Über einparametrische Optimierungsprobleme (spezielle Einbettungen) und einparametrische Variationsungleichungen.

Digitale Dissertation

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
Lic. Walter Gómez Bofill
geboren an: 26 Oktober 1969
in : Havanna.

Präsident der Humboldt-Universität zu Berlin

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät II

Gutachter

1. Prof. Dr. J. Guddat (Humboldt Universität zu Berlin)
2. Prof. Dr. H. T. Jongen (Rheinisch-Westfälische Technische Hochschule Aachen)
3. Prof. Dr. D. Pallaschke (Universität Karlsruhe)
4. Prof. Dr. R. Tichatschke (Universität Trier)

Eingereicht am: 02.09.1998

Tag der mündlichen Prüfung: 28.01.1999

Hiermit möchte ich mich herzlich bei meinem Betreuer Prof. Dr. J. Guddat und bei Prof. Dr. H. Th. Jongen für die fachlichen Diskussionen und nützlichen Hinweise bedanken.

Inhalt

1 Einleitung **1**

2 Allgemeine Optimierungsprobleme. **5**
2.1 Grundlegende Definitionen und Optimalitätsbedingungen. . . 5
2.2 Transversalität, Satz von Sard und Generizität 10

3 Einparametrische Optimierungsprobleme. **14**
3.1 Nichtdegenerierte kritische Punkte. 14
3.2 Die Klasse von Jongen, Jonker und Twilt. 18

4 Anwendung bei Optimierungsverfahren. **39**
4.1 Eine innere Einbettung. 39
4.2 Generische quadratische Penalty-Einbettungen. 57

5 Regularität einparametrischer Variationsungleichungen. **75**
5.1 Variationsungleichungen. 75
5.2 Einparametrische Variationsungleichungen. 78
5.3 Singularitäten und lokale Eigenschaften 81
5.4 Generizität reguläre Variationsungleichungen 98
5.5 Anwendung bei der Konvergenz von Kurvenverfolgungsprozeduren 122
5.6 Weitere Entwicklung 136

1 Einleitung

In dieser Arbeit beschäftigen wir uns mit zwei parametrischen Aufgaben: Optimierungsprobleme und Variationsungleichungen. Sei eine Menge

$$M = \{x \in I\!R^n | h_i(x) = 0, i \in I, g_j(x) \geq 0, j \in J\}$$

gegeben, wobei I und J endliche Indexmengen darstellen.

Optimierungsprobleme werden mittels einer Zielfunktion $f : I\!R^n \to I\!R$ als die Suche von Punkten $x \in M$ mit

$$f(y) \geq f(x), \quad \forall y \in M$$

definiert.

Die Variationsungleichungen werden mittels einer Abbildung $F : I\!R^n \to I\!R^n$ als die Suche von Punkten $x \in M$ mit

$$F^T(x)(y - x) \geq 0, \quad \forall y \in M$$

definiert.

Beide Probleme werden parametrisch betrachtet, falls die jeweiligen Abbildungen von einen zusätzlichen eindimensionalen reellen Parameter abhängen, d.h.

$$(f(x,t), h_i(x,t), i \in I, g_j(x,t), j \in J)$$

und

$$(F(x,t), h_i(x,t), i \in I, g_j(x,t), j \in J).$$

Wir setzen höchstens dreimal stetige Differenzierbarkeit der auftretenden Funktionen voraus.

Wir betrachten beide parametrischen Aufgaben unter der Regularität im Sinne von Jongen, Jonker und Twilt. Diese Regularität wurde ursprünglich für einparametrische Optimierungsprobleme als die Klasse $\mathcal{F}$ in [35] definiert. Wir formulieren im folgenden die Ziele der Arbeit:

- Die Untersuchung der wichtigsten Eigenschaften der Regularität im Sinne von Jongen, Jonker und Twilt für konkrete Einbettungen (eine innere Punkt-Einbettung und eine Penalty-Einbettung).

- Die Erweiterung dieses nützlichen Werkzeuges auf den Fall der einparametrischen Variationsungleichungen. Dazu gehört die Umformulierung der Singularitäten und die Beweise der wichtigsten Eigenschaften.

Das zweite Kapitel enthält grundlegende Definitionen und wohlbekannten Optimalitätsbedingungen. Anschließend werden wichtige Werkzeuge (z.B. der parametrisierte Satz von Sard, die starke Whitney Topologie) für diese Arbeit dargestellt.

Im dritten Kapitel wird die Regularität in Sinne von Jongen, Jonker und Twilt präsentiert (5 Typen). In einem ersten Abschnitt wird der Begriff des nichtdegenerierten kritische Punktes eingeführt (Typ 1). Im Abschnitt 3.2 werden 4 Basisentartungen (Singularitäten) definiert. So erhält man die Klasse $\mathcal{F}$ von Jongen Jonker und Twilt (siehe [34], [35], siehe auch [14]). Es werden die wichtigsten Eigenschaften dieser Regularität erläutert und Ergebnisse, die auf der Voraussetzung dieser Regularität basieren, vorgestellt.

Das vierte Kapitel beschäftigt sich mit zwei speziellen Einbettungen. Im ersten Abschnitt wird folgende innere Punkt-Einbettung untersucht:

$$P^{int}_{(x^0,q)}(f,G,t) \quad \min \left\{ f^{int}(f,x,t) \mid g^{int}_j(g,x,t) \geq 0, j \in J \right\}, \ t \in [0,1]$$

wobei

$$\begin{aligned} f^{int}(f,x,t) &= tf(x) + (1-t)\|x - x^0\|^2, \\ g^{int}_j(g,x,t) &= g_j(x) - (1-t)q_j \ . \end{aligned}$$

Offensichtlich ist $P^{int}_{(x^0,q)}(f,G,1)$ äquivalent zu

$$\min \ \{f(x) \mid g_i(x) \geq 0, j \in J\}$$

Wir setzen voraus, daß für $P^{int}_{(x^0,q)}(f,G,0)$ ein Punkt x^0 mit $g_j(x^0) > 0$, $j \in J$ bekannt ist (es gilt auch $g_j(x^0) > q_j > 0$). Wir sind uns bewußt, daß diese Voraussetzung für nichtkonvexe Probleme sehr einschänkend ist.

Die Erreichbarkeit des Parameterwertes $t = 1$ bei der Benutzung von Kurvenverfolgungsverfahren mit Sprüngen wird studiert. Durch die Unterscheidung der verschiedenen Richtungen in t, in denen ein Rückkehrpunkt auftreten kann, findet man besonders günstige Eigenschaften dieser Einbettung. Außerdem wird für diese Einbettung die Voraussetzung der Regularität erörtert und ein Rechtfertigungssatz bewiesen. Am Ende dieses Abschnittes wird eine in [19] vorgestellte Homotopie, die Ähnlichkeiten mit der in diesem Abschnitt untersuchten Einbettung besitzt, betrachtet. Insbesondere werden die in [19] vorgeschlagenen Bedingung für die Konvergez dieser Homotopie

einer Kritik unterzogen und dabei die Vorteile der Annahme der Regularität hervorgehoben.

Der zweite Abschnitt beschäftigt sich mit einer Penalty-Einbettung. Diese Einbettung wurde in [11] vorgestellt und untersucht. Die Rechtfertigung der Regularitätsvoraussetzung, die in diesem Abschnitt behandelt wird, wurde in [11] nicht betrachtet. Diese Fragestellung und ihre Antwort ist ein weiterer Beitrag zur Untersuchung dieser Penalty-Einbettung. Beim Beweis wird die von Rückmann und Tammer in [61] entwickelte Technik benutzt, wobei an mehreren Stellen neue Argumente erforderlich sind. Das wird dadurch verursacht, daß einerseits, die für die Störung verfügbare Parametermenge kleiner ist, als in [61], andererseits, die Fälle $t = 0$ und $t = 1$ zusätzlich behandelt werden mußten.

Das fünfte Kapitel ist Variationsungleichungen gewidmet. Im ersten Abschnitt werden die grundlegenden Definitionen und Eigenschaften, ohne Betrachtung von Parametern, vorgestellt. Anschnitt 5.2 dient der Vorstellung des einparametrisches Falles und der dritte Abschnitt der Neubeschreibung der 5 Typen für den Fall der einparametrischen Variationsungleichungen. Die Unterschiede zu den Typen bei Optimierungsproblemen ergeben sich aus der Tatsache, daß die Symmetrie der Hesse-Matrix der Lagrangefunktion bezüglich x verlorengeht. Das erfordert die sorfältige Umformulierung einiger Bedingungen, außerdem mußte bei den Beweisen der lokalen Eigenschaften (geometrische Struktur, Änderung der charakteristischen Größen) diese neue Situation berücksichtigt werden. Im Abschnitt 5.3 kommt man zum generischen Charakter dieser Regularität. Dabei wird diese Generizität insbesondere bezüglich der Offenheit ausführlich untersucht. Aussagen über die Existenz verbindender Kurven zwischen $t = 0$ und $t = 1$ unter der Voraussetzung der Regularität werden in letzten Abschnitt bewiesen.

Die gesamte Betrachtung der Regularität im Sinne vom Jongen, Jonker und Twilt für einparametrische Variationsungleichungen und die daraus resultierenden Schlußfolgerungen liefern einen nützlichen Beitrag zur Untersuchung von Variationsungleichungen.

Der Abschnitt 5.6 ist ein Ausblick auf weitere Untersuchungen.

2 Allgemeine Optimierungsprobleme.

2.1 Grundlegende Definitionen und Optimalitätsbedingungen.

Wir definieren in diesem Kapitel die allgemeinen Optimierungsprobleme und mehrere Begriffe, die sich darauf beziehen. Die Bedeutung der Optimierungsprobleme für die angewandte Mathematik ist längst erkannt worden. Optimierungsprobleme treten sehr häufig in den technischen und auch in anderen Bereichen der Wissenschaft auf.

