptung für Contingent-Ableitungen.

(ii) Die Funktion g sei richtungsdifferenzierbar, und es sei ein $w \in Cf(g(x), v)$ gegeben, d.h. $g'(x, u) = v$ und

$$w = \lim_{t_k \searrow 0} \frac{1}{t_k} [f(g(x) + t_k v) - f(g(x))]$$

für eine geeignete Folge $\{t_k\}$, $t_k \searrow 0$. Wir haben dann

$$g(x + t_k u) = g(x) + t_k v + o(t_k)$$

und für w ergibt das

$$w = \lim_{t_k \searrow 0} \frac{1}{t_k} [f(g(x + t_k u) - o(t_k)) - f(g(x))] \in Ch(x, u)$$

(wegen Bemerkung 2.3(iii)).
Nun sei f richtungsdifferenzierbar, $v \in Cg(x, u)$ und $w = f'(g(x), v)$. Für eine bestimmte Folge $\{t_k\}$, $t_k \searrow 0$ erhält man somit

$$g(x + t_k u) = g(x) + t_k v + o(t_k). \tag{3.2}$$

Schreiben wir weiter w mit Hilfe von (3.2) um, erhalten wir analog zu oben $w \in Ch(x, u)$.

(iii) Nun sei $w = Df(g(x))v$ und $v \in Tg(x, u)$. Dann gilt $v = \lim v^k$,

$$v^k = \frac{1}{t_k} \left[g(x^k + t_k u) - g(x^k) \right] \tag{3.3}$$

für bestimmte Folgen $x^k \to x$ und $t_k \searrow 0$. Die Funktion f ist stetig differenzierbar. Zusammen mit (3.3) ergibt das

$$w = \lim_{t_k \searrow 0} \frac{1}{t_k} \left[f(g(x^k) + t_k v) - f(g(x^k)) \right] = \lim_{t_k \searrow 0} \frac{1}{t_k} \left[f(g(x^k + t_k u)) - f(g(x^k)) \right] \in Th(x, u).$$

(iv) Es sei $v = Dg(x)u$ und $w \in Tf(g(x), v)$, d.h.

$$w = \lim_{t_k \searrow 0} \frac{1}{t_k} \left[f(y^k + t_k v) - f(y^k) \right]$$

für $t_k \searrow 0$ und $y^k \to g(x)$. Wegen der Unterhalbstetigkeit von g^{-1} in $(g(x), x)$ finden wir eine Folge $\{x^k\}$, $x^k \to x$ und $y^k = g(x^k)$. Das ergibt

$$w = \lim_{t_k \searrow 0} \frac{1}{t_k} \left[f(g(x^k) + t_k v) - f(g(x^k)) \right].$$

Wegen $v = Dg(x)u$ können wir schreiben

$$g(x^k) + t_k v = g(x^k + t_k u) + o(t_k)$$

und erhalten $w \in Th(x,u)$. #

Aus dem Korollar 2.10 folgt, daß die Pseudo-Regularität einer Funktion die Unterhalbstetigkeit ihrer Inversen zur Folge hat. Dies ist interessant im Hinblick auf die Aussage (iv) aus Proposition 3.5.

Nun untersuchen wir das Problem partieller Ableitungen: Es sei $f : I\!R^n \times I\!R^m \to I\!R^p$ eine lokale Lipschitz-Funktion, $f = f(x,y)$. Wir sind interessiert an Formeln folgender Gestalt für partielle Contingent-Ableitungen und Thibault-Mengen:

$$\begin{aligned} Tf((x,y);(u,v)) &= T_x f((x,y);u) + T_y f((x,y);v) \qquad (3.4)\\ Cf((x,y);(u,v)) &= C_x f((x,y);u) + C_y f((x,y);v) \qquad (3.5) \end{aligned}$$

wobei T_x, T_y bzw. C_x, C_y die partiellen Richtungsableitungen bezeichnen sollen, z.B.: $T_x f((x,y);u) = Tf(.,y)(x,u)$. Diese Formeln sind im allgemeinen nicht richtig (siehe folgendes Beispiel):

Beispiel 3.6 Die Funktion $f : I\!R^2 \to I\!R$ sei gegeben durch

$$f(x,y) := \begin{cases} 0 & \text{für} \quad x \le 0 \\ x & \text{für} \quad 0 \le x \le |y| \\ |y| & \text{für} \quad |y| \le x \end{cases}$$

Wie man leicht sieht, ist $Cf((0,0);(1,1)) = \{\frac{1}{\sqrt{2}}\} \subset Tf((0,0);(1,1))$, andererseits ist $T_x f((0,0);1) = T_y f((0,0);1) = C_x f((0,0);1) = C_y f((0,0);1) = \{0\}$. #

In [23] findet man eine hinreichende Bedingung für die Gültigkeit obengenannter Formeln. Auf analoge Art und Weise kann man eine Aussage auch für Contingent-Ableitungen erzielen. Als Ergebnis erhalten wir ein zu [23] ähnliches, aber schärferes Lemma:

Lemma 3.7 *Es sei $f : I\!R^n \times I\!R^m \to I\!R^p$, $f = f(x,y)$ lokal Lipschitz und stetig differenzierbar bezüglich y.*

(i) $D_y f(.,\theta)$ sei lokal uniform Lipschitz für θ nahe an y^0 und x nahe an x^0, d.h. es existiert eine Konstante $L > 0$ mit: $\forall x_1, x_2 \in x^0 + L^{-1} I\!B\ \forall \theta \in y^0 + L^{-1} I\!B$:

$\|D_y f(x_1,\theta) - D_y f(x_2,\theta)\| \le L\|x_1 - x_2\|$.

Dann gilt:

$$Tf((x^0,y^0);(u,v)) = T_x f((x^0,y^0);u) + D_y f(x^0,y^0)v.$$

(ii) $D_y f(.,\theta)$ *sei gleichmäßig stetig für* θ *nahe an* y^0 *und* x *nahe an* x^0*, d.h. es existiert eine Konstante* $L > 0$ *mit:* $\forall x_1, x_2 \in x^0 + L^{-1}\mathbb{B}\ \forall \theta \in y^0 + L^{-1}\mathbb{B}:$

$\forall \varepsilon > 0\, \exists \delta > 0: \quad \|x_1 - x_2\| \leq \delta \Rightarrow \|D_y f(x_1,\theta) - D_y f(x_2,\theta)\| \leq \varepsilon$.

Dann gilt:

$$Cf((x^0,y^0);(u,v)) = C_x f((x^0,y^0);u) + D_y f(x^0,y^0)v.$$

Beweis: (i): Die Inklusion $(\subset)$. Wir sehen uns die Formel näher an. Die linke Seite besteht aus allen Limita von

$$a^k = \frac{1}{t_k}\left[f(x^k + t_k u, y^k + t_k v) - f(x^k, y^k)\right], \qquad (x^k, y^k) \to (x^0, y^0)$$

und die rechte Seite aus allen Limita von

$$b^k = \frac{1}{\alpha_k}\left[f(z^k + \alpha_k u, y^0) - f(z^k, y^0)\right] + \frac{1}{\beta_k}\left[f(x^0, w^k + \beta_k v) - f(x^0, w^k)\right],$$

wobei $(z^k, w^k) \to (x^0, y^0)$. Jetzt sei die Folge $\{a^k\}$ gegeben, $a^k \to a$. Unser Ziel ist es, die Folge $\{b^k\}$ zu finden, so daß $b^k \to a$. Dazu setzt man $\alpha_k = \beta_k = t_k$, $z^k = x^k$, $w^k = y^k$ und rechnet die Differenz $c^k = a^k - b^k$ aus.

Um die Schreibweise abzukürzen, setzen wir $\xi^k := x^k + t_k u$, $\eta^k := y^k + t_k v$, und erhalten:

$$\begin{aligned} c^k &= \frac{1}{t_k}\left\{\left[f(\xi^k,\eta^k) - f(\xi^k, y^0)\right] - \left[f(x^k,\eta^k) - f(x^k,y^0)\right]\right\} + \\ &+ \frac{1}{t_k}\left\{\left[f(x^k,\eta^k) - f(x^k,y^k)\right] - \left[f(x^0,\eta^k) - f(x^0,y^k)\right]\right\} = \frac{c_1^k}{t_k} + \frac{c_2^k}{t_k} \end{aligned}$$

Den Term in der ersten geschweiften Klammer kann man schreiben als

$$c_1^k = \int_0^1 \left[D_y f(\xi^k,\theta) - D_y f(x^k,\theta)\right](\eta^k - y^0)d\tau$$

wobei $\theta = y^0 + \tau(\eta^k - y^0)$. Wir erhalten:

$$\|c_1^k\| \leq L\|\xi^k - x^k\|\|\eta^k - y^0\| = Lt_k\|u\|\|\eta^k - y^0\|$$

und $\frac{c_1^k}{t_k} \to 0$, da η^k gegen y^0 konvergiert. Analoge Abschätzung bekommen wir für c_2^k.

(ii): Die Inklusion $(\subset)$. Hier ist die Situation einfacher, da $x^k = x^0$ und $y^k = y^0$ gilt. Mit denselben Substitutionen wie vorher ergibt das:

$$\|c_1^k\| \le \varepsilon \|\eta^k - y^0\| = \varepsilon t_k \|v\|$$

für $t_k\|u\| \le \delta$. Analog die zweite Abschätzung.

(i) und (ii): Die Inklusionen $(\supset)$: Jetzt sei die Folge $\{b^k\}$ gegeben,

$$b^k = \frac{1}{\alpha_k}\Big[f(z^k + \alpha_k u, y^0) - f(z^k, y^0)\Big] + \frac{1}{\beta_k}\Big[f(x^0, w^k + \beta_k v) - f(x^0, w^k)\Big], \tag{3.6}$$

wobei $(z^k, w^k) \to (x^0, y^0)$.

Wir können versuchen, den zweiten Term auf der rechten Seite in (3.6) zu ersetzen durch

$$p^k = \frac{1}{\alpha_k}\Big[f(x^0, \mu^k + \alpha_k v) - f(x^0, \mu^k)\Big],\ \eta^k \to y^0.$$

Da die Funktion f in der zweiten Variable stetig differenzierbar ist, können wir das tun und es ist $\|p^k - b_2^k\| \to 0$. Wir setzen $t_k = \alpha_k$, $x^k = z^k$, $y^k = \mu^k$ und c^k konvergiert gegen Null wie früher. Bei den Contingent-Ableitungen erfolgt der Beweis analog, nur das Argument für die Konvergenz von c^k ist jetzt anders. #

Die Forderung, die Funktion f sei in der zweiten Variable stetig differenzierbar, ist sehr stark. Wir können sie auf folgende Art und Weise umgehen: Wir definieren eine neue Funktion $F(x, z) := f(x, g(z))$ (wobei f stetig differenzierbar in der zweiten Variable ist). Für bestimmte Abbildungen g können die Formeln nun auch auf F übertragen werden (dabei ist F nicht differenzierbar in z).

Definition 3.8 *Eine lokale Lipschitz-Funktion $g : \mathbb{R}^k \to \mathbb{R}^m$ nennen wir* **simpel** *in z^0, wenn für jede Folge $\{t_k\}$, $t_k \searrow 0$ und für jedes Paar (v, w) mit $v \in Tg(z^0, w)$ eine Folge $\{z^k\}$, $z^k \to z^0$ existiert, so daß gilt:*

v ist ein Häufungspunkt der Folge $\left\{\frac{1}{t_k}\Big[g(z^k + t_k w) - g(z^k)\Big]\right\}$.

Lemma 3.9 *Es sei $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$, $f = f(x, y)$ lokal Lipschitz und stetig differenzierbar bezüglich y.*

(i) Zusätzlich erfülle f die Voraussetzungen von Lemma 3.7 (i) und die Funktion $g : \mathbb{R}^k \to \mathbb{R}^m$ sei simpel in $z^0 \in g^{-1}(y^0)$.

Dann gilt für die Funktion $F(x, z) := f(x, g(z))$ folgende Formel:

$$TF((x^0, z^0); (u, w)) = T_x f((x^0, g(z^0)); u) + \bigcup_{v \in Tg(z^0, w)} D_y f(x^0, g(z^0))v.$$

(ii) *Zusätzlich erfülle* f *die Voraussetzungen von Lemma 3.7 (ii) und die Funktion* $g : \mathbb{R}^k \to \mathbb{R}^m$ *sei richtungsdifferenzierbar in* $z^0 \in g^{-1}(y^0)$.

Dann gilt für die Funktion $F(x,z) := f(x, g(z))$ *folgende Formel:*

$$CF((x^0, z^0); (u, w)) = C_x f((x^0, g(z^0)); u) + D_y f(x^0, g(z^0)) g'(z^0, w).$$

Beweis: (i) Proposition 3.5 (i) hat folgende Inklusion zur Folge:

$$TF((x^0, z^0); (u, w)) \subset T_x f((x^0, g(z^0)); u) + \bigcup_{v \in Tg(z^0, w)} D_y f(x^0, g(z^0)) v.$$

Nun sei ein b gegeben, $b = b_1 + b_2$, $b_1 \in T_x f((x^0, g(z^0)); u)$, $b_2 = D_y f(x^0, g(z^0)) v$ und $v \in Tg(z^0, w)$. Demnach können wir b_1 schreiben als den Limes der Folge

$$\frac{1}{t_k} \left[f(x^k + t_k u, g(z^0)) - f(x^k, g(z^0)) \right]$$

für bestimmte Folgen $t_k \searrow 0$ und $x^k \to x^0$. Die Funktion g ist simpel, deshalb können wir annehmen, daß v ein Häufungspunkt der Folge

$$\left\{ \frac{1}{t_k} \left[g(z^k + t_k w) - g(z^k) \right] \right\}$$

für eine geeignete Folge $z^k \to z^0$ ist. Das ergibt:

$$g(z^k + t_k w) = g(z^k) + t_k v + o(t_k). \tag{3.7}$$

Nun können wir b_2 mit Hilfe der Folge $\{t_k\}$ umschreiben: Wir setzen

$$b_2^k := \frac{1}{t_k} \left[f(x^0, g(z^k) + t_k v) - f(x^0, g(z^k)) \right]$$

und erhalten $b_2^k \to b_2$. Die linke Seite definieren wir via

$$a^k := \frac{1}{t_k} \left[f(x^k + t_k u, g(z^k) + t_k v) - f(x^k, g(z^k)) \right].$$

Eine geeignete Teilfolge von $\{a^k\}$ konvergiert gegen ein a und wir erhalten zusammen mit (3.7):

$$\begin{aligned} a &= \lim_{t_k \searrow 0} \frac{1}{t_k} \left[f(x^k + t_k u, g(z^k + t_k w)) - f(x^k, g(z^k)) \right] = \\ &= \lim_{t_k \searrow 0} \frac{1}{t_k} \left[F(x^k + t_k u, z^k + t_k w) - F(x^k, z^k) \right] \in TF((x^0, z^0); (u, w)). \end{aligned}$$

Analoge Integralabschätzungen wie in Beweis von Lemma 3.7 (i) ergeben die Behauptung.