Ein Optimierungsproblem wird durch zwei Objekte gekennzeichnet: Die Zielfunktion und die zulässige Menge.

Definition 1
Es seien eine abgeschlossene Menge $X \subset \mathbb{R}^n$ und eine Funktion $f : X \to \mathbb{R}$ gegeben. Das darin bestehende Problem, einen Punkt $x \in X$ zu finden, so daß $\forall y \in X$ gilt $f(x) \leq f(y)$, wird Optimierungsproblem definiert durch (f, X) genannt. Als Notation werden wir $P(f, X)$ benutzen.

Definition 2
Eine lokale Lösung (lokaler Minimalpunkt) des Optimierungsproblems $P(f, X)$ ist ein zulässiger Punkt x (das heißt $x \in X$), so daß eine offene Menge $U \subset \mathbb{R}^n$ existiert mit der folgenden Eigenschaft: für alle $y \in X \cap U$ gilt die Ungleichung $f(x) \leq f(y)$. Falls diese Ungleichung strikt in der Umgebung gilt, wird von einer strikten Lösung (oder strikter lokaler Minimalpunkt) gesprochen. Kann der ganze Raum $\mathbb{R}^n$ als Umgebung genommen werden, dann haben wir einen globalen Minimalpunkt.

Bedingungen für die Existenz eines globalen Minimalpunktes sind zum Beispiel durch den Satz von Weierstraß gegeben, daß f unterhalbstetig und X kompakt ist. Für die Entwicklung von Lösungsverfahren und die Beschreibung von Optimalitätsbedingungen ist von großem Nutzen, wenn die Menge X eine explizite Beschreibung besitzt. Wir werden in dieser Arbeit annehmen, daß X die Gestalt

$$X = \{x \in \mathbb{R}^n | h_i(x) = 0, i \in I, g_j(x) \geq 0, j \in J\}$$

hat, wobei $I = \{1, \ldots, m\}$ und $J = \{1, \ldots, p\}$ endliche Indexmengen sind. Bei dieser Beschreibung der zuläsigen Menge werden wir das Optimierungsproblem auch mit der Notation $P(f, H, G)$ versehen, wobei H und G Abbildungen zwischen den entsprechenden Räumen sind.

Bei verschiedenen Anwendungen und Lösungsansätzen werden unterschiedliche Differenzierbarkeitsgrade für die definierenden Funktionen des Optimierungsproblems $P(f, H, G)$ verlangt.Wir werden normalerweise mit Funktionen aus $C^2(\mathbb{R}^n, \mathbb{R})$ und $C^3(\mathbb{R}^n, \mathbb{R})$ arbeiten.

Gegeben sei ein $\overline{x} \in X$. Wir definieren die aktive Indexmenge in der Form:

$$J_0(\overline{x}) = \{j \in J | g_j(\overline{x}) = 0\}.$$

Ungleichungsrestriktionen mit Indizes außerhalb von $J_0(\overline{x})$ spielen in einer Umgebung von $\overline{x}$ keine Rolle und sind deswegen bei der Beschreibung von Optimalitätsbedingungen lokaler Natur nicht zu berücksichtigen.

Der folgende Satz (siehe [29]) gibt gut bekannte notwendige Optimalitätsbedingungen im Fall differenzierbarer Daten (f, H, G) an.

Satz 1 *(Fritz John)*
Es sei $\overline{x} \in X$ eine lokaler Minimalpunkt des Optimierungsproblems $P(f, H, G)$. Dann existiert eine Tupel von Zahlen $(\lambda_0, \lambda, \mu)$ (genannt Multiplikatoren), die nicht alle Null sind, wobei $(\lambda, \mu) = (\lambda_i, \mu_j)$ mit $i \in I$ und $j \in J_0(\overline{x})$, so daß λ_0 und μ_j, $j \in J_0(\overline{x})$ nicht negativ sind, und die Gleichung gilt

$$\lambda_0 Df(\overline{x}) - \sum_{i \in I} \lambda_i Dh_i(\overline{x}) - \sum_{j \in J_0(\overline{x})} \mu_j Dg_j(\overline{x}) = 0 \tag{1}$$

gilt.

Ein Punkt, in dem solche Multiplikatoren $(\lambda_0, \lambda, \mu)$ existieren, wird als ein Fritz-John-Punkt bezeichnet. Wenn in einem Fritz-John-Punkt der Multiplikator λ_0 gleich Null ist, so ist diese Bedingung unabhängig von der Zielfunktion und kann somit keine Beziehung zur Optimalität besitzen. Um eine solche Situation zu vermeiden, werden Bedingungen an die Restriktionsfunktionen gestellt, die als Constraint Qualification bekannt sind. Wir führen zunächst zwei solche Bedingungen, die in unserer Arbeit eine besondere Rolle spielen, ein:

Definition 3
Es sei $\overline{x} \in X$ ein zulässiger Punkt.

1. *Die Linear Independence Constraint Qualification (LICQ) gilt im Punkt $\overline{x}$, falls die Menge $\{Dh_i(\overline{x}), Dg_j(\overline{x}), i \in I, j \in J_0(\overline{x})\}$ linear unabhängig ist.*

2. *Die Mangasarian-Fromowitz Constraint Qualification (MFCQ) ist im Punkt $\overline{x}$ erfüllt, wenn zwei Bedingungen gelten.*

   - *Die Menge $\{Dh_i(\overline{x}), i \in I\}$ ist linear unabhängig.*
   - *Es existiert ein $\xi \in \mathbb{R}^n$, so daß*

$$\begin{aligned} Dh_i(\overline{x})\xi &= 0, \quad i \in I, \\ Dg_j(\overline{x})\xi &> 0, \quad j \in J_0(\overline{x}). \end{aligned}$$

Diese beiden Bedingungen werden oft auf Mengen angewandt. Zum Beispiel, wenn wir sagen, daß MFCQ auf X gilt, heißt das natürlicherweise die Erfüllung in jedem Punkt von X. Die LICQ ist offensichtlich stärker als die MFCQ. Wenn die Gleichung (1) mit dem Vektor ξ der MFCQ skalar multipliziert wird, ist der folgende Satz ohne Schwierigkeiten unter Berücksichtigung der beiden Bedingungen aus der MFCQ leicht zu beweisen (siehe [51]).

Satz 2 *(Karush-Kuhn-Tucker)*
Wenn die MFCQ in einem lokalen Minimalpunkt $\bar{x} \in X$ erfüllt ist, dann existieren Multiplikatoren (auch Lagrange-Multiplikatoren genannt) $(\lambda, \mu) = (\lambda_i, \mu_j)$ mit $i \in I$ und $j \in J_0(\overline{x})$, so daß $\mu_j \geq 0, \forall j \in J_0(\overline{x})$ und die Gleichung

$$Df(\overline{x}) - \sum_{i \in I} \lambda_i Dh_i(\overline{x}) - \sum_{j \in J_0(\overline{x})} \mu_j Dg_j(\overline{x}) = 0 \tag{2}$$

gilt.

Dieser Satz gibt uns Anlaß für die Definition 4.

Definition 4
Ein zulässiger Punkt $\bar{x} \in X$ heißt ein stationärer Punkt (oder Karush-Kuhn-Tucker-Punkt) , wenn Multiplikatoren $(\lambda, \mu) = (\lambda_i, \mu_j, i \in I, j \in J_0(\overline{x}))$ existieren, so daß $\mu_j \geq 0, \forall j \in J_0(\overline{x})$ und (2) erfüllt sind. Wenn (λ, μ) ein solches Multiplikatorentupel ist, wird der ganze Vektor $(\bar{x}, \lambda, \mu)$ als Karush-Kuhn-Tucker-Vektor (oder KKT-Vektor) bezeichnet.

Die Lagrange-Funktion in Bezug auf einen zulässigen Punkt $\bar{x} \in X$ wird als

$$L(x, \lambda, \mu) = Df(x) - \sum_{i \in I} \lambda_i h_i(x) - \sum_{j \in J_0(\overline{x})} \mu_j g_j(x)$$

definiert.

Es sei bemerkt, daß die KKT-Vektoren Nullstellen der Ableitung der zugehörigen Lagrange-Funktion bilden. Somit haben wir eine notwendige Optimalitätsbedingung (stationärer Punkt und Fritz-John-Satz), die nur Informationen erster Ordnung nutzt.

Für die Beschreibung von Lösungsansätzen und Verfahren, oder von anderen Optimalitätseigenschaften (z.B. lokale Maximalpunkte), werden auch andere Begriffe erster Ordnung eingeführt. Der erste ist eine Abkehr von der Positivität der Multiplikatoren bezüglich der Ungleichungen bei stationären Punkten.

Definition 5
Ein zulässiger Punkt, in dem die LICQ erfüllt ist, heißt kritisch, falls Multiplikatoren $(\lambda, \mu) = (\lambda_i, \mu_j, i \in I, j \in J_0(\overline{x}))$ existieren (nicht unbedingt mit nichtnegativen μ), so daß die Gleichung (2)erfüllt ist.

Für die Formulierung von Optimalitätsbedingungen zweiter Ordnung nehmen wir Minimalpunkte, in denen die LICQ erfüllt ist, das heißt kritische Punkte, die auch stationäre Punkte sind. In solchen Punkten ist der Tangentialraum an die zulässige Menge gut beschreibbar.

Definition 6
Es sei $\bar{x} \in X$ ein zulässiger Punkt, in dem die LICQ erfüllt ist. Der Tangentialraum an $\bar{x}$ bezüglich der zulässigen Menge X wird als

$$T_{\bar{x}}M = \{\eta \in I\!R^n | Dh_i(\bar{x})\eta = 0, Dg_j(\bar{x})\eta = 0, i \in I, j \in J_0(\bar{x})\}$$

definiert.

Da die LICQ erfüllt ist, sind die Multiplikatoren $(\lambda, \mu) = (\lambda_i, \mu_j, i \in I, j \in J_0(\overline{x}))$ eindeutig bestimmt. Wir werden uns auf den Fall der strikten Komplementarität (d.h. $(\mu_j > 0, \forall j \in J_0(\overline{x}))$ beschränken.

Satz 3 *(Notwendige Optimalitätsbedingung zweiter Ordnung)*
Es seien $\bar{x} \in X$ ein kritischer Punkt und $(\bar{\lambda}, \bar{\mu}) = (\bar{\lambda}_i, \bar{\mu}_j, i \in I, j \in J_0(\overline{x}))$ die zugehörigen Multiplikatoren. Die Optimalitätsbedingungen erster Ordnung seien mit strikter Komplementarität erfüllt, d.h. $(\bar{\mu}_j > 0, \forall j \in J_0(\overline{x}))$.

Notw. *Falls $\bar{x}$ ein lokaler Minimalpunkt ist, gilt die folgende Optimalitätsbedingung zweiter Ordnung: $\zeta^T D_x^2 L(\bar{x}, \bar{\lambda}, \bar{\mu})\zeta \geq 0, \ \forall \zeta \in T_{\bar{x}}M$.*

Hinr. *Falls die folgende Optimalitätsbedingung zweiter Ordnung,*

$$\zeta^T D_x^2 L(\bar{x}, \bar{\lambda}, \bar{\mu})\zeta > 0, \ \forall \zeta \in T_{\bar{x}}M$$

gilt, dann ist $\bar{x}$ ein strikter lokaler Minimalpunkt.

Es sei V eine Matrix, deren Spalten eine Basis des linearen Unterraumes $T_{\bar{x}}M$ bilden. Dann lassen sich die Bedingungen zweiter Ordnug in die positive Semi-Definitheit (bzw. positive Definitheit) der Matrix $V^T D_x^2 L(\bar{x}, \bar{\lambda}, \bar{\mu})V$ umformulieren. Diese Definitheit hängt entscheidend von dem Vorzeichen der Eigenwerte von $V^T D_x^2 L(\bar{x}, \bar{\lambda}, \bar{\mu})V$ ab.

Definition 7
Für eine beliebige symmetrische Matrix A wird das Tripel

$$In(A) = (p(A), n(A), z(A))$$

ihr Innertiatripel genannt, wobei $p(A)$, $n(A)$ und $z(A)$ die Anzahl der positiven, negativen und Null-Eigenwerte von A bezeichnen.

Der Satz von Sylvester sagt aus, daß dieses Inertiatripel für alle Matrizen $B^T AB$, wobei B eine quadratische und reguläre Matrix ist, das gleiche bleibt. Speziell folgt, daß das Inertiatripel der Matrix $V^T D_x^2 L(\bar{x}, \bar{\lambda}, \bar{\mu})V$ nicht von der eventuellen Selektion der Basis abhängt. Daraus folgt, daß die Definitheitseigenschaften und sogar die Regularität der Matrix $V^T D_x^2 L(\bar{x}, \bar{\lambda}, \bar{\mu})V$ nicht von der spezifischen Matrix V abhängen.

2.2 Transversalität, Satz von Sard und Generizität

Im diesem Abschnitt führen wir einige Begriffe ein, die die fundamentalen mathematischen Werkzeuge in dieser Arbeit bilden.

Definition 8 *(siehe z.B. [62],[31])*
Es seien k, m, n mit $k \geq 1$ und $0 \leq m \leq n$ drei natürliche Zahlen. Eine Menge M ist eine C^k-Mannigfaltigkeit der Dimension m, falls für jeden Punkt $x \in I\!R^n$ eine Umgebung $U(x)$ und eine Abbildung $\Phi \in C^k(U(x), I\!R^m)$ existieren, so daß folgendes gilt:

- $V = \{\Phi(U(x) \cap M)\}$ *ist eine offene Menge im* $I\!R^m$,
- $\Phi|_{M \cap U(x)} : M \cap U(x) \to V$ *ist bijektiv,*
- $\Phi^{-1} \in C^k(V, I\!R^n)$.

Die Zahl $n - m$ wird die Kodimension der Mannigfaltigkeit genannt. Es ist die Anzahl von Funktionen mit unabhängigen Gradienten, die lokal die Mannigfaltigkeit beschreiben.

Es sei eine Abbildung $\Phi \in C^k(U, I\!R^m)$ mit $k \geq 1$ und U offen gegeben. Ein Punkt $\bar{x} \in U$ heißt Singulärpunkt von Φ, falls die Matrix $D\Phi_{\bar{x}}(\bar{x})$ linear abhängige Zeilen besitzt. Alle anderen Punkte in U heißen regulär. Andererseits wird ein Punkt $y \in I\!R^m$ Singulärwert bezüglich Φ genannt, falls die Menge $\Phi^{-1}(y) \subset U$ mindestens einen singulären Punkt von Φ enthält. Die Punkte in $I\!R^m$, die nicht Singulärwerte bezüglich Φ sind, werden als Regulärwerte bezeichnet.

Bemerkung 1
Es sei $M \subset I\!R^n$, und wir nehmen an, daß man für jedes $\bar{x} \in M$ eine Umgebung $U(\bar{x})$ und eine Abbildung $\Phi_{\bar{x}} : U(\bar{x}) \to I\!R^m$ mit folgenden Eigenschaften findet:

- $\Phi_{\bar{x}} \in C^k(U(\bar{x}), I\!R^m)$,
- $M \cap U(\bar{x}) = \{x \in U(\bar{x}) | \Phi_{\bar{x}}(x) = 0\}$,
- $\bar{x}$ *ist ein Regulärpunkt von* $\Phi_{\bar{x}}$ *(was impliziert, daß* $m \leq n$*).*

Dann ist M eine C^k-Mannigfaltigkeit der Dimension $n-m$ und der Kodimension m.

Speziell können in der vorhergehenden Bemerkung die beschreibenden Funktionen auf diesselbe Art und Weise für jeden Punkt der Mannigfaltigkeit deklariert sein, wie z.B. bei linearen Mannigfaltigkeiten. Damit wird der Fall der durch Gleichungen beschriebenen Mengen, wo überall die LICQ erfüllt ist, berücksichtigt. Jetzt führen wir den Transversalitätsbegriff aus der Differentialgeometrie ein (siehe z.B. [32]).

Definition 9
Zwei C^k-Mannigfaltigkeiten $M_1, M_2 \subset \mathbb{R}^n$ der Kodimensionen m_1 und m_2 schneiden sich transversal (es wird mit $M_1 \overset{\top}{\cap} M_2$ bezeichnet) , wenn für jeden Punkt $\bar{x} \in M_1 \cap M_2$ folgendes gilt:
Sind $\Phi_{\bar{x},1} C^k(U(\bar{x}), \mathbb{R}^{m_1})$ und $\Phi_{\bar{x},2} C^k(U(\bar{x}), \mathbb{R}^{m_2})$ zwei beschreibende Abbildungen der Mannigfaltigkeiten auf der Umgebung $U(\bar{x})$ entsprechend Bemerkung 1, dann ist die Menge aller Zeilenvektoren der beiden Matrizen $D\Phi_{\bar{x},1}(\bar{x})$ und $D\Phi_{\bar{x},2}(\bar{x})$ linear unabhängig.

Aus dieser Definition folgt:

1. Wegen der linearen Unabhängigkeit der Zeilenvektoren gilt, falls $M_1 \cap M_2 \neq \emptyset$, daß $m_1 + m_2 \leq n$.

2. Falls zwei C^k-Mannigfaltigkeiten der Kodimensionen m_1 und m_2 sich transversal schneiden, so ist ihr Durchschnitt entweder leer oder eine C^k-Mannigfaltigkeit der Kodimension $m_1 + m_2$.

Der folgende Satz ist von zentraler Bedeutung für die Untersuchungen in den nächsten Abschnitten.

Satz 4

1. *(Satz von Sard, siehe z.B. [62]) Es seien $U \subset \mathbb{R}^n$ eine offene Menge und $\Phi \in C^k(U, \mathbb{R}^m)$ mit $k > \max(n-m, 0)$. Dann hat die Menge der Singulärwerte von Φ das m-dimensionale Lebesgue-Maß Null.*

2. *(Parametrisierter Satz von Sard, siehe z.B. [32] Satz 7.4.4)Es seien $U \subset I\!R^n$, $V \subset I\!R^m$ offene Mengen und $\Phi(x,y) \in C^k(U \times V, I\!R^p)$ mit $k > \max(n-p, 0)$. Es sei Null ein Regulärwert von Φ, dann hat die folgende Menge*

$$\left\{ y \in V \;\middle|\; \begin{array}{c} 0 \text{ ist ein Singulärwert von} \\ \Phi_y \in C^k(U, I\!R^p) \end{array} \right\}$$

das Lebesgue-Maß Null.

Hier bezeichnet $\Phi_y \in C^k(U, I\!R^p)$ die durch $\Phi_y(x) = \Phi(x,y)$ definierte Abbildung.