(ii) folgt sofort aus Proposition 3.5 (ii). #

3.2 Strenge Regularität und Pseudo-Regularität

Folgendes Beispiel beschreibt eine stückweise lineare Funktion, die gleichzeitig pseudo-regulär, jedoch nicht streng regulär im Nullpunkt ist:

Beispiel 3.10 Wir definieren eine Funktion $f : I\!R^2 \to I\!R^2$. Die erste Komponente f_1 von f sei gegeben durch

$$f_1(x,y) := \begin{cases} y & \text{für} \quad |y| \leq x \\ x & \text{für} \quad |x| \leq y \\ -y & \text{für} \quad |y| \leq -x \\ -x & \text{für} \quad |x| \leq -y \end{cases}$$

Jetzt bezeichne $r : I\!R^2 \to I\!R^2$ die lineare Transformation, die den Punkt (x,y) um den Winkel $-\frac{\pi}{4}$ dreht, d.h. $r(x,y) = \frac{\sqrt{2}}{2}(y+x, y-x)$. Nun sei die zweite Komponente f_2 von f definiert als $f_2(x,y) := f_1(r(x,y))$. Die explizite Gestalt von f_2 ist dann

$$f_2(x,y) = \begin{cases} \frac{\sqrt{2}}{2}(-x+y) & \text{für} \quad x \geq 0,\ y \geq 0 \\ \frac{\sqrt{2}}{2}(x+y) & \text{für} \quad x \leq 0,\ y \geq 0 \\ \frac{\sqrt{2}}{2}(x-y) & \text{für} \quad x \leq 0,\ y \leq 0 \\ \frac{\sqrt{2}}{2}(-x-y) & \text{für} \quad x \geq 0,\ y \leq 0 \end{cases}$$

Man sieht leicht, daß f pseudo-regulär in jedem Punkt ist. Außerdem hat jeder Bildpunkt $c = (a,b) \neq (0,0)$ genau zwei Urbilder: $f^{-1}(c) = \{z(c), -z(c)\}$ und $f^{-1}((0,0)) = \{(0,0)\}$. #

Nächste Aussage vergleicht die Kriterien für strenge bzw. Pseudo-Regularität von lokalen Lipschitz-Funktionen. Die Aussage (i) findet man in [23], die Aussage (ii) ist eine Umformulierung des Satzes 2.8.

Satz 3.11 *Es sei $f : I\!R^n \to I\!R^n$ lokal Lipschitz.*

(i) f ist streng regulär in x^0 $\Leftrightarrow 0 \notin Tf(x^0,u) \quad \forall u \neq 0$.

(ii) f ist pseudo-regulär in x^0 $\Leftrightarrow \exists \alpha > 0 : I\!B \subset \bigcup_{\|u\| \leq \alpha} Cf(x,u) \qquad \forall x \in x^0 + \alpha^{-1} I\!B$.

Bemerkung 3.12 $f \in C^1(I\!R^n I\!R^n)$ und f sei pseudo-regulär in x^0. Dann ist f streng regulär in x^0.

Beweis: Wenn man bedenkt, daß $Tf(x^0)u = Cf(x^0,u) = \{Df(x^0)u\}$, bekommt man aus der Pseudo-Regularität von f, daß $Df(x^0)$ regulär ist. Damit ist f streng regulär. #

3.3 Pseudo-Regularität für stückweise glatte Funktionen

Eine wichtige Klasse von Lipschitz-Funktionen bilden die stückweise glatten Funktionen. Derartige Abbildungen haben sehr "gute" Eigenschaften (im Vergleich zu einer beliebigen Lipschitz-Funktion), so daß man viele wichtige Aussagen über stetig differenzierbare Funktionen auf stückweise glatte Funktionen erweitern kann.

Definition 3.13 *Wir nennen eine Funktion $f : \mathbb{R}^n \to \mathbb{R}^n$* **stückweise glatt** *oder eine* **$\mathbf{PC^1}$-Funktion,** *wenn endlich viele C^1- Funktionen $f_i : \mathbb{R}^n \to \mathbb{R}^n,\ i = 1, \ldots, p$ existieren, so daß f eine stetige Auswahlfunktion von $f_1, \ldots, f_p$ ist, d.h. f ist stetig und*

$$f(x) \in \{f_1(x), \ldots, f_p(x)\} \qquad \forall x \in \mathbb{R}^n.$$

Zu einem beliebigen Punkt $x \in \mathbb{R}^n$ definieren wir die entsprechende **aktive Indexmenge** als $I_f(x) := \{i \,|\, i \in \{1, \ldots, p\},\, f(x) = f_i(x)\}$. Aus Stetigkeitsgründen können auf einer geeigneten Umgebung eines Punktes x^0 nur die Funktionen f_i mit $i \in I_f(x^0)$ aktiv werden. Wir sagen, die Funktion f_i ist in x^0 **wesentlich aktiv,** falls $x^0 \in \mathrm{cl}\ \mathrm{int}\{x \,|\, f(x) = f_i(x)\}$ gilt. OBdA können wir jetzt annehmen, daß alle Funktionen $f_i,\ i \in I_f(x^0)$ wesentlich aktiv sind.

Andernfalls gibt es eine Umgebung U von x^0 und ein Index $i \in I_f(x^0)$ mit der Eigenschaft, daß $U \cap \mathrm{int}\{x \,|\, f(x) = f_i(x)\} = \emptyset$ gilt. Deshalb kann es in U keinen Punkt $\bar{x}$ geben, in dem nur die Funktion f_i aktiv wäre, sonst gäbe es eine Umgebung von $\bar{x}$, wo das auch der Fall ist. Wir erhalten, daß f auf U eine Auswahlfunktion von $f_i,\ i \in I_f(x^0) \setminus \{i\}$ ist.

Aus Sicht unserer Untersuchungen ist der Übergang von C^1- Funktionen zu PC^1- Funktionen sehr interessant: während für C^1-Funktionen Pseudo-Regularität zugleich strenge Regularität bedeutete, ist dies bei PC^1-Funktionen anders (siehe Beispiel 3.10). Deshalb sind PC^1-Funktionen die Kandidaten mit einfachster Struktur, bei denen man die Unterschiede zwischen den Regularitätsbegriffen studieren kann.

Um das Kriterium aus dam Satz 3.11 (ii) anwenden zu können, brauchen wir eine Aussage über Contingent-Ableitungen stückweise glatter Funktionen. Diese wird vom nächsten bekannten Lemma geliefert:

Lemma 3.14 *([27], [42])*
Es sei $f : \mathbb{R}^n \to \mathbb{R}^n$ eine stetige Auswahlfunktion der C^1-Funktionen $f_i : \mathbb{R}^n \to \mathbb{R}^n$, $i = 1, \ldots, p$. Dann ist f richtungsdifferenzierbar und es gilt:

$$f'(x^0, u) \in \{Df_i(x^0)u \,|\, i \in I_f(x^0)\}.$$

Weiterhin benötigen wir einige nützliche Eigenschaften stückweise linearer Funktionen. Wir sagen, $f : \mathbb{R}^n \to \mathbb{R}^n$ ist **stückweise linear,** falls f eine stetige Auswahlfunktion einer endlichen Anzahl von linearen Funktionen $f_i(x) = A_i x,\ i = 1, \ldots, p$ ist. Insbesondere sind Richtungsableitungen von stückweise glatten Funktionen stückweise linear.

Unter einer **Kegelzerlegung** $\mathcal{C}$ von $I\!R^n$ verstehen wir eine Familie von n-dimensionalen (konvexen) polyhedralen Kegeln, so daß sich je zwei Kegel der Familie in einer gemeinsamen Seite schneiden und die Vereinigung aller Kegel den ganzen Raum $I\!R^n$ bildet. Zwei Kegel der Kegelzerlegung heißen **adjazent,** wenn ihr Durchschnitt eine $(n-1)$-dimensionale Seite ist. Die Vereinigung aller k-dimensionaler Seiten bezeichnen wir mit $skel_k\mathcal{C}$.

Man überlegt sich leicht, daß für eine beliebige stückweise lineare Funktion $f : I\!R^n \to I\!R^n$ eine Kegelzerlegung $\mathcal{C}$ von $I\!R^n$ existiert, so daß für jeden Kegel $C \in \mathcal{C}$ eine Matrix $A(C)$ existiert mit $f_{|C}(x) = A(C)x$. Wir sagen dann, daß f stückweise linear bzgl. $\mathcal{C}$ ist.

Definition 3.15 *Es sei $f : I\!R^n \to I\!R^n$ stückweise linear bzgl. einer Kegelzerlegung $\mathcal{C}$ von $I\!R^n$. Wir nennen f* **koherent orientiert,** *falls alle Matrizen $A(C),\ C \in \mathcal{C}$ regulär sind und ihre Determinanten dasselbe Vorzeichen besitzen.*

Das nächste Lemma (siehe [22]) charakterisiert die koherente Orientierung:

Lemma 3.16 *Es sei $f : I\!R^n \to I\!R^n$ stückweise linear bzgl. einer Kegelzerlegung $\mathcal{C}$ von $I\!R^n$. Die Funktion f ist koherent orientiert genau dann, wenn das Bild jedes Kegels der Kegelzerlegung $\mathcal{C}$ wieder ein Kegel ist und adjazente Kegel auch adjazente Bilder besitzen.*

Die Surjektivität ist eine notwendige Bedingung für Pseudo-Regularität. Wir werden später sehen, daß in diesem Zusammenhang folgendes Lemma wichtig ist:

Lemma 3.17 *([22])*
Es sei $\mathcal{C}$ eine Kegelzerlegung von $I\!R^n$ und $f : I\!R^n \to I\!R^n$ eine bzgl. $\mathcal{C}$ stückweise lineare Funktion. Wenn f koherent orientiert ist, so ist f surjektiv.

Beweis: Zuerst zeigen wir, daß die Funktion f folgende Eigenschaft besitzt:
Ist die Menge $A \subset I\!R^n$ abgeschlossen, so ist auch ihr Bild $f(A)$ abgeschlossen. Es sei also eine Folge $\{a_n\} \subset A$ gegeben und die Folge $\{f(a_n)\}$ konvergiere gegen ein y. Die Funktion f sei eine stetige Auswahlfunktion der Funktionen $f_i,\ i = 1, \ldots, p$. Die Menge $\{f(a_n)\}_{n=1}^{\infty} \cup \{y\}$ ist kompakt. Da f koherent orientiert ist, sind auch alle Mengen $f_i^{-1}(\{f(a_n)\}_{n=1}^{\infty} \cup \{y\})$ kompakt. D.h., auch die Menge $f^{-1}(\{f(a_n)\}_{n=1}^{\infty} \cup \{y\})$ ist kompakt und $\{a_n\}$ konvergiert gegen ein a mit $a \in A$ (A ist abgeschlossen). Wir erhalten so $f(a) = y \in f(A)$, d.h. $f(A)$ ist abgeschlossen. Daraus folgt: $f(I\!R^n)$ ist ebenfalls abgeschlossen.

Die Funktion f erhält nach Lemma 3.16 die Adjazenz von Kegeln. Daraus folgt, f ist ein lokaler Homöomorphismus in jedem Punkt $x \in X := I\!R^n \setminus f^{-1}f(skel_{n-2}\mathcal{C})$ (d.h., es gibt

Umgebungen U von x und V von $f(x)$, so daß $f_{|U} : U \to V$ bijektiv ist). Da X offen ist, erhalten wir dadurch, daß auch $f(X)$ offen ist.

Daraus folgt, daß der Rand von $f(\mathbb{R}^n)$ in der Menge $f(\mathbb{R}^n) \setminus f(X) = f(skel_{n-2}\mathcal{C})$ enthalten ist. Diese Menge ist wiederum $(n-2)$-dimensional, d.h. sie trennt $\mathbb{R}^n$ nicht, und das bedeutet $f(\mathbb{R}^n) = \mathbb{R}^n$. #

Folgender Satz charakterisiert die Pseudo-Regularität von stückweise glatten Funktionen:

Satz 3.18 *Es sei $f : \mathbb{R}^n \to \mathbb{R}^n$ eine stetige Auswahlfunktion der C^1-Funktionen $f_i : \mathbb{R}^n \to \mathbb{R}^n$, $i = 1, \ldots, p$. . OBdA seien alle f_i in einem Punkt x^0 wesentlich aktiv.*

(i) Wenn alle Matrizen $Df_i(x^0)$ regulär sind und ihre Determinanten dasselbe Vorzeichen besitzen, dann ist f pseudo-regulär in x^0.

(ii) Wenn die Funktion f pseudo-regulär im Punkt x^0 ist, so ist die Richtungsableitung $f'(x^0, .)$ koherent orientiert.

Beweis: (i): OBdA können wir annehmen, daß alle Determinanten $\det Df_i(x^0)$ positiv sind. Aus Stetigkeitsgründen gibt es eine Umgebung U von x^0, wo dies erhalten bleibt und $I_f(x) \subset I_f(x^0)\ \forall x \in U$, d.h. $f'(x, .)$ ist koherent orientiert und damit surjektiv $\forall x \in U$. Weiter gilt:

$$\min_{i \in I_f(x^0)} \min_{u \in S} \|Df_i(x^0)u\| > 0$$

(wobei S die Einheitssphäre bezeichnet) und es existiert eine Konstante $c > 0$ mit

$$\min_{i \in I_f(x^0)} \min_{u \in S} \|Df_i(x)u\| \geq c \qquad \forall x \in U.$$

Nun setzen wir $\alpha := c^{-1}$ und es sei ein $v \in \mathbb{B}$ gegeben. Wegen der Surjektivität gibt es für jedes $x \in U$ eine Richtung u mit $v = f'(x, u)$ und

$$\|v\| = \|f'(x, u)\| = \|u\| \left\| f'\left(x, \frac{u}{\|u\|}\right) \right\| \geq \|u\| c = \frac{\|u\|}{\alpha}.$$

Damit gilt: $\mathbb{B} \subset \bigcup_{\|u\| \leq \alpha} f'(x, u)\ \forall x \in U$. Gegebenenfalls vergrößern wir α und erhalten die Behauptung.