Abschließend führen wir die sogenannte strong- oder Whitney-C^r_s-Topologie im Raum $C^k(I\!R^n, I\!R)$ ein, wobei $r \leq k$. Eine offene Umgebung des Nullpunktes in dieser Topologie ist durch eine stetige positive Funktion $\varepsilon : I\!R^n \to I\!R$ gekennzeichnet, und als

$$V^r_\varepsilon = \left\{ \psi \in C^k(I\!R^n, I\!R) \,|\, |D_\alpha \psi(y)| < \varepsilon(y), \forall y \in I\!R^n, \forall |\alpha| \leq r \right\} \tag{3}$$

definiert.

Hier werden die Indizes $\alpha = (\alpha_1, \ldots, \alpha_n) \in I\!N^n$ und $|\alpha| = \sum_{i=1}^n \alpha_i$ benutzt. Mit $D_\alpha \psi = \frac{\partial^{|\alpha|}\psi}{\partial^{\alpha_1} x_1 \ldots \partial^{\alpha_n} x_n}$ wird die α-fache partielle Ableitung von ψ bezeichnet. Die offenen Umgebungen einer anderen Funktion g sind von der Form $g + V^r_\varepsilon$. Analog wird diese Topologie für die Fälle definiert, daß ein oder beide Indizes r und k, unter Berücksichtigung der Bedingung $r \leq k$, unendlich sind. Andererseits ist die C^r_s-Topologie für endliche Produkträume als die Produkttopologie deklariert. Bezüglich mehr Details über diese Topologie und ihre Eigenschaften verweisen wir z.B. auf [26] und [32].

Auf diese Topologie wird zugegriffen, falls globale Aussagen bezüglich der Stabilität (Offenheit) von Eigenschaften oder ihre Allgemeinheit (Dichtheit) gemacht werden.

Definition 10
Eine Untermenge des Raumes $C^k(I\!R^n, I\!R)$, versehen mit der C^r_s-Topologie, heißt generisch, falls sie den Durchschnitt einer abzählbaren Familie offener und dichter Mengen enthält. Andererseits heißt eine Eigenschaft P, gestellt auf Elemente des Raumes $C^k(I\!R^n, I\!R)$, generisch, falls die Menge, wo P gültig ist, generisch ist.

3 Einparametrische Optimierungsprobleme.

3.1 Nichtdegenerierte kritische Punkte.

Wir beginnen diesen Abschnitt mit der Einführung nichtdegenerierter kritischer Punkte.

Definition 11 *(siehe z.B. [31])*
Es seien $\bar{x} \in X$ ein kritischer Punkt und $(\bar{\lambda}, \bar{\mu})$ die zugehörigen Multiplikatoren. Der Punkt heißt nichtdegeneriert, wenn die folgenden zwei Bedingungen erfüllt sind:

ND1 $\bar{\mu}_j \neq 0, \forall j \in J_0(\overline{x})$

ND2 $D_x^2 L(\bar{x}, \bar{\lambda}, \bar{\mu})|_{T_{\bar{x}}M}$ *ist regulär.*

Die Notation $D_x^2 L(\bar{x}, \bar{\lambda}, \bar{\mu})|_{T_{\bar{x}}M}$ bezieht sich auf eine Matrix der Form $V^T D_x^2 L(\bar{x}, \bar{\lambda}, \bar{\mu}) V$, in der die Spalten von V unabhängig sind und $T_{\bar{x}}M$ linear erzeugen.Wie vorher schon erwähnt, hängt die Regularität der Matrix $V^T D_x^2 L(\bar{x}, \bar{\lambda}, \bar{\mu}) V$ nicht von der eventuell gewählten Basis V ab.

Eine Motivation für diesen Begriff wird jetzt anhand der lokalen Analyse der kritischen Punkte eines parametrischen Optimierungsproblems gegeben. Zuerst charakterisieren wir die kritischen Punkte mindestens lokal durch ein einfaches Gleichungssystem (siehe z.B. [7],[6], [32]).

Nach Definition 5 von kritischen Punkten ist leicht zu sehen, daß diese als Lösungsmenge eines Gleichungssystems aufgefaßt werden können. Ein kritischer Punkt $\bar{x}$ bildet zusammen mit den Multiplikatoren $(\bar{x}, \bar{\lambda}, \bar{\mu}) \in I\!R^n \times I\!R^m \times I\!R^{|J_0|}$ eine Nullstelle der folgenden Abbildung.

$$\mathcal{H}_{J_0(\bar{x})}(x, \lambda, \mu) = \begin{pmatrix} Df(x) - \sum_{i \in I} \lambda_i h_i(x) - \sum_{j \in J_0(\overline{x})} \mu_j g_j(x) \\ -H(x) \\ -G_{J_0(\bar{x})}(x) \end{pmatrix} \tag{4}$$

Diese Abbildung erfolgt zwischen Räumen gleicher Dimension, nämlich $I\!R^{n+m+|J_0|}$. Falls im kritischen Punkt $\bar{x}$ alle Multiplikatoren bezüglich der Ungleichungen $\bar{\mu}$ verschieden von Null sind, dann charakterisieren in einer Umgebung von $\bar{x}$ die Nullstellen der Abbildung (4) die kritischen Punkte des Optimierungsproblems. Diese Charakterisierung der kritischen Punkte

erfolgt in einer Umgebung, in der zu J_0 keine neuen Indizes hinzukommen und die LICQ erfüllt bleibt.

Dieses Gleichungsystem ist nützlich, wenn man für die Untersuchung von Stabilitätseigenschaften der kritischen Punkte den Satz über Implizite Funktionen benutzt.

Dann betrachten wir ein parameterabhängiges Optimierungsproblem der Form $P(d)$, wobei $d \in I\!R^s$. Ein solches parametrisches Problem wird durch die Zielfunktion $f(x,d) : I\!R^{n+s} \to I\!R$, die Gleichungsrestriktionen $H(x,d) : I\!R^{n+s} \to I\!R^m$, und die Ungleichungsrestriktionen $G(x,d) : I\!R^{n+s} \to I\!R^p$ definiert. Für jeden fixierten Wert des Parametervektors d betrachtet man das Optimierungsproblem, das durch die Daten $(f(.,d), H(.,d), G(.,d))$ gegeben ist. Die besondere Rolle des Parametervektors wird in der Bezeichnung $P(d)$ beschrieben. Wenn der Parameter zum eindimensionalen Raum $I\!R$ gehört, wird er mit t bezeichnet.

Betrachten wir die Menge

$$\Sigma_{krit} = \{(x,y) | x \text{ ist kritischer Punkt von } P(y)\}.$$

In diesem Zusammenhang machen wir darauf aufmerksam, daß die aktive Indexmenge J_0 auch von d abhängt. Wir nutzen deswegen die Bezeichnung $J_0(x,d)$, falls das Problem parametrisch ist und die Änderungen der Parameter berücksichtigt werden sollen.

Jetzt wird die interessante und wichtige Frage gestellt, wie sich ein kritischer Punkt eines Optimierungsproblems für einen festen Parameterwert in Abhängigkeit von der Veränderung der Parameter verhält. Diese Frage hängt natürlich eng mit der lokalen Struktur der Menge Σ_{krit} zusammen. Um diese lokale Analyse durchzuführen, betrachten wir ein Paar $(\bar{x}, \bar{d})$ und die zugehörigen Multiplikatoren $(\bar{\lambda}, \bar{\mu})$. Man kann jetzt das System (4) auch in Abhängigkeit vom Parameter d und bezüglich $\bar{x}$ betrachten.

$$\mathcal{H}_{J_0(\bar{x},\bar{d})}(x,\lambda,\mu,d) = \begin{pmatrix} D_x f_{(x,d)} - \sum_{i \in I} \lambda_i D_x h_{i(x,d)} - \sum_{j \in J_0(\bar{x},\bar{d})} \mu_j D_x g_{j(x,d)} \\ -H_{(x,d)} \\ -G_{J_0(\bar{x},\bar{d})(x,d)} \end{pmatrix} \tag{5}$$

Es folgt, $\mathcal{H}_{J_0(\bar{x},\bar{d})}(\bar{x},\bar{\lambda},\bar{\mu},\bar{d}) = 0$. Gilt ND1 (erste Bedingung der Nichdegenerierheit), dann gibt es eine Umgebung $U_{(\bar{x},\bar{d})}$von $(\bar{x},\bar{d})$, in der die aktive Indexmenge $J_0(x,d)$ eine Untermenge von $J_0(\bar{x},\bar{d})$ bleibt und die LICQ

erfüllt ist. In dieser Umgebung $U_{(\bar{x},\bar{d})}$ ist die Menge Σ_{krit} die Projektion der folgenden Menge

$$\left\{(x,\lambda,\mu,d)|\mathcal{H}_{J_0(\bar{x},\bar{d})}(x,\lambda,\mu,d)=0\right\} \tag{6}$$

auf die Variablen (x,d).

Lokal um $(\bar{x},\bar{\lambda},\bar{\mu},\bar{d})$ ist die Menge (6) eine Mannigfaltigkeit der Dimension d, falls alle Zeilen der Jacobi-Matrix $D\mathcal{H}_{J_0(\bar{x},\bar{d})}(\bar{x},\bar{\lambda},\bar{\mu},\bar{d})$ linear unabhängig sind. Die Regularität der Matrix $D_{(x,\lambda,\mu)}\mathcal{H}_{J_0(\bar{x},\bar{d})}(\bar{x},\bar{\lambda},\bar{\mu},\bar{d})$ impliziert schon die Unabhängigkeit der genannten Zeilen und ist, wie folgendes Lemma zeigt, unter ND1 äquivalent zur Bedingung ND2.