(ii): Wir definieren $X_i := \{x \mid f(x) = f_i(x)\}$. Der Satz 3.11 (ii) angewendet auf Punkte $x \in \operatorname{int} X_i \cap (x^0 + \alpha^{-1}\mathbb{B})$ ergibt $\|Df_i(x)^{-1}\| \leq \alpha$. Mit anderen Worten:

$$\min_{u \in S} \|Df_i(x)u\| \geq \alpha^{-1},$$

d.h. insbesondere sind alle Matrizen $Df_i(x^0)$, $i \in I_f(x^0)$ regulär. Weiter verfahren wir indirekt.

Die Funktion $g := f'(x^0, .)$ ist stückweise linear, d.h. es gibt eine Kegelzerlegung $\mathcal{C} = \{C_i\}_{i \in I_f(x^0)}$ von $I\!R^n$, so daß auf dem Kegel C_i die Abbildung $g_i(u) = Df_i(x^0)u$ aktiv ist. Angenommen, es existieren Indizes $i, j \in I_f(x^0)$ mit $\det\, Df_i(x^0) < 0 < \det\, Df_j(x^0)$. OBdA können wir voraussetzen, daß C_i und C_j adjazent sind (sonst bilden wir eine Kette von adjazenten Kegeln $C_i = A_1, \ldots, A_k = C_j$ und es müssen zwei Kegel in der Kette mit der gewünschten Eigenschaft existieren). Wir wählen einen Punkt $\bar{u} \in C_i \cap C_j \setminus \{0\}$ beliebig nahe am Nullpunkt. Da $\det\, Df_i(x^0) < 0 < \det\, Df_j(x^0)$ ist, sind die Mengen

$$Cg(\bar{u}, C_i) := \bigcup_{w \in C_i} Cg(\bar{u}, w)$$

und $Cg(\bar{u}, C_j)$ im denselben Halbraum enthalten. D.h., die Abbildung $Cg(\bar{u}, .)$ ist nicht surjektiv. Aus Satz 3.11(ii) folgt damit, daß g nicht pseudo-regulär im Punkt $\bar{u}$ sein kann. Schließlich ergibt Korollar 2.7 einen Widerspruch zur Pseudo-Regularität von f. #

Korollar 3.19 *Es sei $f : I\!R^n \to I\!R^n$ eine stetige Auswahlfunktion der C^1-Funktionen $f_i : I\!R^n \to I\!R^n,\ i = 1, \ldots, p.$. OBdA seien alle f_i in einem Punkt x^0 wesentlich aktiv.*

Ist f pseudo-regulär in einem Punkt x^0, dann liegt Pseudo-Regularität vom Typ 3 vor.

Beweis: Aus der Pseudo-Regularität von f folgt insbesondere, daß f^{-1} lokal durch reguläre Funktionen $f_i^{-1},\ i \in I_f(x^0)$ gegeben ist. D.h., für y hinreichend nahe an $f(x^0)$ erhalten wir endlich viele, und zwar höchstens p Urbilder $x \in f^{-1}(y)$. Der Satz über die implizite Funktion ergibt die Behauptung. #

In dem Fall, daß zur Bildung der Richtungsableitung $f'(x^0, .)$ alle Matrizen $Df_i(x^0)$, $i \in I_f(x^0)$ benötigt werden, bekommen wir aus dem Satz 3.18 eine notwendige und hinreichende Bedingung:

Korollar 3.20 *Es sei $f : I\!R^n \to I\!R^n$ stückweise glatt und $\mathcal{C}$ sei die Kegelzerlegung, bzgl. der $f'(x^0, .)$ stückweise linear ist. Zusätzlich existiere für jedes $i \in I_f(x^0)$ ein Kegel $C_i \in \mathcal{C}$, so daß $f'(x^0, u) = Df_i(x^0)u$ auf dem Kegel C_i gilt.*

Dann ist f pseudo-regulär in x^0 genau dann, wenn alle Matrizen $Df_i(x^0),\ i \in I_f(x^0)$ regulär sind und ihre Determinanten dasselbe Vorzeichen besitzen.

Beweis: In diesem Fall ist die koherente Orientierung von $f'(x^0, .)$ äquivalent zu der Bedingung, daß alle Matrizen $Df_i(x^0),\ i \in I_f(x^0)$ regulär sind und dasselbe Determinantenvorzeichen besitzen. #

In [34] sind Bedingungen aufgeführt, unter denen eine stückweise glatte Funktion f streng regulär ist. Es handelt sich dabei um die Bedigung (ii) aus Satz 3.18 mit zusätzlichen Forderungen (Invertierbarkeit der Richtungsableitung oder Bedingungen an den Abbildungsgrad von f bzw. $f'(x, .)$). Außerdem sind positive Aussagen mit Hilfe der koherenten

Orientierung in den Fällen möglich, wo die Mengen $\{x \mid f(x) = f_i(x)\}$ eine Kegelzerlegung von $\mathbb{R}^n$ bilden und die Anzahl der Kegel höchstens vier ist (dazu siehe [42]). In diesem Spezialfall sind Pseudo-Regularität und strenge Regularität äquivalent.

Folgendes Beispiel zeigt, daß im allgemeinen die koherente Orientierung der Richtungsableitung $f'(x^0, .)$ nicht hinreichend für die Pseudo-Regularität ist :

Beispiel 3.21 Es sei eine Funktion $f : \mathbb{R}^2 \to \mathbb{R}^2$ definiert durch

$$f_1(x,y) = \begin{cases} (x - \sqrt{y})\sqrt{y} + \sqrt{y} & x \in \left[\sqrt{y}, \sqrt{y} + \frac{1}{2}(\sqrt[3]{y} - \sqrt{y})\right], \ y \in (0,1) \\ \sqrt[3]{y} - (\sqrt[3]{y} - x)(2 - \sqrt{y}) & x \in \left[\sqrt{y} + \frac{1}{2}(\sqrt[3]{y} - \sqrt{y}), \sqrt[3]{y}\right], \ y \in (0,1) \\ x & \text{sonst} \end{cases}$$

$$f_2(x,y) = y.$$

Man kann nachweisen, daß $f'((x,y);.)$ koherent orientiert ist für alle x nahe am Nullpunkt:

$$f'((x,y);u) \in \{A_1 u, A_2 u, A_3 u\}$$

$$A_1 = \begin{pmatrix} \sqrt{y} & -1 + \frac{x+1}{2\sqrt{y}} \\ 0 & 1 \end{pmatrix} \qquad A_2 = \begin{pmatrix} 2 - \sqrt{y} & -\frac{1}{3y^{2/3}} + \frac{1}{6y^{5/6}} - \frac{1}{2}xy^{-1/2} \\ 0 & 1 \end{pmatrix}$$

$$A_3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Es ist $f'((0,0);u) = A_3 u = u$, jedoch ist f nicht pseudo-regulär in $(0,0)$. #

3.4 Einige Eigenschaften von Lipschitz-Funktionen

In diesem Abschnitt beweisen wir das angekündigte Dimensionslemma, das die Dimensionen von Urbildmengen und Bildmengen bei Lipschitz-Funktionen vergleicht.

Die Peano-Abbildung ist ein bekanntes Beispiel einer Funktion, die das Einheitsintervall $[0,1]$ auf das Quadrat $[0,1] \times [0,1]$ abbildet. Unser Ziel ist es zu zeigen, daß ähnliche (in dem Sinne, daß das Bild größere Dimension als der Urbildraum hat) Abbildungen nicht lokal Lipschitz sein können.

Beispiel 3.22 (siehe [13], Beispiel 4.4.5) Wir erhalten die Peano-Abbildung als den Limes einer Folge von Funktionen $f_n : [0,1] \to [0,1] \times [0,1]$. Dazu teilen wir das Einheitsintervall $[0,1]$ in 9 gleiche Intervalle mit der Länge 1/9 und das Quadrat $[0,1] \times [0,1]$ in 9 gleiche Quadrate mit der Seitenlänge 1/3. Die Funktion f_1 sei gegeben durch:

$$f_1(0) = (0,0), \quad f_1\left(\tfrac{1}{9}\right) = \left(\tfrac{1}{3},\tfrac{1}{3}\right), \quad f_1\left(\tfrac{2}{9}\right) = \left(0,\tfrac{2}{3}\right), \quad f_1\left(\tfrac{1}{3}\right) = \left(\tfrac{1}{3},1\right), \quad f_1\left(\tfrac{4}{9}\right) = \left(\tfrac{2}{3},\tfrac{2}{3}\right),$$
$$f_1\left(\tfrac{5}{9}\right) = \left(1,\tfrac{1}{3}\right), \quad f_1\left(\tfrac{2}{3}\right) = \left(\tfrac{2}{3},0\right), \quad f_1\left(\tfrac{7}{9}\right) = \left(\tfrac{1}{3},\tfrac{1}{3}\right), \quad f_1\left(\tfrac{8}{9}\right) = \left(\tfrac{2}{3},\tfrac{2}{3}\right), \quad f_1(1) = (1,1).$$

und durch lineare Extrapolation im Innern der Intervalle.

Jetzt setzen wir voraus, daß wir das Intervall $[0,1]$ bereits in 9^n gleiche Intervalle und das Quadrat $[0,1]\times[0,1]$ in 9^n gleiche Quadrate mit der Seitenlänge 3^{-n} geteilt haben und daß eine stetige Funktion $f_n : [0,1] \to [0,1]\times[0,1]$ gegeben ist. Wir nehmen ein beliebiges Intervall $[a,b]$, das durch die Teilung des Intervalls $[0,1]$ entstanden ist und es sei $f_n(a) = (x_a, y_a),\ f_n(b) = (x_b, y_b)$. Der Vektor $f_n(b) - f_n(a)$ gehöre dem i-ten Quadranten an und r sei die Rotationsabbildung, die eine Rotation um den Mittelpunkt des Quadrates $[0,1]\times[0,1]$ um den Winkel $\frac{\pi}{2}$ realisiert. Für t mit $a \le t \le b$ definieren wir f_{n+1} wie folgt:

$$f_{n+1}(t) := (x_a, y_a) + 3^{-n} r^{i-1}(f_1(t)). \tag{3.8}$$

Auf diese Weise erhalten wir eine stetige Abbildung $f_{n+1} : [0,1] \to [0,1]\times[0,1]$. Die Funktionenfolge $\{f_n\}$ konvergiert (punktweise) gegen eine stetige Funktion $f : [0,1] \to [0,1]\times[0,1]$ mit $f([0,1]) = [0,1]\times[0,1]$ (siehe [13]).

Wir sehen: $f(0) = 0$ und $f(9^{-n}) = (3^{-n}, 3^{-n})$ für alle $n \in I\!N$ und erhalten ($x = 0$, $y = 9^{-n}$ und $\|.\|$ ist die Maximumnorm):

$$\frac{\|f(y) - f(x)\|}{\|y - x\|} = \frac{3^{-n}}{9^{-n}} = 3^n. \tag{3.9}$$

Mit anderen Worten, die Funktion f ist nicht lokal Lipschitz im Nullpunkt. #

Der Beweis des nächsten Lemmas nutzt denselben Effekt: Wegen des Dimensionsunterschiedes kann die potentielle Lipschitz-Konstante nicht endlich sein.

Lemma 3.23 *(Dimensionslemma)*
Es sei $f : C \to I\!R^k$ lokal Lipschitz, $C \subset I\!R^n$ sei abgeschlossen und int $f(C) \neq \emptyset$. *Dann ist $n \ge k$.*

Beweis: Angenommen, $n < k$. Wir zerlegen den Urbildraum $I\!R^n$ durch affine Hyperebenen $x_i = c,\ i = 1,\dots n,\ c \in \mathbb{Z}$ in Würfel mit der Seitenlänge 1. Dann existiert ein Würfel A mit $\lambda^k(f(C\cap A)) > 0$, wobei λ^k das Lebesgue-Maß in $I\!R^k$ bezeichnet. OBdA können wir annehmen, daß $\lambda^k(f(C\cap A)) = 1$.

Jetzt zerlegen wir den Würfel A in 2^n gleiche Würfel $A_1,\dots A_{2^n}$ mit der Seitenlänge 1/2. Dann existiert ein Index $i \in \{1,\dots,2^n\}$, so daß gilt:

$$\lambda^k(f(C\cap A_i)) \ge \frac{1}{2^n}. \tag{3.10}$$

Wir definieren $d := \operatorname{diam} f(C \cap A_i) = \sup\{\|y_1 - y_2\| \mid y_1, y_2 \in f(C \cap A_i)\} < \infty$ und finden zwei Punkte $y_1, y_2 \in f(C \cap A_i)$ mit $\|y_1 - y_2\| \geq d/2$ und $x_j \in f^{-1}(y_j), x_j \in C \cap A_i, j = 1, 2$.

Wir wissen: $f(C \cap A_i) \subset y_1 + d\mathbb{B}$ und daraus folgt

$$\lambda^k(f(C \cap A_i)) \leq (2d)^k. \tag{3.11}$$

Zusammen mit (3.10) ergibt das

$$d \geq \frac{1}{2^{1+\frac{n}{k}}}. \tag{3.12}$$

Andererseits gilt: $\|x_1 - x_2\| \leq \dfrac{1}{2}$ wegen $x_1, x_2 \in A_i$. Wir erhalten

$$\frac{\|f(x_1) - f(x_2)\|}{\|x_1 - x_2\|} \geq d \geq \frac{1}{2^{\frac{n}{k}+1}}. \tag{3.13}$$

Nun setzen wir diese Konstruktion induktiv fort. Es sei $A^1_{i(1)} := A_i$ und angenommen, im s-ten Schritt, haben wir einen Würfel $A^{s-1}_{i(s-1)}$ mit der Seitenlänge $2^{-(s-1)}$, so daß gilt:

$$\lambda^k(f(C \cap A^{s-1}_{i(s-1)})) \geq \frac{1}{2^{n(s-1)}}$$

Jetzt zerlegen wir den Würfel $A^{s-1}_{i(s-1)}$ in 2^n gleiche Würfel $A^s_1, \ldots, A^s_{2^n}$ mit der Seitenlänge 2^{-s}. Es gibt einen Index $i(s) \in \{1, \ldots, 2^n\}$, so daß gilt:

$$\lambda^k(f(C \cap A^s_{i(s)})) \geq \frac{1}{2^{ns}}. \tag{3.14}$$

Analog definieren wir $d_s := \operatorname{diam} f(C \cap A^s_{i(s)})$ und finden Punkte $y^s_1, y^s_2 \in f(C \cap A^s_{i(s)})$ mit $\|y^s_1 - y^s_2\| \geq d_s/2$ und $x^s_j \in f^{-1}(y^s_j), x^s_j \in C \cap A^s_{i(s)}, j = 1, 2$.