Lemma 1 *[37]*
Es sei A eine symmetrische Matrix mit n Zeilen und Spalten und B eine Matrix mit n Zeilen und m Spalten. Weiter sei $\bar{m}$ der Rang von B. Die Matrix C mit $n+m$ Zeilen und Spalten sei als

$$C=\begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix}$$

definiert. Dann gilt die folgende Beziehung der Inertiatripel (siehe Definition 7).

$$In(C)=In(C|_{KernB^T})+In(\bar{m},\bar{m},m-\bar{m})$$

Die Bezeichnung $C|_{KernB^T}$ steht wieder für die Einschränkung von C auf den linearen Unterraum $Kern\ B^T=\left\{y\in I\!R^n|B^Ty=0\right\}$. Die Matrix $D_{(x,\lambda,\mu)}\mathcal{H}_{J_0(\bar{x},\bar{d})}(\bar{x},\bar{\lambda},\bar{\mu},\bar{d})$ hat eine ähnliche Gestalt wie C, wobei $A=D_x^2L(\bar{x},\bar{\lambda},\bar{\mu},\bar{d})$ und B die $n\times(m+|J_0(\bar{x},\bar{d})|)$ Matrix ist, deren Spalten die Vektoren $\left\{D_xh_i(\bar{x},\bar{d}),D_xg_j(\bar{x},\bar{d}),i\in I,j\in J_0(\overline{x},\bar{d})\right\}$ sind. Nach Definition folgt, daß $KernB^T=T_{\bar{x}}M$ ist. Die LICQ heißt jetzt nichts anderes als $m=\bar{m}$ im vorigen Lemma, und daraus folgt, daß unter ND1 die Bedingung ND2 äquivalent mit der Regularität der Matrix $D_{(x,\lambda,\mu)}\mathcal{H}_{J_0(\bar{x},\bar{d})}(\bar{x},\bar{\lambda},\bar{\mu},\bar{d})$ ist.

Die vorige Regularität impliziert nach dem Satz über Implizite Funktionen die Existenz einer Umgebung $U_{\bar{x}}\times U_{\bar{\lambda}}\times U_{\bar{\mu}}\times U_{\bar{d}}$ und einer eindeutigen Funktion $\Phi:U_{\bar{d}}\to U_{\bar{x}}\times U_{\bar{\lambda}}\times U_{\bar{\mu}}$, so daß die Gleichung $\Phi(d)=(x,\lambda,\mu)$ die Menge (6) in der Umgebung $U_{\bar{x}}\times U_{\bar{\lambda}}\times U_{\bar{\mu}}\times U_{\bar{d}}$ beschreibt. Ohne Beschränkung der Allgemeinheit können wir annehmen, daß $U_{\bar{x}}\times U_{\bar{d}}=$

$U_{(\bar{x},\bar{d})}$ eine Umgebung ist, in der die LICQ und die aktive Indexmenge stabil bleiben. In diesem Fall beschreibt, wegen der Eindeutigkeit der Multiplikatoren unter der LICQ, die Gleichung $\Phi(d) = (x, \lambda, \mu)$ die genannte Menge (6) nicht nur auf $U_{\bar{x}} \times U_{\bar{\lambda}} \times U_{\bar{\mu}} \times U_{\bar{d}}$, sondern auf der größeren offenen Menge $U_{\bar{x}} \times I\!R^m \times I\!R^{|J_0(\bar{x},\bar{d})|} \times U_{\bar{d}}$. Das heißt mit anderen Worten, daß die Menge Σ_{krit} in der Umgebung $U_{(\bar{x},\bar{d})}$ durch das Gleichungssystem $x - \Phi_x(d) = 0$ beschrieben wird. Die Menge Σ_{krit} ist um den nichtdegenerierten kritischen Punkt eine differenzierbare Mannigfaltigkeit und läßt sich nach d parametrisieren, was die differenzierbare Abhängigkeit der kritischen Punkte bedeutet.

Die Stabilitätseigenschaften von stationären Punkten in Abhängigkeit von Störungen in den beschreibenden Funktionen des Problems sind genau untersucht worden. Der von Robinson [59] eingeführte Begriff der strengen Regularität einer Lösung eines Systems verallgemeinerter Gleichungen führt im Fall von stationären Vektoren zum Begriff eines streng regulären stationären Vektors $(\bar{x}, \bar{\lambda}, \bar{\mu})$. Dieser Begriff impliziert die Erfüllung der LICQ, aber andererseits ist die Bedingung ND1 nicht erforderlich.

Ein anderer wichtiger Beitrag ist der von Kojima [44] eingeführte Begriff eines streng stabilen stationären Punktes x. In diesem Fall sind ND1 und auch die LICQ nicht erforderlich. In beiden Fällen wird die Existenz und Eindeutigkeit von Lösungen in Umgebungen von $(\bar{x}, \bar{\lambda}, \bar{\mu})$ bzw. $\bar{x}$ und die Lipschitzstetigkeit für hinreichend wenig gestörte Probleme charakterisiert. Zahlreiche Untersuchungen dieser beiden Begriffe, unter anderem ihre Äquivalenz unter der LICQ, und äquivalente Charakterisierungen sind in der Literatur zu finden, siehe z.B. [36], [37], [40], [41], [42], [43], [46], [47], [64].

Definition 12
Es seien $\bar{x}$ ein nichtdegenerierter kritischer Punkt und $(\bar{\lambda}, \bar{\mu})$ die dazugehörigen Multiplikatoren. Der lineare Index (Ko-Index) wird definiert als die Anzahl der negativen (bzw. positiven) Multiplikatoren $\bar{\mu}_j$, $j \in J_0(\bar{x})$. Diese beiden linearen Indizes werden bezeichnet als LI (bzw. LCI). Der quadratische Index (Ko-Index) wird definiert als die Anzahl der negativen (bzw. positiven) Eigenwerte von $D_x^2 L(\bar{x}, \bar{\lambda}, \bar{\mu})|_{T_{\bar{x}}M}$. Die beiden quadratischen Indizes werden mit QI (bzw QCI) bezeichnet.

Nach Definition gilt $LI + LCI = |J_0(\bar{x})|$ und $QI + QCI = n - (m + |J_0(\bar{x})|)$. Die Tupel der vier Indizes (LI, LCI, QI, QCI) bestimmen das Verhalten der

Zielfunktion über der zulässigen Menge X um den Punkt $\bar{x}$ (siehe [31]). Zum Beispiel ist ein nichtdegenerierter kritischer Punkt ein lokaler Minimalpunkt (Maximalpunkt), falls $LI + QI = 0$ (bzw. $LCI + QCI = 0$).

Die nichtdegenerierten kritischen Punkte sind der generische Zustand von kritischen Punkten, wie der folgende Satz andeutet. Die Menge der regulären Optimierungsprobleme ist als

$$\mathcal{F}_{OP} = \left\{ (f, H, G) \in C^\infty(\mathbb{R}^n, \mathbb{R}^{m+p+1}) \;\middle|\; \begin{array}{l} \text{Bezüglich } P(f, H, G) \text{ gilt die} \\ \text{LICQ im zulässigen Bereich und} \\ \text{alle kritischen Punkte sind} \\ \text{nichtdegeneriert} \end{array} \right\}$$

zu definieren.

Satz 5 *[32]*
$\mathcal{F}_{OP}$ ist eine offene und dichte Menge in $C^\infty(\mathbb{R}^n, \mathbb{R}^{m+p+1})$, versehen mit der C^k-strong-Topologie, mit $k \geq 2$.

3.2 Die Klasse von Jongen, Jonker und Twilt.

Dieser Abschnitt ist der Untersuchung einparametrischer Optimierungsprobleme gewidmet. Dazu beschränken wir uns auf parametrische Probleme der Form $P(t)$, wobei $t \in \mathbb{R}$. Diese Untersuchung kann als die Beschreibung von Lösungsmengen, aber nicht auf einer lokalen Ebene wie im vorigen Abschnitt, sondern von einem globalen Gesichtpunkt aus, verstanden werden.

Dafür gibt es zwei wichtige Zugänge. Der von Kojima und Hirabayashi (siehe [45]) basiert auf einer geeigneten Erweiterung der Begriffe *regulärer Punkt* und *regulärer Wert* für eine stückweise stetig differenzierbare Abbildung, deren Nullstellen die Menge der stationären Vektoren widerspiegeln.

Wir werden uns in diesem Abschnitt und allgemein in dieser Arbeit auf den von Jongen, Jonker und Twilt eingeführten Zugang beschränken. Zunächst führen wir detailliert die entsprechenden Begriffe ein.

Definition 13 *(siehe [34], [35])*
Es sei $\bar{x}$ ein zulässiger Punkt des einparametrischen Optimierungsproblems $P(t)$. Der Punkt $(\bar{x}, \bar{t})$ heißt ein verallgemeinerter kritischer Punkt, falls die Vektoren $\{D_x f(\bar{x}, \bar{t}), D_x h_i(\bar{x}, \bar{t}), D_x g_j(\bar{x}, \bar{t}), i \in I, j \in J_0(\bar{x}, \bar{t})\}$ linear abhängig sind.

Unter diesen Begriff fallen alle bisher eingeführten Definitionen bezüglich Optimalitätsbedingungen erster Ordnung, wie F-John-Punkt, stationärer oder kritischer Punkt. Es sei folgende Menge

$$\Sigma_{gc} = \{(x,t) | x \text{ ist verallgemeinerter kritischer Punkt von } P(t)\} \qquad (7)$$

definiert.

Die Menge Σ_{gc} in ihrer lokalen und globalen Struktur wird möglichst genau unter bestimmten Voraussetzungen untersucht. Ein erster Schritt ist gewährleistet, indem wir die Menge Σ_{krit} um die nichtdegenerierten kritischen Punkte untersucht haben. In einer Umgebung eines kritischen Punktes (LICQ ist erfüllt) stimmen die Mengen Σ_{gc} und Σ_{krit} überein. Im Fall der Nichtdegenerierheit ist die Menge Σ_{gc} lokal eine nach t parametrisierte Kurve.