Ähnlich wie in (3.11) ergibt das $\lambda^k(f(C \cap A^s_{i(s)})) \leq (2d_s)^k$ und

$$d_s \geq \frac{1}{2^{1+\frac{ns}{k}}}. \tag{3.15}$$

Wegen $x^s_1, x^s_2 \in A^s_{i(s)}$ gilt $\|x^s_1 - x^s_2\| \leq \frac{1}{2^s}$ und

$$\frac{\|f(x^s_1) - f(x^s_2)\|}{\|x^s_1 - x^s_2\|} \geq \frac{2^s d_s}{2} \geq 2^{s - \frac{ns}{k} - 2} = 2^{s(1-\frac{n}{k})-2}. \tag{3.16}$$

Nun betrachten wir die Folgen $\{x^s_1\}$ und $\{x^s_2\}$. Aus $x^s_j \in A^s_{i(s)} \subset A^{s-1}_{i(s-1)}$ folgt

$$\|x^s_j - x^{s-1}_j\| \leq \frac{1}{2^{s-1}},$$

so daß beide Folgen gegen den Punkt $x^* \in \bigcap_{s=1}^{\infty} A^s_{i(s)}$ konvergieren.

Wenn wir uns daran erinnern, daß wir $n < k$ angenommen haben, sehen wir, daß der rechte Term in (3.16) gegen ∞ konvergiert (für $s \to \infty$). D.h., f ist nicht lokal Lipschitz in x^*. #

Die folgenden Überlegungen werden angestellt, um zu zeigen, daß die von uns benutzten Standard-Transformationen nicht aus der Klasse der lokalen Lipschitz-Funktionen herausführen.

Bemerkung 3.24

$$\forall x, y \in I\!R^n,\ x, y \neq 0: \quad \left\| \frac{x}{\|x\|} - \frac{y}{\|y\|} \right\| \leq \frac{\|x-y\|}{\min\{\|x\|, \|y\|\}}. \tag{3.17}$$

OBdA sei $\|x\| \leq \|y\|$. Wir projezieren den Punkt y auf die Kugel $xI\!B$ und erhalten aus der Nichtexpansivität der Projektionsabbildung: $\left\| x - \frac{y}{\|y\|}\|x\| \right\| \leq \|x-y\|$. #

Bemerkung 3.25 *Inversion* (siehe [13])
Es sei eine positive reelle Zahl r gegeben. Jedem Punkt $x \in I\!R^n \setminus \{0\}$ entspricht genau ein Punkt $i(x)$, der auf der Halbgeraden liegt, die im Nullpunkt beginnt und durch x geht, so daß gilt: $\|i(x)\|\|x\| = r^2$. Wir setzen also $i(x) = tx,\ t \in I\!R_+$ und erhalten $\|tx\|\|x\| = r^2$ und $t = \frac{r^2}{\|x\|^2}$. Das ergibt $i(x) = \frac{r^2}{\|x\|^2}x$.

Die Abbildung $i : I\!R^n \setminus \{0\} \to I\!R^n \setminus \{0\}$ wird **Inversion** in der Sphäre mit dem Mittelpunkt im Nullpunkt und mit Radius r genannt. Es handelt sich um eine bijektive Abbildung $I\!R^n \setminus \{0\} \to I\!R^n \setminus \{0\}$ mit $i^2 = id$. #

Bemerkung 3.26 *Stereographische Projektion* (siehe [13])
Es sei eine positive Zahl r gegeben. Wir betrachten die Sphäre S in $I\!R^n$ mit dem Mittelpunkt $(0, \ldots, 0, r)$ und dem Radius r. Mit H bezeichnen wir die Hyperebene, die durch die Gleichung $x_n = 2r$ definiert ist. Jede Gerade, die durch den Nullpunkt und durch einen Punkt $x \in S \setminus \{0\}$ führt, trifft H in genau einem Punkt, den wir mit $s(x)$ bezeichnen. Um $s(x)$ zu bestimmen, setzen wir $s(x) = tx,\ t \in I\!R_+$. Aus $tx_n = 2r$ erhalten wir $t = \frac{2r}{x_n}$ und

$$s(x) = \frac{2r}{x_n}x.$$

Die Abbildung $s : S \setminus \{0\} \to H$ wird **stereographische Projektion** der Sphäre S aus dem Pol 0 auf die Hyperebene H genannt.

Den Durchschnitt der Sphäre S und des Halbraumes, der durch die Ungleichung $x_n \geq r$ definiert ist, bezeichnen wir mit T. Nun schränke man die Abbildung s auf die Menge T

ein. Wir möchten zeigen, daß $s_{|_T}$ lokal Lipschitz und Lipschitz invertierbar ist (d.h. $s_{|_T}^{-1}$ ist eine lokale Lipschitz-Funktion). Dazu betrachten wir die Inversion $i_{|_T}$ in der Sphäre mit dem Mittelpunkt im Nullpunkt und mit Radius $2r$ eingeschränkt auf die Menge T. Es sei $x \in T$, dann gilt: $\|x - (0, \ldots, 0, r)\|^2 = r^2$, d.h. $\|x\|^2 - 2x_n r + r^2 = r^2$ und $\|x\|^2 = 2x_n r$. Wir erhalten:

$$i_{|_T}(x) = \frac{4r^2}{\|x\|^2} x = \frac{2r}{x_n} x = s_{|_T}(x).$$

Daher gilt $s_{|_{s(T)}}^{-1} = i_{|_{s(T)}}^{-1} = i_{|_{s(T)}}$. Deshalb genügt es zu zeigen, daß $i_{|_{T \cup s(T)}}$ lokal Lipschitz ist.

Es seien Punkte $x, y \in T \cup s(T)$ gegeben, dann gilt $r \leq \|x\|$ and $r \leq \|y\|$. Außerdem haben wir:

$$\begin{aligned}
\|i(x) - i(y)\| &= 4r^2 \left\| \frac{x}{\|x\|^2} - \frac{y}{\|y\|^2} \right\| \leq 4r^2 \left\| \frac{x}{\|x\|^2} - \frac{x}{\|y\|^2} \right\| + 4r^2 \left\| \frac{x}{\|y\|^2} - \frac{y}{\|y\|^2} \right\| \leq \\
&\leq 4\|x\| r^2 \frac{|\,\|y\|^2 - \|x\|^2\,|}{\|x\|^2 \|y\|^2} + \frac{4r^2}{\|y\|^2} \|x - y\| \leq \\
&\leq 4\|x\| r^2 \|x - y\| \frac{1}{\|x\|\,\|y\|} \left(\frac{1}{\|x\|} + \frac{1}{\|y\|} \right) + 4\|x - y\| \leq 12\|x - y\|.
\end{aligned}$$

Nun sei $r = 1$ und wir setzen $t(x) := (s_1(x + e_n) - 2, \ldots, s_{n-1}(x + e_n) - 2)$ (wobei $e_n := (0, \ldots, 0, 1)$ und s_i die i-te Komponente von s ist). Wenn wir die Hyperebene H mit $\mathbb{R}^{n-1}$ identifizieren, erhalten wir dadurch eine bijektive Abbildung, die die "obere Hälfte" der Einheitssphäre S auf $t(S) \subset \mathbb{R}^{n-1}$ abbildet und die in beiden Richtungen lokal Lipschitz ist. #

3.5 Urbildmengen von Lipschitz-Funktionen

Das Ziel dieses Abschnitts ist es, hinreichende Bedingungen für die Isoliertheit der Urbildmengen von lokalen Lipschitz-Funktionen zu formulieren. Dazu sei eine lokale Lipschitz-Funktion $f : \mathbb{R}^n \to \mathbb{R}^n$ gegeben, die in x^0 pseudo-regulär ist. Wann ist x^0 ein isoliertes Urbild von $f(x^0)$?

Bemerkung 3.27 Angenommen, in x^0 häufen sich die Urbilder von $f(x^0)$, dann gibt es eine Folge $\{x^k\}$, die gegen x^0 konvergiert und $f(x^k) = f(x^0)\ \forall k$. Wenn wir die Richtungen

$$u^k = \frac{x^k - x^0}{\|x^k - x^0\|}$$

definieren, können wir annehmen, daß die Folge $\{u^k\}$ gegen ein geeignetes u^0, $\|u^0\| = 1$ konvergiert (mindestens als Teilfolge). Wir setzen $t_k := \|x^k - x^0\|$ und erhalten:

$$\frac{1}{t_k} \left[f(x^0 + t_k u^k) - f(x^0) \right] = \frac{1}{t_k} \left[f(x^k) - f(x^0) \right] = 0$$

und $0 \in Cf(x^0, u^0)$.

Der folgende bekannte Satz charakterisiert die Differenzierbarkeitseigenschaften von lokalen Lipschitz-Funktionen (siehe [14]). Mit der Formulierung "für fast alle" ist gemeint, daß die Menge der Werte, die das nicht erfüllt, eine Nullmenge ist.

Satz 3.28 *(Satz von Rademacher):*
Es sei $f : \mathbb{R}^n \to \mathbb{R}^n$ lokal Lipschitz. Dann ist f fast überall im $\mathbb{R}^n$ differenzierbar.

Das nächste Lemma ist eine einfache Folgerung aus dem Satz von Rademacher:

Lemma 3.29 *Es sei $f : \mathbb{R}^n \to \mathbb{R}^n$ lokal Lipschitz und pseudo-regulär in x^0. Dann existieren Umgebungen U von x^0 und V von $f(x^0)$, so daß gilt:*
für fast alle $y \in V$ sind die Mengen $f^{-1}(y) \cap U$ endlich.

Beweis: Aus Bemerkung 1.4(ii) erhalten wir eine Umgebung U von x^0, so daß f pseudo-regulär in x für alle $x \in U$ ist. Wir setzen $V := f(U)$.

Aus dem Satz von Rademacher folgt: Die Funktion f ist differenzierbar in fast allen Punkten $x \in U$. Wir bezeichnen mit U_d die Menge aller Punkte $x \in U$ mit dieser Eigenschaft und erhalten: $\forall x \in U_d :\ Df(x)$ hat den vollen Rang (siehe Satz 3.11) und x ist ein isoliertes Urbild von $f(x)$ (siehe Bemerkung 3.27). Andererseits gilt $\lambda^n(U \setminus U_d) = 0$ und $\lambda^n(f(U \setminus U_d)) = 0$. (Man überdecke die Menge $U \setminus U_d$ durch eine Familie von Kugeln mit kleinen Radien. Wir können leicht eine Lipschitz-Abschätzung für die Inhalte der Bilder der Kugeln angeben; also ist die Menge $f(U \setminus U_d)$ auch eine Nullmenge.)

Auf diese Weise haben wir gezeigt: Für fast alle $y \in V = f(U)$ gilt: $f^{-1}(y) \cap U \subset U_d$, d.h. jedes $x \in f^{-1}(y) \cap U$ ist ein isoliertes Urbild von y und die Menge $f^{-1}(y) \cap U$ ist endlich.
#

Nach diesem Lemma können wir die rechte Seite der Gleichung $f(x) = y$ so stören, daß wir isolierte Lösungen erhalten. Unser Ziel ist aber eine allgemeine Aussage für alle möglichen Bilder. Die Antwort gibt der folgende Satz:

Satz 3.30 *Es sei eine lokale Lipschitz-Funktion $f : \mathbb{R}^n \to \mathbb{R}^n$ gegeben, die in x^0 pseudo-regulär ist. Unter der zusätzlichen Voraussetzung, daß f richtungsdifferenzierbar ist, existieren Umgebungen U von x^0 und V von $f(x^0)$, so daß gilt:*

(i) $\forall u \neq 0 \quad \forall x \in U :\ f'(x, u) \neq 0$ *(d.h. x ist ein isoliertes Urbild von $f(x)$)*

(ii) $\forall y \in V :$ *die Mengen $f^{-1}(y) \cap U$ sind endlich.*

Wäre $f'(x^0, u) = 0$ für ein $u \neq 0$, würde die Funktion $f'(x^0, .)$ auf dem ganzen Strahl $\{tu \,|\, t \geq 0\}$ den Wert Null annehmen. Man "spürt" intuitiv, daß eine Dimension im Bildraum "fehlt", um die angenommenen Eigenschaften aufrecht erhalten zu können.

Beweis: Um die Schreibweise abzukürzen, schreiben wir $g(u) = f'(x^0, u)$. Wir nehmen das Gegenteil der Behauptung (i) an, d.h. wir haben eine Richtung u^0 mit $g(u^0) = 0$ und $\|u^0\| = 1$. Wir wissen, (siehe Korollar 2.7), daß g bzw. g^{-1} ist eine Lipschitzstetige Funktion bzw. Multifunktion ist. Mit L_g bzw. L sollen die entsprechenden Lipschitz-Konstanten bezeichnet werden. Wir definieren eine positive reelle Zahl r durch

$$r := \frac{1}{2LL_g} \tag{3.18}$$

und betrachten folgende Menge:

$$A := \left\{ \, u \in I\!R^n \mid \|u - u^0\| = r, \; (u - u^0)^T u^0 = 0 \right\}. \tag{3.19}$$

Man sieht, daß die Menge A der Durchschnitt der Sphäre mit dem Mittelpunkt u^0 und Radius r mit der $(n-1)$-dimensionalen affinen Hyperebene H ist, die senkrecht zu der Geraden steht, die durch u^0 und den Nullpunkt geht. Es sei eine feste positive Zahl ε gegeben mit

$$L\varepsilon < \frac{r}{4 + 2r}. \tag{3.20}$$

Jetzt definieren wir die Menge

$$A_\varepsilon := \{ \, u \in A \mid \|g(u)\| \geq \varepsilon \} \,. \tag{3.21}$$

Wir unterscheiden weiter zwischen zwei Fällen.

1. Angenommen, die Menge A_ε ist nicht leer. Dann erklären wir eine Funktion $h : A_\varepsilon - u^0 \to I\!R^n$ durch

$$h(w) := \frac{g(u^0 + w)}{\|g(u^0 + w)\|}.$$

Mit Hilfe von Bemerkung 3.24 können wir schreiben: $\forall w_1, w_2 \in A_\varepsilon - u^0$

$$\begin{aligned} \|h(w_1) - h(w_2)\| &= \left\| \frac{g(u^0 + w_1)}{\|g(u^0 + w_1)\|} - \frac{g(u^0 + w_2)}{\|g(u^0 + w_2)\|} \right\| \leq \\ &\leq \frac{L_g \|w_1 - w_2\|}{\min\{\|g(u^0 + w_1)\|, \|g(u^0 + w_2)\|\}} \leq \frac{L_g}{\varepsilon} \|w_1 - w_2\|. \end{aligned} \tag{3.22}$$

Wir zeigen nun als wichtiges Zwischenergebnis: Es existiert ein Punkt $v \in I\!R^n$, $\|v\| = 1$ mit $v \notin h(A_\varepsilon - u^0)$. Angenommen, so ein v gibt es nicht, d.h. $h(A_\varepsilon - u^0)$ überdeckt die ganze Einheitssphäre S^n in $I\!R^n$. Wir wollen die stereographische Projektion auf die

Mengen A und $h(A_\varepsilon - u^0)$ anwenden. Auf diese Weise würden wir eine Lipschitz-Funktion von $I\!R^{n-2}$ auf eine Teilmenge von $I\!R^{n-1}$ mit nichtleerem Inneren erhalten. Das würde einen Widerspruch zu Lemma 3.23 bilden.

Dazu identifizieren wir die $(n-1)$-dimensionale Hyperebene H mit $I\!R^{n-1}$, so daß der Punkt u^0 dem Nullpunkt in $I\!R^{n-1}$ entspricht. Den Durchschnitt der Menge $h(A_\varepsilon - u^0)$ mit dem Halbraum, der durch die Ungleichung $x_n \geq 0$ in $I\!R^n$ definiert ist, bezeichne man mit D. Wir untersuchen das Urbild $h^{-1}(D)$ in $H = I\!R^{n-1}$. Wenn wir $h^{-1}(D)$ mit dem Halbraum schneiden, der durch die Ungleichung $x_{n-1} \geq 0$ bzw. $x_{n-1} \leq 0$ definiert ist, erhalten wir Mengen C_1 bzw. C_2. Für mindestens eine von den beiden gilt: Ihr Bild hat ein nichtleeres Inneres bezüglich der Sphäre S^n (eine Folgerung aus dem Baire'schen Satz). Mit anderen Worten: $\exists\, i \in \{1,2\} : \ \mathrm{int}_{S^n} h(C_i) \neq \emptyset$. OBdA sei $i = 1$. Dadurch erhalten wir die Mengen $C_1 \subset S^{n-1} \subset I\!R^{n-1}$ bzw. $h(C_1) \subset S^n \subset I\!R^n$, die wir auf $I\!R^{n-2}$ bzw. $I\!R^{n-1}$ stereographisch projezieren können.

Die Abbildung t_n beschreibe die Transformation aus der Bemerkung 3.26, wobei n die Dimension des Urbildraumes der Projektionsabbildung bezeichnen soll. Nun definieren wir:

$$\tilde{h} := t_n \circ h \circ t_{n-1}^{-1}, \quad \tilde{h} : \tilde{C} := t_{n-1}(C_1) \subset I\!R^{n-2} \to I\!R^{n-1}, \tag{3.23}$$

und $\tilde{h}$ ist Lipschitz (siehe (3.22)). Wir wissen, daß $t_n(h(C_1)) = \tilde{h}(\tilde{C})$ und daß t_n ein Homöomorphismus ist, also ergibt das: int $(\tilde{h}(\tilde{C})) \neq \emptyset$. Nach Lemma 3.23 kann die Funktion $\tilde{h}$ nicht existieren, wir erhalten also einen Widerspruch.

Daraus folgt in der Tat: Es gibt einen Punkt $v \in S^n, \ v \notin h(A_\varepsilon - u^0)$.

Für dieses spezielle v können wir die folgende Menge definieren

$$K_v := \{u \in \mathrm{conv}\ A |\ g(u) = \gamma v, \gamma \geq 0\} \tag{3.24}$$

und erhalten:

$$K_v \cap A_\varepsilon = \emptyset. \tag{3.25}$$

2. Sei $A_\varepsilon = \emptyset$. In diesem zweiten Fall wählen wir einen beliebigen Punkt v, $\|v\| = 1$ und definieren die entsprechende Menge K_v wie oben. Trivialerweise erhalten wir auch in diesem Fall die Beziehung (3.25).

Die Abbildung g^{-1} ist Lipschitz, also gibt es einen Punkt $u \in g^{-1}(2\varepsilon v)$, so daß gilt $\|u - u^0\| \leq 2L\varepsilon$. Wir bestimmen diejenige Zahl $\alpha > 0$, mit der αu in der Hyperebene H liegt, d.h. $(\alpha u - u^0)^T u^0 = 0$. Dies ist gerade $\alpha = (u^T u^0)^{-1}$. Wegen

$$(u - u^0)^T u^0 \leq \|u - u^0\| \leq 2L\varepsilon$$

(weil $\|u^0\| = 1$) erhalten wir:

$$1 - 2L\varepsilon \leq \|u^0\|^2 - (u^0 - u)^T u^0 = u^T u^0 \leq \|u\| \leq \|u^0\| + \|u - u^0\| \leq 1 + 2L\varepsilon \tag{3.26}$$

damit bekommen wir weiter

$$\frac{1}{1+2L\varepsilon} \leq \alpha \leq \frac{1}{1-2L\varepsilon} \qquad \text{und} \qquad |\alpha - 1| \leq \frac{2L\varepsilon}{1-2L\varepsilon} \tag{3.27}$$

und

$$\|\alpha u - u^0\| \leq \alpha \|u - u^0\| + |\alpha - 1| \, \|u^0\| \leq \frac{2L\varepsilon}{1-2L\varepsilon} + \frac{2L\varepsilon}{1-2L\varepsilon} = \frac{4L\varepsilon}{1-2L\varepsilon} < r \tag{3.28}$$

wegen (3.20). Wir sehen so: $\alpha u \in K_v$ und $g(\alpha u) = 2\alpha\varepsilon v$.

Die Definition von K_v impliziert, daß K_v kompakt ist. Deswegen existiert das Maximum $\max\limits_{u \in K_v} \|g(u)\|$ und

$$\max_{u \in K_v} \|g(u)\| \geq \|g(\alpha u)\| = 2\alpha\varepsilon \geq \frac{2\varepsilon}{1+2L\varepsilon} > \varepsilon \tag{3.29}$$

(wegen (3.20)). Wir betrachten den Punkt $u^* \in K_v$, wo das Maximum angenommen wird, und zeigen als nächstes $\|u^* - u^0\| < r$. Angenommen, das ist nicht wahr, d.h. $\|u^* - u^0\| = r$ und $u^* \in K_v \cap A$. Dann hätten wir $u^* \in K_v \cap A_\varepsilon$ (weil $\|g(u^*)\| > \varepsilon$). Andererseits ist $K_v \cap A_\varepsilon = \emptyset$ (siehe (3.25)). Also gilt wirklich $\|u^* - u^0\| < r$ und $d(u^*, A) > 0$.

Wenn wir mit T den Kegel $\{\, tA \mid t \geq 0\}$ bezeichnen, erhalten wir dadurch: $d(u^*, T) > 0$.

Jetzt sei β eine feste Zahl mit $1 < \beta < 1 + d(u^*, T)$. Wir studieren mögliche Urbilder u für den Wert $v^* = \beta g(u^*)$.

Wieder unterscheiden wir zwischen zwei Fällen:

1. $u \in \mathbb{R}^n \setminus \text{conv}\ T$. Dies ergibt $\|u - u^*\| > \beta - 1$, weil $u^* \in \text{conv}\ T$, $d(u^*, T) > \beta - 1$.

2. $u \in \text{conv}\ T$.
   Um dieselbe Abschätzung für $\|u - u^*\|$ zu bekommen, untersuchen wir die Projektion $p(u)$ des Punktes u auf die Gerade durch u^0 und den Nullpunkt. Es gilt:

   $$p\left(\frac{u}{\|p(u)\|}\right) = \frac{\|p(u)\| u^0}{\|p(u)\|} = u^0.$$

   Zusammen mit der Bedingung $u \in \text{conv}\ T$ erhalten wir

   $$\frac{u}{\|p(u)\|} \in \text{conv}\ A.$$

   Wegen

   $$g\left(\frac{u}{\|p(u)\|}\right) = \frac{\beta}{\|p(u)\|} g(u^*)$$

ergibt das

$$\frac{u}{\|p(u)\|} \in K_v$$

und

$$\left\| g\left(\frac{u}{\|p(u)\|}\right)\right\| = \frac{\beta}{\|p(u)\|}\|g(u^*)\| \leq \max_{u\in K_v}\|g(u)\| = \|g(u^*)\|.$$

Auf diese Weise haben wir gezeigt, daß $\|p(u)\| \geq \beta$ und

$$\|u - u^*\| \geq \|p(u) - p(u^*)\| \geq \|p(u)\| - \|p(u^*)\| \geq \beta - 1 \tag{3.30}$$

wie in Fall 1.

Schließlich erhalten wir für alle $u \in g^{-1}(v^*)$:

$$L \geq \frac{\|u - u^*\|}{\|\beta g(u^*) - g(u^*)\|} \geq \frac{\beta - 1}{(\beta - 1)\|g(u^*)\|} \geq \frac{1}{L_g\|u^* - u^0\|} \geq \frac{1}{L_g r} = 2L \tag{3.31}$$

wegen (3.18) und wir haben einen Widerspruch. #

Als eine wichtige Folgerung des Satzes 3.30 erhalten wir:

Korollar 3.31 *Es sei $f : I\!R^n \to I\!R^n$ lokal Lipschitz, pseudo-regulär in x^0. Weiterhin sei f richtungsdifferenzierbar und streng B-differenzierbar in x^0.*

Dann gilt: die Menge $Cf^{-1}((f(x^0), x^0); v)$ ist endlich $\forall v,\ \|v\| = 1$ und es liegt die Pseudo-Regularität vom Typ 4 vor.

Beweis: Es sei ein $v \in S$ beliebig gegeben. Angenommen, es gibt eine Folge $\{u^k\}$ mit paarweise verschiedenen Gliedern und $f'(x^0, u^k) = v$. Aus dem Satz 3.30 folgt: Es existiert das Minimum $\varepsilon := \min_{u\in S}\|f'(x^0, u)\| > 0$. Das ergibt

$$1 = \|v\| = \|f'(x^0, u^k)\| \geq \|u^k\|\varepsilon$$

also ist die Folge $\{u^k\}$ beschränkt und hat oBdA einen Limes u^*.

Wir wissen: Die Richtungsableitung $g(.) := f'(x^0, .)$ ist lokal Lipschitz und pseudo-regulär überall (also auch in u^*). Nach Lemma 3.4 ist g selber richtungsdifferenzierbar (wegen der strengen B-differenzierbarkeit von f). Damit erfüllt g alle Voraussetzungen von Satz 3.30. Das bedeutet wiederum, in einer gewissen Umgebung von u^* kann es nur endlich viele Lösungen von $g(u) = v$ geben, Widerspruch.
Insbesondere ist die Pseudo-Regularität von f vom Typ 4. #

Unter den Bedingungen von Satz 3.30 gibt es für jedes y hinreichend nahe an $f(x^0)$ in einer Umgebung von x^0 nur endlich viele Urbilder von y. Dies impliziert eine weitere Frage: Ist diese Anzahl der Urbilder durch eine universelle Konstante nach oben beschränkt?

Man könnte sich folgende hypothetische Situation vorstellen: Im Bildraum sei die Folge $\{y^k\}$, $y^k \to f(x^0)$ gegeben. Wir betrachten die Mengen $X^k := f^{-1}(y^k)$. Für $k \to \infty$ ist $\operatorname{diam} X^k \to 0$ und $\operatorname{card} X^k \to \infty$ bei Erhaltung der Abschätzung

$$d(x^0, X^k) \leq L\|f(x^0) - y^k\|.$$

Können wir diese Situation ausschließen?

Dies scheint im Moment noch ein offenes Problem zu sein. Obwohl man unter den Voraussetzungen von Satz 3.30 einen Abbildungsgrad auf einer geeigneten Umgebung von x^0 definieren kann und auch seine Endlichkeit bekommt, kennt der Autor im allgemeinen keine Möglichkeit, wie man an der Größe des Abbildungsgrades die Anzahl der Urbilder ablesen könnte.

Folgendes Beispiel illustriert das beschriebene Phänomen; es liegt aber keine Pseudo-Regularität vor (sonst hätten wir bereits ein Gegenbeispiel):

Beispiel 3.32 Wir definieren folgende Funktion $g : [0,1] \to I\!R$:

$$g(x) := \begin{cases} \frac{9}{2^{n+1}}x - \frac{9}{2^{2n+1}} & \text{für} \quad x \in I_n := \left[\frac{1}{2^n}, \frac{3}{2^{n+1}}\right], \ n \in I\!N \\ -\frac{9}{2^{n+1}}x + \frac{9}{2^{2n}} & \text{für} \quad x \in J_n := \left[\frac{3}{2^{n+1}}, \frac{1}{2^{n-1}}\right], \ n \in I\!N \\ 0 & \text{für} \quad x = 0 \end{cases}$$

Die Funktion g hat in jedem Punkt $a_n := \frac{1}{2^{n-1}}$ ein lokales Minimum mit dem Wert 0 und in jedem Punkt $b_n := \frac{3}{2^{n+1}}$ ein lokales Maximum mit dem Wert $\left(\frac{3}{2^{n+1}}\right)^2$. Wir sehen auch, daß mit wachsendem n die lokalen Maximalwerte immer kleiner werden.