Wie im vorigen Abschnitt erwähnt wurde, ist die Nichtdegeneriertheit der typische Zustand von Optimierungsproblemen ohne Parameter. Das ist aber nicht der Fall, wenn wir einparametrische Probleme betrachten. Hier ist nicht typisch, daß alle kritischen Punkte aller Probleme in jedem Parameterwert $t \in I\!R$ nichtdegeneriert sind. Man muß die lokale Struktur von Σ_{gc} in den anderen Fällen, als den nichtdegenerierten, auch untersuchen.

Ausgangspunkt ist der nichtdegenerierte Fall selbst. Man betrachtet die Verletzung jeder einzelnen Bedingung, nämlich ND1, ND2 oder LICQ. Dafür geben wir die Definition von 5 Typen verallgemeinerter kritischer Punkte an. Der erste Typ ist nichtdegeneriert und die anderen vier sind verschiedene Spezialfälle von degenerierten Punkten. Es sei bemerkt, daß es unter der LICQ in einem verallgemeinerten kritischen Punkt eindeutig bestimmte Multiplikatoren (λ, μ) gibt , die die folgende Gleichung

$$L(x,\lambda,\mu,t) = D_x f(x,t) - \sum_{i \in I} \lambda_i D_x h_i(x,t) - \sum_{j \in J_0} \mu_j D_x g_j(x,t) = 0$$

realisieren.

In den folgenden Definitionen wird die Bezeichnung $z = (x,t)$ benutzt.

Definition 14 *(siehe [35])*
Ein verallgemeinerter kritischer Punkt $\bar{z} = (\bar{x}, \bar{t})$ ist vom <u>Typ 1</u>, falls $\bar{x}$ ein nichtdegenerierter kritischer Punkt des Problems $P(\bar{t})$ ist.

Die lokale Struktur von Σ_{gc} ist im folgenden Bild 1 gezeigt. In einer Umgebung von $\bar{z}$ sind alle Punkte von Σ_{gc} nichtdegeneriert, und das Indextupel (LI, LCI, QI, QCI) bleibt unverändert.

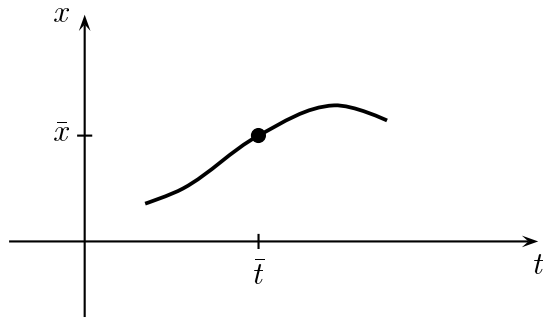


Bild 1: Σ_{gc} in einer Umgebung von $\bar{z} \in \Sigma_{gc}^1$.

Definition 15 *(siehe [35])*
Ein verallgemeinerter kritischer Punkt $\bar{z} = (\bar{x}, \bar{t})$ ist vom Typ 2, falls folgendes gilt:

1. *LICQ ist erfüllt im $\bar{z}$.*

2. *Alle Multiplikatoren $\bar{\mu}_j, j \in J_0(\bar{z})$, bis auf einen einzigen, sind verschieden von Null. Der verschwindende Multiplikator sei z.B. $\bar{\mu}_p$, wobei $p \in J_0(\bar{z})$.*

3. *$D_x^2 L(\bar{x}, \bar{\lambda}, \bar{\mu}, \bar{t})|_{T_{\bar{x}}M}$ ist regulär.*

4. *$D_x^2 L(\bar{x}, \bar{\lambda}, \bar{\mu}, \bar{t})|_{T_{\bar{x}}^+ M}$ ist regulär, wobei*

$$T_{\bar{x}}^+ M = \left\{\eta \in I\!R^n | Dh_i(\bar{x}, \bar{t})\eta = 0, Dg_j(\bar{x}, \bar{t})\eta = 0, i \in I, j \in J_0^+(\bar{x}, \bar{t})\right\}$$

und $J_0^+(\bar{z}) = \{j \in J_0(\bar{z}) | \bar{\mu}_j \neq 0\}$.

5. *Es seien*

- *$\Pi : I\!R^{n+1} \to I\!R^{m+|J_0^+(\bar{z})|}$ definiert als $\Pi = (H, g_j, j \in J_0^+(\bar{z}))$, und*
- *W eine Matrix, deren Spalten eine lineare Basis des Unterraumes $T_{\bar{x}}^+ M$ sind.*

Dann ist die folgende Zahl γ

$$\gamma = D_x g_p(\bar{z})(\alpha + \beta) + D_t g_p(\bar{z})$$

verschieden von Null, wobei:

$$\begin{aligned} \alpha &= -((D_x\Pi(\bar{z}))^\dagger)^T \cdot D_t\Pi(\bar{z}), \\ \beta &= -W(W^T \cdot D_x^2 L \cdot W)^{-1} W^T \cdot \left[D_x^2 L \cdot \alpha + D_t D_x^T L\right]. \end{aligned}$$

Hierbei ist $B^\dagger = (B^T B)^- 1 B^T$ die Moore-Penrose Inverse einer Matrix B, die nur linear unabhängige Spalten besitzt.

Da der Multiplikator μ_p in der vorigen Definition Null ist, kann man die Restriktion g_p lokal um $\bar{z}$ einerseits als Gleichung betrachten und andererseits streichen. Damit entstehen zwei Gleichungssysteme ähnlich (5). Wegen der LICQ und der dritten und vierten Bedingungen in Definition 15 erhält man aus beiden Systemen lokale kritische Kurven $x^0(t)$ und $x^1(t)$, wobei $x^0(t)$ durch die Entfernung von g_p und $x^1(t)$ durch ihre Betrachtung als Gleichung entstehen. Die fünfte Bedingung der vorigen Definition kann dann als die folgende Ungleichung

$$Dg_p(x^0(\bar{t}), \bar{t}) = D_x g_p(\bar{z}) \cdot D_t x^0(\bar{t}) + D_t g_p(\bar{z}) = \gamma \neq 0$$

geschrieben werden.

Diese beiden Kurven schneiden sich transversal in $\bar{z}$. Die lokale Struktur der Menge Σ_{gc} ist in einer Umgebung von $\bar{z}$ im folgenden Bild 2 gezeigt. Alle Punkte von Σ_{gc} in der Umgebung von $\bar{z}$, außer $\bar{z}$ selbst, sind nichtdegenerierte kritische Punkte. Die Beziehungen zwischen den jeweiligen Index-Tupel (LI, LCI, QI, QCI) auf jedem Zweig von Σ_{gc} hängt von zwei sogenannten charakteristischen Zahlen ab (siehe [35]). Diese sind $\delta = \delta_1 - \delta_2$ und $sign(\gamma)$, wobei δ_1 (bzw. δ_2) die Anzahl negativer Eigenwerte von $D_x^2 L(\bar{x}, \bar{\lambda}, \bar{\mu}, \bar{t})|_{T_{\bar{x}}^+ M}$ (bzw. $D_x^2 L(\bar{x}, \bar{\lambda}, \bar{\mu}, \bar{t})|_{T_{\bar{x}} M}$) ist.

Definition 16 *(siehe [35])*
Ein verallgemeinerter kritischer Punkt $\bar{z} = (\bar{x}, \bar{t})$ ist vom Typ 3, falls folgendes gilt:

1. *LICQ ist erfüllt in $\bar{z}$.*

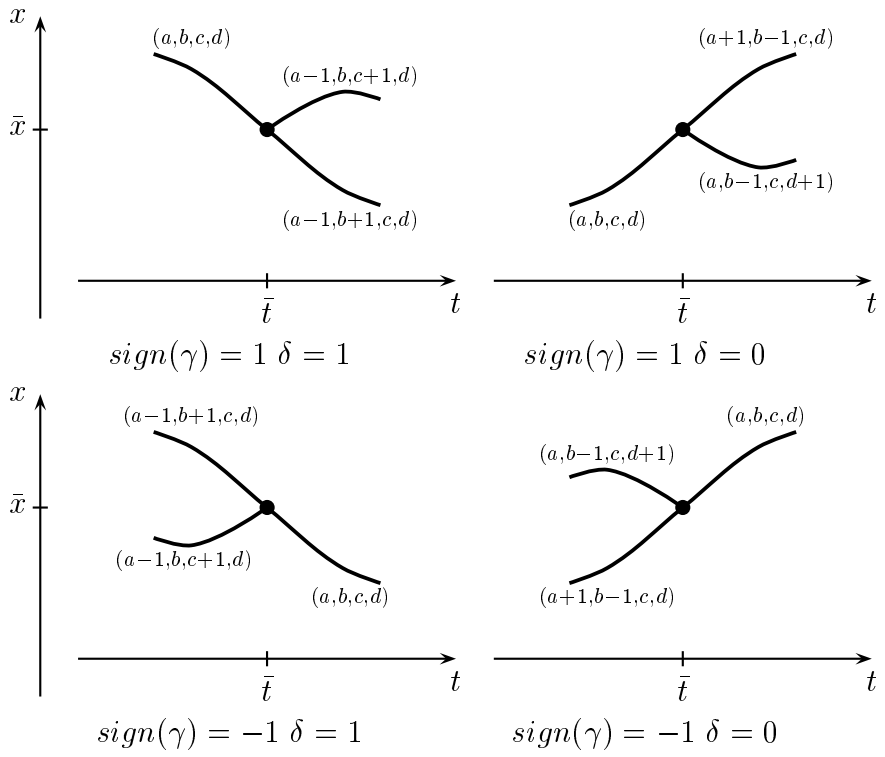


Bild 2: Σ_{gc} in einer Umgebung von $\bar{z} \in \Sigma_{gc}^2$.