Nun definieren wir die Funktion $f : [0,1] \to I\!R$ als $f(x) := g(x) + x^5$. Man überlegt sich leicht, daß die Funktion f richtungsdifferenzierbar ist. Wenn wir uns mit unseren Untersuchungen auf das Intervall $[0, 1/2]$ einschränken, bleiben die lokale Minima bzw. Maxima (jetzt mit anderen Extremalwerten) immer noch in den Punkten a_n bzw. b_n erhalten (da die Ableitungen der Funktion x^5 hinreichend klein sind). Man sieht leicht:

$$f\left(\frac{3}{2^{2n}}\right) > g\left(\frac{3}{2^{2n}}\right) = \frac{9}{2^{4n}} > \frac{8}{2^{5n-2}} = \left(\frac{1}{2^{n-1}}\right)^5 = f\left(\frac{1}{2^{n-1}}\right)$$

für $n > 2$.

Wir untersuchen die Anzahl der Urbilder zu den Werten

$$c_n := f\left(\frac{3}{2^{2n}}\right) = f(b_{2n-1}).$$

In allen lokalen Maximalpunkten b_k, $k < 2n - 1$ nimmt die Funktion f einen größeren Wert als c_n an, und in allen lokalen Minimalpunkten a_k mit $n \leq k \leq 2n - 1$ nimmt f

einen kleineren Wert als c_n an. Das bedeutet: Die Gerade $y = c_n$ hat einen nichtleeren Durchschnitt mit jedem der Intervalle $\mathrm{int} J_k, \; n \leq k \leq 2n-2$ und $\mathrm{int} I_k, \; n-1 \leq k \leq 2n-2$. Außerdem schneidet sie das Intervall J_{2n-1} in einem seiner Endpunkte. Damit gibt es mindestens $2n$ verschiedene Urbilder für den Wert c_n. Dabei bleibt jedoch die Anzahl der Urbilder für jeden Wert endlich. #

4 Einige Anwendungen

In diesem Kapitel möchten wir die Ergebnisse aus dem letzten Abschnitt auf einige Probleme anwenden. Als erstes untersuchen wir die Kojima-Funktion, deren Nullstellen die Karush-Kuhn-Tucker-Punkte (kurz KKT-Punkte) einer Optmierungsaufgabe beschreiben. In diesem Kontext vergleichen wir strenge Regularität mit der Pseudo-Regularität und formulieren einige Folgerungen. Insbesondere sind wir dann in der Lage (unter zusätzlichen Voraussetzungen) in einem Optimierungsproblem die Nebenbedingungen zu streichen, deren entsprechende Lagrange-Multiplikatoren Null sind, und das Regularitätsverhalten der Kojima-Funktion bleibt unverändert.

Unsere zweite Anwendung betrifft einen Newton-ähnlichen Algorithmus zum Lösen von nichtglatten Gleichungen. Dieses Verfahren funktioniert unter zwei zusätzlichen Voraussetzungen. Als eine Folgerung von Satz 3.30 können wir die erste Bedingung im richtungsdifferenzierbaren Fall erfüllen.

4.1 Regularität der Kojima-Funktion

Es sei ein Optimierungsproblem

$$(P) \qquad \min\{f(x) \,|\, g(x) \leq 0\}$$

gegeben, dabei ist $f \in C^{1.1}(\mathbb{R}^n, \mathbb{R})$, $g \in C^{1.1}(\mathbb{R}^n, \mathbb{R}^m)$. (d.h., die Ableitungen $Df(x)$ und $Dg(x)$ sind lokale Lipschitz-Funktionen). Wir definieren folgende Abbildung:

$$F(x,y) := \begin{pmatrix} Df(x) & + & \sum_{i=1}^m y_i^+ Dg_i(x) \\ g(x) & - & y^- \end{pmatrix}, \tag{4.1}$$

wobei $y_i^+ := \max\{0, y_i\}$ und $y_i^- := \min\{0, y_i\}$ sei.

Die Nullstellen von F nennen wir kritische Punkte zum Problem (P). Die Abbildung F wurde von M. Kojima in [20] eingeführt und ist unter dem Namen Kojima-Funktion bekannt. Sie ist eng mit folgender verallgemeinerten Gleichung verbunden:

$$H(x,y) \in N_C(x,y), \qquad (x,y) \in C,$$

wobei

$$H(x,y) = \begin{pmatrix} Df(x) + Dg(x)^T y \\ g(x) \end{pmatrix}$$

und $C = \mathbb{R}^n \times \mathbb{R}^m_+$ und

$$N_C(x,y) = \{(0,\eta) \in \mathbb{R}^n \times \mathbb{R}^m \,|\, \eta_i = 0 \text{ für } y_i > 0 \text{ und } \eta_i \leq 0 \text{ für } y_i = 0\}.$$

Die Lösungen dieser verallgemeinerten Gleichung werden KKT-Punkte zum Problem (P) genannt. Folgende einfache Beziehung ergibt sich zwischen den kritischen und den KKT-Punkten:

$$\begin{array}{llll} (x,y) & \text{KKT-Punkt} & \Rightarrow & (x,y+g(x)) & \text{kritischer Punkt,} \\ (x,y) & \text{kritischer Punkt} & \Rightarrow & (x,y^+) & \text{KKT-Punkt.} \end{array}$$

Wir definieren ein gestörtes Problem $P(a,b)$:

$$P(a,b): \qquad \min\{f(x) - a^T x \mid g(x) - b^T x \leq 0\}.$$

Man sieht leicht, daß die kritischen Punkte diese Problems den Lösungen von $F(x,y) = (a,b)^T$ entsprechen.

Wir sind an Aussagen interessiert, die charakteristisch für die strenge bzw. Pseudo-Regularität der Funktion F sind. Dazu schreiben wir F in folgender, leichter zu handhabender Gestalt:

$$F(x,y) = \begin{pmatrix} F_1(x,y) \\ F_2(x,y) \end{pmatrix} = \begin{pmatrix} Df(x) & Dg(x)^T & 0 \\ g(x) & 0 & -E \end{pmatrix} \begin{pmatrix} 1 \\ y^+ \\ y^- \end{pmatrix} = M(x)N(y).$$

Nun sei ein Punkt (x^0,y^0) gegeben, in dem wir F untersuchen. Wir brauchen die Mengen $TF((x^0,y^0);(u,v))$ und $CF((x^0,y^0);(u,v))$. Schreibt man jetzt die Kojima-Fuktion F in der Form aus Lemma 3.9, erhält man dadurch

$$TF((x^0,y^0);(u,v)) = TM(x^0,u)N(y^0) + M(x^0)TN(y^0,v), \tag{4.2}$$

(weil die Funktion $N(.)$ simpel ist, siehe [23]) und analog für Contingent-Ableitungen

$$CF((x^0,y^0);(u,v)) = CM(x^0,u)N(y^0) + M(x^0)CN(y^0,v), \tag{4.3}$$

da N richtungsdifferenzierbar ist. Dabei ist $TN(y^0,v)$ gegeben durch alle Vektoren $\{(0,\alpha,v-\alpha)\}$, die folgende Bedingungen erfüllen: $\alpha_i = r_i v_i$, $i = 1,\dots,m$ und

$$r_i = 1 \text{ für } y_i^0 > 0; \qquad r_i \in [0,1] \text{ für } y_i^0 = 0; \qquad r_i = 0 \text{ für } y_i^0 < 0,$$

und $CN(y^0,v)$ besteht aus allen Vektoren $\{(0,\alpha,v-\alpha)\}$ mit folgenden Bedingungen : $\alpha_i = r_i v_i$, $i = 1,\dots,m$ und

$$\begin{array}{l} r_i = 1 \;\text{ für }\; y_i^0 > 0 \;\text{ oder }\; y_i^0 = 0;\; v_i \geq 0, \\ r_i = 0 \;\text{ für }\; y_i^0 < 0 \;\text{ oder }\; y_i^0 = 0;\; v_i \leq 0. \end{array}$$

Wir definieren die Indexmengen:

$$I^+ := \{i \mid y_i^0 > 0\},\; I^0 := \{i \mid y_i^0 = 0\},\; I^- := \{i \mid y_i^0 < 0\}.$$

Unter Verwendung der Substitutionen

$Q_T(u) = TF_1(.,y^0)(x^0,u)$
$Q_C(u) = CF_1(.,y^0)(x^0,u)$

und

$$\alpha_i = r_i v_i, \qquad \beta_i = (1-r_i)v_i$$

und umgekehrt

$$v_i = \alpha_i + \beta_i, \quad r_i = \tfrac{\alpha_i}{v_i} \text{ für } v_i \neq 0, \quad r_i = 1 \text{ für } v_i = 0$$

bekommen wir folgende Aussage (Teil (i) siehe [19], [25], Teil (ii) folgt aus dem Satz 3.30):

Satz 4.1 *Es sei F die Kojima-Funktion zu dem Optimierungsproblem (P) mit $f \in C^{1.1}(I\!R^n, I\!R)$, $g \in C^{1.1}(I\!R^n, I\!R^m)$.*

(i) Die Funktion F ist streng regulär in $x^0 \Leftrightarrow$ folgendes System hat nur die triviale Lösung $(u,\alpha,\beta) = (0,0,0)$:

$$\begin{array}{lcl} Q_T(u) + Dg(x^0)^T\alpha & \ni & 0 \\ Dg(x^0)u - \beta & = & 0 \\ \alpha_i\beta_i \geq 0 \quad (i \in I^0) & & \\ \alpha_i = 0 \quad (i \in I^-) & & \\ \beta_i = 0 \quad (i \in I^+). & & \end{array}$$

(ii) Die Ableitungen Df und Dg seien zusätzlich richtungsdifferenzierbar, und F sei pseudo-regulär in x^0. Dann gilt:

(a) folgendes System hat nur die triviale Lösung $(u,\alpha,\beta) = (0,0,0)$:

$$\begin{array}{lcl} Q_C(u) + Dg(x^0)^T\alpha & \ni & 0 \\ Dg(x^0)u - \beta & = & 0 \\ \alpha_i\beta_i = 0,\ \beta_i \leq 0 \leq \alpha_i & & (i \in I^0) \\ \alpha_i = 0 \quad (i \in I^-) & & \\ \beta_i = 0 \quad (i \in I^+). & & \end{array}$$

(b) in einer geeigneten Umgebung U des Punktes (x^0,y^0) gibt es nur endlich viele kritische Punkte zum Problem (P)

(c) für (a,b) hinreichend nahe am Nullpunkt gibt es in U nur endlich viele kritische Punkte zum gestörten Problem $P(a,b)$. Insbesondere bleibt die konvexe Menge der entsprechenden Lagrange-Vektoren stets einelementig.

Beweis: Man sieht leicht, daß die (α,β) in (i) bzw. (ii) genau die Ableitungen $TN(y^0,v)$ bzw. $CN(y^0,v)$ sind. Der Rest bekommt man durch Anwendung des Satzes 3.11 und des Satzes 3.30. #

Die folgenden Überlegungen werden wir im Beweis des nächsten Lemmas brauchen:

Bemerkung 4.2

(i) Es seien $f : I\!R^n \to I\!R^n$ eine stetige Funktion, U eine offene, beschränkte Menge im R^n und $h : U \to I\!R^n$ eine stetige Auswahlfunktion von f^{-1} auf U. Dann gilt $f^{-1}(y) \cap h(U) = \{h(y)\}\ \forall y \in U$. (siehe [25]). Das bedeutet: f^{-1} ist lokal eindeutig und stetig in der Nähe eines Punktes $y \in U$.

(ii) Es sei $f : I\!R^n \to I\!R^n$ pseudo-regulär in x^0 und x^0 sei ein isoliertes Urbild von $y^0 := f(x^0)$. Dann hat die Multifunktion $H(y) := f^{-1}(y) \cap (x^0 + 2L\|y - y^0\| I\!B)$ (wobei L die pseudo-Lipschitz-Konstante ist) für y nahe an y^0 folgende Eigenschaften:
H hat kompakte Bilder, $H(f(x^0)) = \{x^0\}$ und H ist stetig (d.h. unter und oberhalb stetig) auf einer hinreichend kleinen Umgebung U von x^0. (siehe [25])

Aus der nächsten Aussage (siehe [25]) geht hervor, daß für die Beziehungen zwischen beiden Regularitätskonzepten die Lagrange-Multiplikatoren mit dem Wert 0 keine große Rolle spielen.

Lemma 4.3 *([25]):*
Es sei F die Kojima-Funktion zu dem Optimierungsproblem (P) ($f \in C^{1.1}(I\!R^n, I\!R)$, $g \in C^{1.1}(I\!R^n, I\!R^m)$) und F sei pseudo-regulär in einem kritischen Punkt (x^0, y^0) mit $y^0_m = g_m(x^0) = 0$. Wir definieren $F^{(m)}$ als die Kojima-Funktion zum Problem $(P^{(m)})$, das wir aus (P) durch Streichen der letzten Bedingung $g_m(x) \leq 0$ bekommen und $z^0 = (y^0_1, \dots, y^0_{m-1})$. Dann gilt:

(i) $(F^{(m)})$ ist pseudo-regulär in dem Punkt (x^0, z^0).

(ii) Zusätzlich sei (x^0, y^0) ein isolierte Nullstelle von F. Dann gilt:
F ist nicht streng regulär in $(x^0, y^0) \Rightarrow F^{(m)}$ ist nicht streng regulär in (x^0, z^0).

Beweis:
(i): Im Folgenden sei die Norm als die Maximum-Norm definiert. Die Funktion F ist pseudo-regulär in (x^0, y^0) und die entsprechenden Umgebungen aus Definition 1.3 seien gegeben durch $U = \delta I\!B^0$ und $V = (x^0, y^0) + \varepsilon I\!B^0$, wobei ε so klein gewählt wird, daß gilt:

$$\|x^1 - x^0\| \leq \varepsilon \Rightarrow g_m(x^1) < \frac{\delta}{3}. \tag{4.4}$$

Außerdem gibt es ein $r \in (0, \delta)$ mit der Eigenschaft, daß

$$\|x^2 - x^0\| \leq \varepsilon + Lr \Rightarrow g_m(x^2) < \frac{\delta}{2}. \tag{4.5}$$

Jetzt seien zwei Punkte $(a^1, b^1), (a^2, b^2) \in I\!R^{n+m-1}$ gegeben mit

$$\|(a^1, b^1)\|, \|(a^2, b^2)\| < \frac{r}{2}.$$

Wir fixieren einen Punkt $(x^1, y^1) \in I\!R^{n+m-1}$,

$$(x^1, y^1) \in (F^{(m)})^{-1}(a^1, b^1) \cap [(x^0, y^0) + \varepsilon I\!B^0].$$

Wir haben zu zeigen: Es existiert ein Punkt $(x^2, y^2) \in (F^{(m)})^{-1}(a^2, b^2)$ mit

$$\|(x^1, y^1) - (x^2, y^2)\| \le L\|(a^1, b^1) - (a^2, b^2)\| \tag{4.6}$$

für ein geeignetes $L > 0$. Nun definieren wir zwei Vektoren $(a^1, \beta^1), (a^2, \beta^2) \in I\!R^{n+m}$ dadurch, daß wir $\beta^1 := (b^1, \frac{\delta}{2})$ und $\beta^2 := (b^2, \frac{\delta}{2})$ setzen. Man sieht:

$$\|(a^1, \beta^1)\|,\ \|(a^2, \beta^2)\| < \delta.$$

Wegen $g_m(x^1) < \frac{\delta}{3} < \frac{\delta}{2}$ ist die m-te Ungleichung nicht aktiv in x^1. Deshalb gilt:

$$(x^1, \eta^1) := (x^1, y^1, g_m(x^1) - \tfrac{\delta}{2}) \in F^{-1}(a^1, \beta^1).$$

Die Pseudo-Regularität von F ergibt einen Punkt $(x^2, \eta^2) \in F^{-1}(a^2, \beta^2)$ mit

$$\|(x^1, \eta^1) - (x^2, \eta^2)\| \le L\|(a^1, \beta^1) - (a^2, \beta^2)\| = L\|(a^1, b^1) - (a^2, b^2)\| \le Lr.$$

Die Konstante r haben wir so gewählt (siehe (4.5)), daß $g_m(x^2) < \frac{\delta}{2}$. Deshalb wird die m-te Komponente von η^2 gerade

$$\eta^2_m = g_m(x^2) - \frac{\delta}{2} < 0.$$

Durch das Streichen der letzten Komponente bekommen wir so aus dem Vektor (x^2, η^2) einen Vektor $(x^2, y^2) \in (F^{(m)})^{-1}$, für den (4.6) gilt.

(ii) Wir benutzen die Multifunktion H aus Bemerkung 4.2(ii) und definieren folgende Abbildung: $M(a, b) := \arg\min\{y_m \,|\, (x, y) \in H(a, b)\}$. Angenommen, $M(a, b)$ ist ist einelementig auf einer Kugel $rI\!B$. Dann ist die Abbildung M (wegen der Eigenschaften von H) eine stetige Auswahlfunktion von F^{-1} und $M(0, 0) = \{(x^0, y^0)\}$. Die Bemerkung 4.2(i) impliziert die strenge Regularität von F in (x^0, y^0).

D.h., $M(a, b)$ muß für eine geeignete Folge $(a, b) \to (0, 0)$ mehrwertig sein. Damit haben wir zwei verschiedene Punkte (x^1, y^1), $(x^2, y^2) \in M(a, b)$, und es gilt

$$(u, v) := (x^1, y^2) - (x^2, y^2) \neq 0, \tag{4.7}$$

sowie $y^1_m = y^2_m$ (wegen der Definition von M). Außerdem erhalten wir $(u, v) \to (0, 0)$ und $y^1_m \to y^0_m = 0$.

Wir untersuchen die Wirkung von $(F^{(m)})^{-1}$ auf die folgenden zwei Punkte:

$p^1 := (a - (y^1_m)^+ Dg_m(x^1), \beta)$,
$p^2 := (a - (y^2_m)^+ Dg_m(x^2), \beta)$,

wobei $\beta = (b_1, \ldots, b_{m-1})$. Durch das Streichen der letzten Koordinate in den Vektoren (x^1, y^1), (x^2, y^2) definieren wir Punkte (x^1, η^1), (x^2, η^2) und es gilt: $(x^i, \eta^i) \in (F^{(m)})^{-1}(p^i)$, $i = 1, 2$. Wir wissen: $v_m = 0$, und das ergibt:

$$p^1 - p^2 = (y^2_m)^+ Dg_m(x^2) - (y^1_m)^+ Dg_m(x^1) = (y^1_m)^+ [Dg_m(x^1) - Dg_m(x^2)].$$

Jetzt nehmen wir an, $F^{(m)}$ sei streng regulär in (x^0, z^0). Mit einer Lipschitz- Konstante K von Dg_m und $(u, w) := (u, v_1, \ldots, v_{m-1})$ liefert das:

$$\|(u, w)\| \leq L\|p^1 - p^2\| = L(y^1_m)^+ \|Dg_m(x^2) - Dg_m(x^1)\| \leq LK(y^1_m)^+ \|u\|.$$

Wegen $y^1_m \to 0$ folgt daraus: $(u, w) = 0$ und mit $v_m = 0$ auch $(u, v) = 0$. Das ist ein Widerspruch zu (4.7). #

In [25] konnte Kummer folgenden Satz nur für stückweise glatte Funktionen formulieren (oder man benötigte explizit die Forderung, daß (x^0, y^0) ein isolierter kritischer Punkt ist). Nun können wir mit Hlife von Satz 3.30 die Aussage auf richtungsdifferenzierbare Lipschitz-Funktionen erweitern:

Satz 4.4 *Es sei $f \in C^{1.1}(I\!R^n, I\!R)$, $g \in C^{1.1}(I\!R^n, I\!R^m)$ und die Ableitungen Df, Dg seien richtungsdifferenzierbar. Weiter sei (x^0, y^0) ein kritischer Punkt des Problems (P) und $I^+ := \{i \,|\, y^0_i > 0\}$. Wir definieren das reduzierte Problem (P^r), indem wir in (P) alle Bedingungen streichen, die zu einem $i \notin I^+$ gehören.*

Wenn F pseudo-regulär, aber nicht streng regulär in (x^0, y^0) ist, so ist die Kojima-Funktion F^r von P^r ebenfalls pseudo-regulär, aber nicht streng regulär in dem Punkt (x^0, z^0). (den Punkt z^0 bekommt man aus y^0 durch Streichen der Komponenten y_i, $i \notin I^+$.)

Beweis: Die Funktion F erfüllt die Voraussetzungen von Satz 3.30 und damit ist (x^0, y^0) ein isolierter kritischer Punkt. Nun folgt aus Lemma 4.3, daß wir nacheinander alle Restriktionen mit $y^0_i = 0$ streichen können. Der Punkt (x^0, z^0) ist dann ein kritischer Punkt von (P^r) mit den behaupteten Eigenschaften. #

Wenn wir mit C^2- Funktionen arbeiten, erhalten wir auf diese Weise einen kritischen Punkt, in dem die strikte Komplementaritätsbedingung für (P^r) erfüllt ist. Das System $F^r = 0$ ist dann lokal eine C^1-Gleichung, und aus der Pseudo-Regularität von F folgt so die strenge Regularität. Diese Äquivalenz beider Regularitäten im C^2- Fall ist ein Ergebnis aus [11].

Beispiel 4.5 Wir konstruieren eine Funktion $h \in C^1(I\!R2, I\!R)$, so daß die Ableitung Dh stückweise linear und pseudo-regulär, aber nicht streng regulär (im Nullpunkt) ist.
Dazu nehmen wir uns die Funktion f aus dem Beispiel 3.10 her und setzen für einen Punkt $z = (x, y)$: $h(z) := \sqrt{2} f_1(z) f_2(z)$. Nun schreiben wir z in Polarkoordinaten

$z = r(\cos\phi, sin\phi)$ und betrachten die Kegel $C(k) := \{z \mid \phi \in [(k-1)\frac{\pi}{4}, k\frac{\pi}{4}]\},\ 1 \leq k \leq 8$. Dann bekommen wir folgende Darstellungen:

$$h(x,y) = \begin{cases} y(y-x) & \text{falls} \quad (x,y) \in C(1) \cup C(5) \\ x(y-x) & \text{falls} \quad (x,y) \in C(2) \cup C(6) \\ x(y+x) & \text{falls} \quad (x,y) \in C(3) \cup C(7) \\ -y(y+x) & \text{falls} \quad (x,y) \in C(4) \cup C(8) \end{cases}$$

$$Dh(x,y) = \begin{cases} (-y, 2y-x) & \text{falls} \quad (x,y) \in C(1) \cup C(5) \\ (-2x+y, x) & \text{falls} \quad (x,y) \in C(2) \cup C(6) \\ (2x+y, x) & \text{falls} \quad (x,y) \in C(3) \cup C(7) \\ (-y, -2y-x) & \text{falls} \quad (x,y) \in C(4) \cup C(8) \end{cases}$$

Mit einfachen Mitteln erhält man die Stetigkeit und die Pseudo-Regularität von Dh. Außerdem kann man zeigen, daß für jedes $a \in I\!R^2 \setminus \{0\}$ genau drei Lösungen z von $Dh(z) = a$ existieren. #

4.2 Ein Newton-ähnlicher Algorithmus für pseudo-reguläre Funktionen

In den letzten zehn Jahren wurden - beginned mit einer Arbeit von Kojima und Shindoh [21] zahlreiche Konzepte für Newton-Verfahren zum Lösen nichtglatter Gleichungen entwickelt. Der folgende Ansatz stammt aus [24]:
Die Gleichung $f(x+u) = 0$ wird mittels verallgemeinerter Richtungsableitungen approximiert. Als mögliche Ableitungskonzepte werden u.a. klassische Richtungsableitungen, Contingent-Ableitungen, Thibault-Mengen genutzt. Der Algorithmus funktioniert unter zwei Bedingungen: eine uniforme Injektivitätsbedingung in Bezug auf die benutzten Ableitungen und eine Approximationsbedingung. Mit Hilfe der Resultate aus Abschnitt 3 sind wir unter Pseudo-Regularität in der Lage, die Injektivitätsbedingung zu erfüllen, wenn klassische Richtungsableitungen als Approximationsansatz benutzt werden. Die zweite (ziemlich abstrakte) Bedingung kann man für viele der betrachteten Ableitungskonzepte auf eine einfachere Bedingung reduzieren, die allerdings recht stark auch in der Klasse aller richtungsdifferenzierbaren Lipschitz-Funktionen bleibt (in [33] als semismoothness bezeichnet). Wir formulieren die Konvergenzaussage für den Fall endlicher Dimension.

Es sei eine lokale Lipschitz-Funktion $f : I\!R^n \to I\!R^n$ gegeben. Mit x^* bezeichnen wir die Nullstelle von f, die wir finden möchten.

Weiter sei $F : I\!R^n \times I\!R^n \to I\!R^n$ eine Multifunktion, die jedem Paar (x, u) eine nichtleere Menge $F(x,u) \subset R^n$ zuordnet und $F(x, 0) = \{0\}\ \forall x \in I\!R^n$ erfülle. Wir studieren den iterativen Prozess ALG(α):

Zu x^k finde man eine Lösung u von

ALG (α) $$\left[f(x^k) + F(x^k, u)\right] \cap \alpha \|f(x^k)\| \mathbb{B} \neq \emptyset$$

und setze $x^{k+1} := x^k + u$.

Definition 4.6 *Das Tripel (f, x^*, F) heiße* **zulässig,** *wenn für jedes $\varepsilon > 0$ Konstanten $\alpha, r > 0$ existieren, so daß der Algorithmus ALG(α) eine unendliche Folge $\{x^k\}$ generiert mit*

$$\|x^{k+1} - x^*\| \leq \varepsilon \|x^k - x^*\| \qquad \textit{für} \quad \|x^0 - x^*\| \leq r.$$

Die Definition beinhaltet u.a., daß jedes Unterproblem lösbar ist.

Satz 4.7 *([24])*
Das Tripel (f, x^*, F) ist zulässig, wenn eine Umgebung Ω von x^*, eine Konstante $c > 0$ und eine Funktion $o(.)$ existieren, so daß gilt:

$\forall x \in \Omega \quad \forall u \in \mathbb{R}^n$:

$$\begin{array}{ll} (CI) & c\|u\| \leq \inf\{\|v\| \mid v \in F(x, u)\} \\ (CA) & f(x) + F(x, u) \subset F(x, u + x - x^*) + o(\|x - x^*\|)\mathbb{B}. \end{array}$$

Nun betrachten wir eine richtungsdifferenzierbare lokale Lipschitz-Funktion $f : \mathbb{R}^n \to \mathbb{R}^n$ und setzen $F(x, u) := f'(x, u)$.
Beide Bedingungen zusammen implizieren, daß x^* eine isolierte Nullstelle von f ist. Letzteres wird nach Satz 3.30 durch die Pseudo-Regularität von f in x^* gesichert. Außerdem bekommt man zugleich $f'(x, u) \neq 0$ für alle $u \neq 0$ und $x \in \Omega$, denn die Pseudo-Regularität überträgt sich von x^* auf eine Umgebung. Damit wissen wir, daß für jedes $x \in \Omega$ die Zahl

$$c(x) := \min_{u \in S} \|f'(x, u)\| \tag{4.8}$$

positiv ist. Allerdings verlangt die Bedingung (CI) eine positive untere Schranke c für alle $c(x)$ in $x \in \Omega$.
Die zweite Bedingung kann im vorliegenden Fall $F(x, u) = f'(x, u)$, durch eine vereinfachte Forderung (CA*) äquivalent ersetzt werden (siehe[24]). Die Bedingung (CA^*) verlangt eine $o(.)$-Approximation von f durch die Richtungsableitung in x auf allen Strahlen, welche die (unbekannte) Lösung x^* mit x verbinden. Sie sichert, daß die Richtung $u = x^* - x$ die Hilfsaufgabe löst, aber sagt nichts über das lokale Verhalten von f in anderen Richtungen.