2. *Alle Multiplikatoren* $\bar{\mu}_j, j \in J_0(\bar{z})$ *sind verschieden von Null.*

3. $D_x^2 L(\bar{x}, \bar{\lambda}, \bar{\mu}, \bar{t})|_{T_{\bar{x}}M}$ *hat genau einen Eigenwert, der gleich Null ist.*

4. *Es seien*

- $\Pi : I\!R^{n+1} \to I\!R^{m+|J_0(\bar{z})|}$ *definiert als* $\Pi = (H, G_{J_0(\bar{z})})$,
- W *eine Matrix, deren Spalten eine lineare Basis des Unterraumes* $T_{\bar{x}}M$ *sind, und*
- $w \neq 0$ *ein Vektor mit* $W^T D_x^2 L(\bar{x}, \bar{\lambda}, \bar{\mu}, \bar{t}) W \cdot w = 0$ *und* $v = W \cdot w$.

Dann sind die beiden folgenden Zahlen

$$\begin{aligned} \beta_1 &= v^T (D_x^3 L \cdot v) v - 3 v^T D_x^2 L((D_x^T \Pi)^\dagger)^T \cdot (v^T D_x^2 \Pi \cdot v) \\ \beta_2 &= D_t (D_x L \cdot v) - D_t^T \Pi \cdot (D_x^T \Pi)^\dagger \cdot^T D_x^2 L \cdot v \end{aligned}$$

verschieden von Null.

Betrachten wir das Optimierungssproblem (siehe [60]).

$$\begin{aligned} \mathcal{P} \quad & \min_{(x,\lambda,\mu,t) \in \mathcal{X}} \quad \Lambda(x, \lambda, \mu, t) = t, \\ & \mathcal{X} = \left\{ (x, \lambda, \mu, t) \in I\!R^{n+m+|J_0(\bar{z})|+1} | \mathcal{H}_{J_0(\bar{x},\bar{d})}(x, \lambda, \mu, t) = 0 \right\} \end{aligned} \tag{8}$$

Der Punkt $(\bar{x}, \bar{\lambda}, \bar{\mu}, \bar{t}) \in \mathcal{X}$ ist ein zulässiger Punkt von $\mathcal{P}$. Wegen der Gültigkeit von LICQ und ND1 bildet die Projektion von $\mathcal{X}$ auf (x, t) die Menge Σ_{gc}. Unter der ersten und dritten Bedingung der vorigen Definition 16 ist die Gültigkeit der LICQ für die Menge $\mathcal{X}$ im Punkt $(\bar{x}, \bar{\lambda}, \bar{\mu}, \bar{t})$ äquivalent zur Bedingung $\beta_2 \neq 0$. Ebenfalls äquivalent dazu ist die Eigenschaft von $(\bar{x}, \bar{\lambda}, \bar{\mu}, \bar{t})$, kritischer Punkt von $\mathcal{P}$ zu sein. Ein kurzer rechnerischer Aufwand zeigt, daß dann $\beta_1 \neq 0$ äquivalent zur Bedingung ND2 der Nichtdegeneriertheit von $(\bar{x}, \bar{\lambda}, \bar{\mu}, \bar{t})$ als kritischer Punkt ist. Mit anderen Worten ist wegen der Präsenz von nur Gleichungsrestriktionen die auf den ersten Blick rätselhafte vierte Bedingung von Definition 16 äquivalent zur Bedingung, daß $(\bar{x}, \bar{\lambda}, \bar{\mu}, \bar{t})$ ein nichtdegenerierter kritischer Punkt von $\mathcal{P}$ sei (siehe [60], [61], [34]).

Diese andere Charakterisierung gibt uns einen geometrischen Blick auf die lokale Struktur der Kurve $\mathcal{X}$ um $(\bar{x}, \bar{\lambda}, \bar{\mu}, \bar{t})$: sie besitzt einen Maximalpunkt

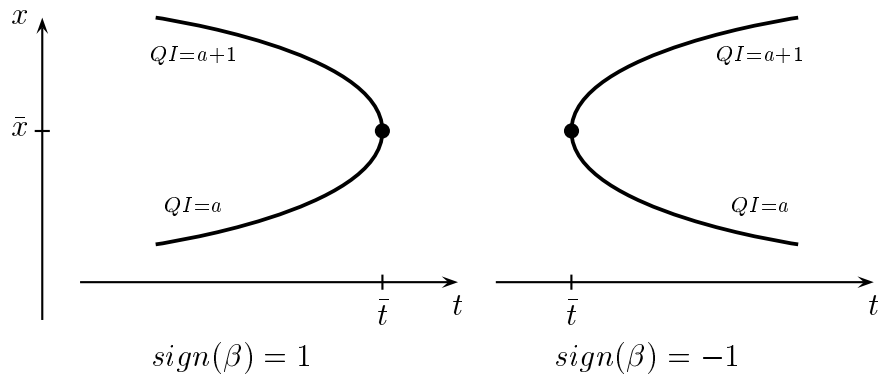


Bild 3: Σ_{gc} in einer Umgebung von $\bar{z} \in \Sigma_{gc}^3$.

oder einen Minimalpunkt ($QI + QCI = 1$) bezüglich der Komponente t. Da $\mathcal{X}$ mit jeder der Komponenten von x lokal parametrisiert werden kann, besitzt ihre Projektion auf (x, t), nämlich die Menge Σ_{gc}, die in der folgenden Abbildung 3 dargestellte geometrische Gestalt. Wieder sind außer $\bar{z}$ alle Punkte von Σ_{gc} in einer Umgebung nichtdegenerierte kritische Punkte. Das lokale Verhalten des Indextupel (LI, LCI, QI, QCI) auf der Kurve Σ_{gc} hängt in diesem Fall von den folgenden zwei charakteristischen Zahlen ab: α und $sign(\beta)$, wobei α die Anzahl der negativen Eigenwerte von $D_x^2 L(\bar{x}, \bar{\lambda}, \bar{\mu}, \bar{t})|_{T_{\bar{x}}M}$ und $\beta = \beta_1 \cdot \beta_2$ ist.

Mit den beiden ersten Singularitäten wurde die Verletzung der Bedingungen ND1 und ND2 der Nichtdegenerierheit betrachtet. Anschließend ist die Verletzung der LICQ zu untersuchen. Dabei sind die Punkte aus den betrachteten Kurven nicht mehr unbedingt kritische Punkte, sondern es kommen auch reine verallgemeinerte kritische Punkte vor.

Definition 17 *(siehe [35])*
Ein verallgemeinerter kritischer Punkt $\bar{z} = (\bar{x}, \bar{t})$ ist vom Typ 4, falls folgendes gilt:

1. $0 < m + |J_0(\bar{z})| < n + 1$.

2. *Der lineare Unterraum, erzeugt von den Vektoren*

$$\{D_x h_i(\bar{z}), D_x g_j(\bar{z}), i \in I, j \in J_0(\bar{z})\},$$

hat die Dimension $m + p - 1$.

3. *Nach der vorhergehenden Bedingung existieren bis auf eine multiplikative Konstante eindeutig bestimmte Multiplikatoren* $\bar{\lambda}_i, i \in I$ *und* $\bar{\mu}_j, j \in J_0(\bar{z})$*, die nicht alle Null sind und*

$$\sum_{i \in I} \bar{\lambda}_i D_x h_i(\bar{z}) + \sum_{j \in J_0(\bar{z})} \bar{\mu}_j D_x g_j(\bar{z}) = 0 \tag{9}$$

erfüllen. Dabei sind alle $\bar{\mu}_j, j \in J_0(\bar{z})$ *verschieden von Null. Ein solcher Multiplikatorenvektor sei* $(\bar{\lambda}_i, \bar{\mu}_j, i \in I, j \in J_0(\bar{z}))$.

4. *Es seien* $\mathbf{L}(z) = \sum_{i \in I} \bar{\lambda}_i h_i(z) + \sum_{j \in J_0(\bar{z})} \bar{\mu}_j g_j(z)$ *und* W *eine Matrix, deren Spalten eine lineare Basis des folgenden* $(n - m - |J_0(\bar{z})| + 1)$*-dimensionalen Unterraumes*

$$T = \{\eta \in I\!R^n | Dh_i(\bar{z})\eta = 0, Dg_j(\bar{z})\eta = 0, i \in I, j \in J_0(\bar{z})\}$$

erzeugen.

Dann ist die Zahl $D_t\mathbf{L}(\bar{z})$ *verschieden von Null, und die Matrix* $A = D_t L(\bar{z}) \cdot W^T D_x^2 \mathbf{L}(\bar{z}) W$ *ist regulär.*

5. *Es sei* $w = W^T \cdot D_x^T f(\bar{z})$.