Dadurch erhalten wir insgesamt als eine hinreichende Bedingung:

Korollar 4.8 *Es sei $f : \mathbb{R}^n \to \mathbb{R}^n$ eine lokale Lipschitz-Funktion, $f(x^*) = 0$. Zusätzlich sei f richtungsdifferenzierbar und pseudo-regulär in x^*. Dann ist das Tripel (f, x^*, f') zulässig, wenn eine Umgebung Ω von x^* und eine Funktion $o(.)$ existieren, so daß gilt:*

$\forall x \in \Omega \quad \forall u \in X$:

$$(CA^*) \qquad\qquad f(x) + f'(x, x^* - x) \subset o(\|x - x^*\|)\mathbb{B}.$$

Beweis: Wir zeigen, daß die Bedingung (CI) aus dem Satz 4.7 erfüllt ist. Angenommen, die positive untere Schranke aus (CI) existiert nicht, d.h., es existiert eine Folge $\{x^k\}$, $x^k \to x^*$ und eine Folge $\{u^k\}$, $u^k \in S$, so daß gilt:

$$\|f'(x^k, u^k)\| = \min_{u \in S} \|f'(x^k, u)\| = c(x^k) \to 0 \qquad (x^k \to x^*).$$

OBdA können wir annehmen, daß die Folge $\{u^k\}$ gegen ein geeignetes $u \in S$ konvergiert. Da die Richtungsableitung $f'(x, .)$ lokal Lipschitz mit der Konstante L_f für jedes x ist (wenn L_f die Lipschitz-Konstante von f bezeichnet), konvergiert auch die Folge $\|f'(x^k, u)\|$ gegen Null. Nun möchten wir die Funktion f mit Hilfe von Satz 1.7 so stören, daß wir auf diese Weise eine im Punkt x^k pseudo-reguläre Funktion mit der Richtungsableitung Null erhalten. Das wäre ein Widerspruch zu Satz 3.30.

Man kann OBdA annehmen, daß die Funktion f pseudo-regulär in x^* mit einer geeigneten Konstante L und den entsprechenden Umgebungen $V = x^* + \delta\mathbb{B}^0$ und $U = \delta\mathbb{B}^0$, $\delta < 1$ ist. Das bedeutet, f ist pseudo-regulär in allen Punkten x, falls $x \in V$ und $f(x) \in U$ gilt. Wir können auch die entsprechenden Umgebungen $V(x)$ und $U(x)$ bestimmen:
$V(x) = x + (\delta - \|x - x^*\|)\mathbb{B}^0$ und $U(x) = f(x) + (\delta - L_f\|x - x^*\|)\mathbb{B}^0$
Nun definieren wir

$$\delta(x) := \min\{\delta - \|x - x^*\|, \delta - L_f\|x - x^*\|\}$$

und finden einen Punkt $x^k \in x^* + \min\{\delta, \frac{\delta}{L_f}\}\mathbb{B}^0$ (d.h. $x^k \in V$ und $f(x^k) \in U$) mit

$$\|f'(x^k, u)\| < \frac{\delta(x^k)}{3(L+1)}. \tag{4.9}$$

Das ist möglich dank der Gestalt von $\delta(.)$ und der Tatsache, daß $\|f'(x^k, u)\|$ gegen Null konvergiert. Insbesondere ist f pseudo-regulär in x^k mit den entsprechenden Umgebungen $\bar{V} = x^k + \delta(x^k)\mathbb{B}^0$ und $\bar{U} = f(x^k) + \delta(x^k)\mathbb{B}^0$.

Die Funktion $g : \mathbb{R}^n \to \mathbb{R}^n$ sei erklärt als

$$g(x) := -\|x - x^k\| f'(x^k, u).$$

Die Lipschitz-Konstante von g ist $\|f'(x^k, u)\|$ und es gilt:

$$\sup_{x \in \bar{V}} \|g(x)\| \leq \delta(x^k)\|f'(x^k, u)\| < \|f'(x^k, u)\|.$$

Dadurch erhalten wir

$$|g| = \max\{\sup_{x\in\bar{V}} \|g(x)\|, L_g\} \leq \|f'(x^k, u)\| < \frac{\delta(x^k)}{3(L+1)}. \tag{4.10}$$

Damit haben wir alle Voraussetzungen von Satz 1.7 erfüllt. Als Ergebnis bekommt man, daß die Funktion $(f+g)$ pseudo-regulär im Punkt x^k ist.

Andererseits gilt für die Richtungsableitung $(f+g)'(x^k, u)$:

$$(f+g)'(x^k, u) = f'(x^k, u) + g'(x^k, u) = f'(x^k, u) - f'(x^k, u) = 0,$$

was im Widerspruch zum Satz 3.30 steht. #

Im Folgenden beschreiben wir typische Situationen, die bei den meisten Anwendungen auftreten.

(i) Die Funktion f ist stückweise glatt. Dann ist die Bedingung (CA^*) automatisch erfüllt (wir haben zur Verfügung die Approximation für alle Auswahlfunktionen, die in x^* wesentlich aktiv sind).

(ii) Die Funktion f ist mit Hilfe von sog. NCP-Funktionen erklärt, die zur Beschreibung von nichtlinearen Komplementaritätsproblemen benutzt werden.

NCP-Funktionen sind gewöhnlich definiert als positiv homogene Funktionale auf $I\!R^2$, die genau die beiden nichtnegativen Koordinatenachsen als Nullstellen besitzen und auf $I\!R^2 \setminus \{0\}$ stetig differenzierbar sind. Insbesondere ist g damit überall richtungsdifferenzierbar.
Es sei $g : I\!R^2 \to I\!R$ so ein Funktional, dann läßt sich ein Komplementaritätsproblem

$$u_i(x) \geq 0, \qquad x_i \geq 0, \qquad u_i(x)x_i = 0 \qquad \forall i = 1, \dots, n, \qquad u \in C^1(I\!R^n, I\!R^n)$$

in der Form

$$f_i(x) := g(u_i(x), x_i) = 0 \qquad \forall i = 1, \dots, n$$

schreiben. Man kann sich leicht überlegen, daß die Bedingung (CA^*) in diesem Fall erfüllt ist.

(iii) Point Based Approximations (dazu siehe [38]).
In diesem Fall wird die Existenz einer Konstante $K > 0$ gefordert, so daß gilt:

$$\|f(x+u) - f(x) - f'(x, u)\| \leq \frac{K}{2}\|u\|^2.$$

Wenn man $u := x^* - x$ setzt, erhält man die Gültigkeit von (CA^*).

Literatur

[1] J.-P. Aubin: "Contingent derivatives of set-valued maps and existence of solutions to nonlinear inclusions and differential inclusions", in Mathematical Analysis and Applications, Part A; ed. L. Nachbin, Advances in Mathematics:Supplementary Studies, 7A, Academic Press, New York, 1981, S. 160-232

[2] J.-P. Aubin: "Lipschitz behavior of solutions to convex minimization problems",*Math. Oper. Res.* **9**(1984) 87-111

[3] J.-P. Aubin, I. Ekeland: "Applied nonlinear analysis",Wiley, New York, 1984

[4] J.-P. Aubin, H. Frankowska: "Set-Valued Analysis", Birkhäuser, Boston, (1990)

[5] J.-P. Aubin: "Viability Theory",Birkhäuser, Boston, (1991)

[6] J.M. Borwein, D.M. Zhuang: "Verifiable necessary and sufficient conditions for regularity of set-valued and single-valued maps", *Journal Math. Anal. Appl.* **134**(1988) 441-459

[7] F.H. Clarke: "Optimization and Nonsmooth Analysis", Wiley, New York, 1983

[8] R. Cominetti: "Metric regularity, tangent sets, and second-order optimality conditions",*Appl. Math. Optim.***21** (1990) 265-287

[9] A. Dmitruk, A.A. Miliutin, N. Osmolovskii: "Lyusterniks theorem and the theory of extrema", *Russian Mathematical Surveys* **35**(1981) 11-51

[10] A.L. Dontchev, W.W. Hager: "Implicit functions, Lipschitz maps and stability in optimization", *Math. of Operations Research* **19**(1994) 753-768

[11] A. L. Dontchev, R. T. Rockafellar: "Characterizations of strong regularity for variationhal inequalities over polyhedral convex sets", *SIAM Journal Optimization* **6**(1996) 1087-1105

[12] I. Ekeland: "On the variational principle", *Journal Math. Anal. and Appl.* **47** (1974) 324-353

[13] R. Engelking, K. Sieklucky: "Topology. A geometric approach", Heldermann Verlag, Berlin, 1992

[14] H. Federer: "Geometric Measure Theory", Springer, 1969

[15] L. M. Graves: "Some mapping theorems", *Duke Math. Journal* **17**(1950) 111-114

[16] R. Henrion: "The approximate subdifferential and parametric optimization", Habilitationsschrift, 1998, Humboldt-Universität zu Berlin

[17] A.D. Ioffe: "Nonsmooth analysis: differentiable calculus of nondifferentiable mappings", *Trans. Amer. Math. Soc.* **266** (1981) 1-56

[18] A.D. Ioffe: "On the local surjection property", *Nonlinear Analysis: Theory, Methods and Applications* **11** (1987) 565-592

[19] D. Klatte, B. Kummer: "Generalized Kojima-functions and Lipschitz stability of critical points", 1997

[20] M. Kojima: "Strongly stable solutions in nonlinear programs", in: S.M. Robinson (ed.), Analysis and Computation of Fixed Points, Academic Press, New York, 1980, S. 93-138

[21] M. Kojima, S. Shindoh: "Extensions of Newton and quasi-Newton methods to systems of PC1-equations",*Journal of OR Soc. of Japan* **29** (1987) 352-374

[22] D. Kuhn, R. Löwen: "Piecewise affine bijections of $I\!R^n$, and the equation $Sx^+ - Tx^- = y$", *Linear Algebra and its Applications* **96** (1987) 109-129

[23] B. Kummer: "Lipschitzian inverse functions, directional derivatives and application in $C^{1.1}$-optimization",*Journal of Opt. Theory and Appl.* **158** (1991) 35-46

[24] B. Kummer: "Newton's method based on generalized derivatives for nonsmooth functions: convergence analysis", in: W. Oettli und D. Pallaschke, eds., Advances in Optimization, Springer, 1992, S. 171-194

[25] B. Kummer: "Lipschitzian and pseudo-Lipschitzian inverse functions and applications to nonlinear optimization", in "Lecture notes in pure and applied mathematics", Vol. 195(1997), Math. programming with data perturbations, ed. A. V. Fiacco, 201-222

[26] B. Kummer: "Metric regularity: characterizations, nonsmooth variations and successive approximation", erscheint in Optimization

[27] L. Kuntz, S. Scholtes: "Structural analysis of nonsmooth mappings, inverse functions, and metric projections", *Journal of Math. Anal. and Appl.* **188** (1994) 346-386

[28] B.S. Mordukhovich: "Approximation methods in problems of optimization and control", Nauka, Moskau, 1988

[29] B.S. Mordukhovich: "Sensitivity analysis in nonnsmooth optimization, in Theoretical Aspects of Industrial Design, ed. D.A. Field und V. Komkov, *SIAM Proceedings in Applied Mathematics* **58**(1992) 32-46

[30] B.S. Mordukhovich: "Complete characterization of openness, metric regularity and Lipschitzian properties of multifunctions", *Transactions of American Mathematical Society* **340**(1993) 1-36

[31] B.S. Mordukhovich: "Lipschitzian stability of constraint systems and generalized equations", *Nonlinear Analysis: Theory, Methods and Applications* **22**(1994) 173-206

[32] J.-P. Penot: "Metric regularity, openness and Lipschitz behavior of maps", *Nonlinear Analysis* **13**(1989) 629-643

[33] L. Qi, J. Sun: "A nonsmooth version of Newton's method", *Math. Programming* **58** (1993) 353-367

[34] D. Ralph, S.Scholtes: "Sensitivity analysis of composite piecewise smooth equations", *Math. Programming* **76** (1997) 593-612

[35] S. M. Robinson: "Stability theory for systems of inequalities, Part II: differentiable nonlinear systems", *SIAM Journal Numer. Anal.* **13**(1976) 497-513

[36] S.M. Robinson: "Strongly regular generalized equations", *Mathematics of Operations Research* **5**(1980) 43-62

[37] S.M. Robinson: "Local structure of feasible sets in nonlinear programming III: Stability and sensitivity", *Math. Prog. Study* **30** (1987) 109-129

[38] S.M. Robinson: "Newton's method for a class of nonsmooth functions", Working paper, (1988), Univ. of Wisconsin-Madison, Department of Industrial Engineering, Madison, WI 53706

[39] S.M. Robinson: "An implicit-function theorem for a class of nonsmooth functions", *Math. of Operations Research* **16** (1991) 292-309

[40] R.T. Rockafellar: "Lipschitzian properties of multifunctions",*Nonlin. anal.: Theory, Methods and Applications* **8** (1985) 867-885

[41] R.T. Rockafellar, R.J.-B. Wets: "Variational Analysis", Springer, 1998

[42] S. Scholtes: "Introduction to piecewise differentiable equations", Habilitationsschrift, Preprint No. 53/1994, Institut für Statistik und Mathematische Wirtschaftstheorie, Universität Karlsruhe, 1994

[43] A. Shapiro: "Perturbation theory of nonlinear programs when the set of optimal solutions is not a singleton", *Applied Math. and Optimization* **18** (1988) 215-229

[44] A. Shapiro: "Perturbation analysis of optimization problems in Banach spaces", *Numer. Funct. Anal. and Optimiz.* **13** (1992) 97-116

[45] L. Thibault: "On generalized differentials and subdifferentials of Lipschitz vector-valued functions", *Nonlin. Anal.: Theory, Methods and Applications* **6** (1982) 1037-1053

[46] C. Ursescu: "Multifunctions with closed convex graphs", *Czech. Math. J.* **25** (1975) 438-441

Erklärung

Ich erkläre, daß ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Literatur angefertigt habe.

Berlin, den 07.12.1998 Peter Fúsek

Tabellarischer Lebenslauf

Angaben zur Person:

Name:	Peter Fúsek
Geburtsdatum:	28. November 1970
Geburtsort:	Kežmarok, Slowakei
Staatsangehörigkeit:	Slowakische Republik
Familienstand:	ledig

Schulausbildung:

1977 - 1985	Grundschule in Poprad
1985 - 1988	Gymnasium in Poprad
1988 - 1989	Gymnasium in Banská Štiavnica

Hochschulausbildung:

1989 - 1994	Humboldt-Universität zu Berlin Studiengang Mathematik-Diplom Spezialisierung: Mathematische Optimierung Nebenfach: Wirtschaftswissenschaften Diplomarbeit: "Über Kettenregeln in Gleichungsform für Ableitungen nichtglatter Funktionen"
1995 -	Promotionsstudium am Institut für Mathematik der Humboldt-Universität zu Berlin
Juli 1995 - Juni 1997	Promotionsstipendium gem. NaFöG
Okt. 1997 -	Stipendiat des Graduiertenkollegs 'Geometrie und nichtlineare Analysis' der HUB

Berlin, den 07.12.1998