Dann ist die Zahl $w^T A^{-1} w$ *nicht gleich Null.*

Bemerkung 2
Wir wählen einen Index $p \in J_0(\bar{z})$ *(falls* $J_0(\bar{z}) = \emptyset$*, sei* $p \in I$*) mit* $\bar{\mu}_p \neq 0$*, die Variable* $\theta = (\theta_0, \theta_i, \theta_j, i \in I, j \in J_0(\bar{z}) \setminus \{p\})$ *und den Punkt* $\bar{\theta} = (0, \bar{\lambda}_i, \bar{\mu}_j, i \in I, j \in J_0(\bar{z}) \setminus \{p\})$*. Betrachten wir die beiden folgenden Abbildungen*

$$\mathcal{L}(z, \theta) = \theta_0 D_x f(z) + \sum_{i \in I} \theta_i D_x h_i(z) + \sum_{j \in J_0(\bar{z}) \setminus \{p\}} \theta_j D_x g_j(z) + \bar{\mu}_p D_x g_p(z)$$

$$\Gamma(z, \theta) = \begin{pmatrix} \mathcal{L}(z, \theta) \\ H(z) \\ G_{J_0(\bar{z})}(z) \end{pmatrix}$$

und das Problem

$$\widetilde{\mathcal{P}} \quad \min_{(z,\theta) \in \widetilde{\mathcal{X}}} \quad \tilde{\Lambda}(z, \theta) = t,$$

$$\widetilde{\mathcal{X}} = \left\{(z, \theta) \in I\!R^{n+m+|J_0(\bar{z})|+1} | \Gamma(z, \theta) = 0\right\}.$$

Nach Definition gilt die Zulässigkeit des Punktes $(\bar{z}, \bar{\theta})$ für das Problem $\widetilde{\mathcal{P}}$. Rechnerische Überlegungen (siehe [60]) zeigen, daß unter der Gültigkeit der ersten drei Voraussetzungen der vorigen Definition ihre letzten zwei Bedingungen äquivalent dazu sind, daß $(\bar{z}, \bar{\theta})$ ein nichtdegenerierter kritischer Punkt für das Problem $\widetilde{\mathcal{P}}$ ist. In ihm ist damit die LICQ auch erfüllt.

Die Projektion von $\widetilde{\mathcal{X}}$ auf die Komponenten (x, t) bildet lokal um $\bar{z}$ die Menge Σ_{gc}. Daher gibt uns diese letzte Charakterisierung geometrische Information über Σ_{gc}. Der Punkt $\bar{z}$ selber ist kein kritischer Punkt, denn $w \neq 0$, aber alle Punkte von Σ_{gc} in einer Umgebung sind nichtdegenerierte kritische Punkte. Die lokale Struktur wird in der nächsten Abbildung 4 gezeigt. Die zwei charakteristischen Zahlen, von denen das lokale Verhalten des Indextupels (LI, LCI, QI, QCI) in den nichtdegerierten Punkten von Σ_{gc} abhängt, sind in diesen Fall $sign(\alpha)$, β, $sign(\delta)$ und γ, wobei $\alpha = w^T A^{-1} w$, β die Anzahl positiver Eigenwerte von A ist, $\delta = D_t \mathbf{L}(\bar{z})$, und γ die Anzahl der negativen Multiplikatoren $\bar{\mu}_j, j \in J_0(\bar{z}) \setminus \{p\}$. Im Fall $|J_0(\bar{z})| \neq 0$ ist in den beiden letzten charakteristischen Zahlen der Multiplikatorenvektor $(\bar{\lambda}_i, \bar{\mu}_j, i \in I, j \in J_0(\bar{z}))$, der die Beziehung (9) erfüllt, so festgelegt, daß $\bar{\mu}_p = 1$ für den festen Index $p \in J_0(\bar{z})$.

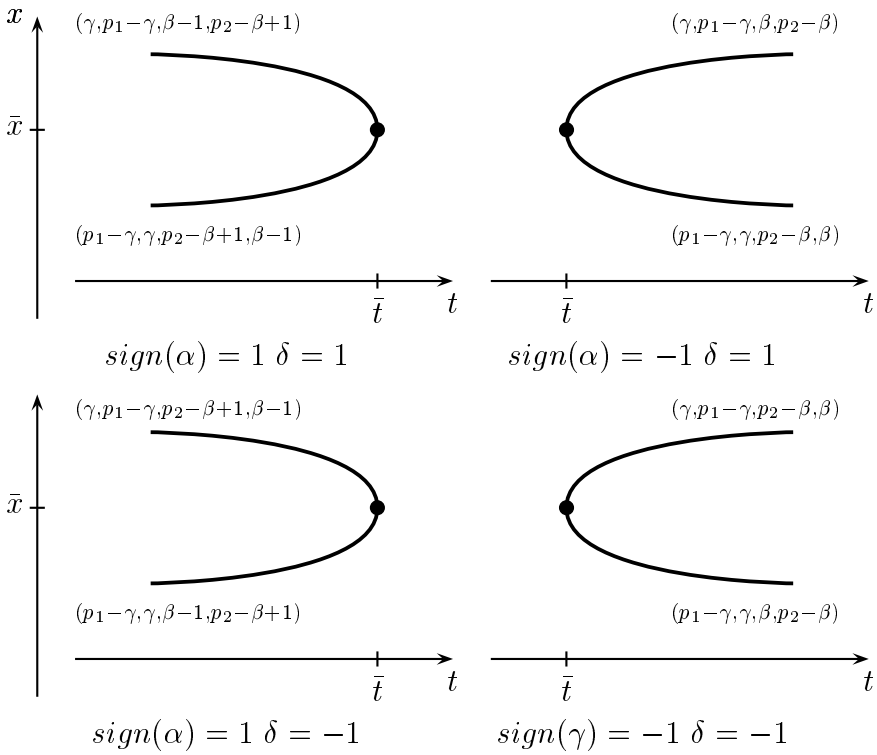
Obwohl die Struktur der Menge Σ_{gc} um die Punkte vom Typ 3 und Typ 4 ähnlich ist, haben diese Punkte ganz unterschiedliche Natur (siehe [14], [35]).

Definition 18 *(siehe [35])*
Ein verallgemeinerter kritischer Punkt $\bar{z} = (\bar{x}, \bar{t})$ ist vom Typ 5, falls folgendes gilt:

1. $m + |J_0(\bar{z})| = n + 1$.

2. *Die Vektoren $\{Dh_i(\bar{z}), Dg_j(\bar{z}), i \in I, j \in J_0(\bar{z})\}$ sind linear unabhängig (es sei bemerkt, daß die Ableitungen in allen Variablen von $\mathbb{R}^{n+1}$ berücksichtigt werden).*

3. *Ähnlich wie bei Punkten vom Typ 4 existieren auch hier, bis auf eine multiplikative Konstante, eindeutig bestimmte Multiplikatoren, die nicht alle gleich Null sind, $\bar{\lambda}_i, i \in I$ und $\bar{\mu}_j, j \in J_0(\bar{z})$, die die Beziehung (9) realisieren. Alle $\bar{\mu}_j, j \in J_0(\bar{z})$ seien verschieden von Null.*



$p_1 = |J_0(\bar{z})| \quad p_2 = n - m - |J_0(\bar{z})|$

Bild 4: Σ_{gc} in einer Umgebung von $\bar{z} \in \Sigma_{gc}^4$.

4. *Nach der zweiten Bedingung existieren eindeutig bestimmte Zahlen* $\bar{\alpha}_i$, $\bar{\beta}_j$, $i \in I$, $j \in J_0(\bar{z})$ *mit*

$$Df(\bar{z}) - \sum_{i \in I} \bar{\alpha}_i Dh_i(\bar{z}) - \sum_{j \in J_0(\bar{z})} \bar{\beta}_j Dg_j(\bar{z}) = 0.$$

Es seien die Zahlen der Form $\Delta_{jk} = \bar{\beta}_j - \bar{\beta}_k \cdot \frac{\bar{\mu}_j}{\bar{\mu}_k}$, $\forall j, k \in J_0(\bar{z})$, $j \neq k$, *verschieden von Null.*

Man kann schnell sehen (siehe [60]), daß unter der Gültigkeit der ersten drei Bedingungen die vierte äquivalent zu der folgenden ist:

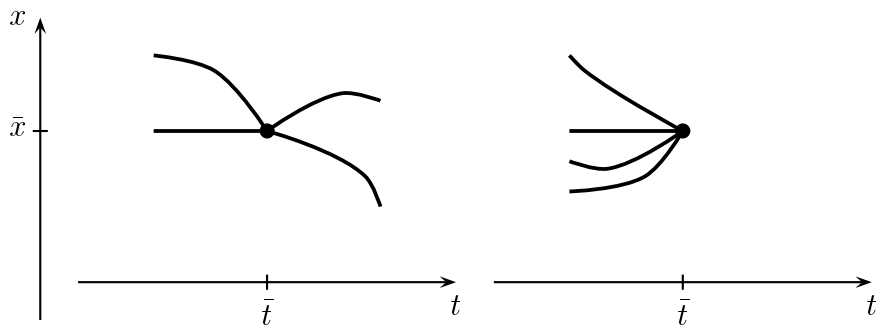
Für jede Lösung $(\theta_i, \vartheta_j, i \in I, j \in J_0(\bar{z}))$ des linearen Gleichungssystems

$$D_x f(\bar{z}) - \sum_{i \in I} \theta_i D_x h_i(\bar{z}) - \sum_{j \in J_0(\bar{z})} \vartheta_j D_x g_j(\bar{z}) = 0 \tag{10}$$

gilt, daß höchstens einer der Multiplikatoren $\vartheta_j, j \in J_0(\bar{z})$ gleich Null ist.

Da alle $\bar{\mu}_j, j \in J_0(\bar{z})$ ungleich Null sind, gibt es genau $|J_0(\bar{z})|$ Vektoren $(\theta_i, \vartheta_j, i \in I, j \in J_0(\bar{z}))$, so daß die Beziehung (10) erfüllt ist und genau ein ϑ_j verschwindet. Für jedes $p \in J_0(\bar{z})$ gibt es eine solche Lösung mit $\vartheta_p = 0$ und zu dieser assoziiert einen Zweig von Σ_{gc} mit $g_p > 0$. Der Punkt $\bar{z}$ ist kein kritischer Punkt, aber anders als für den Typ 4 ist er ein stationärer Punkt. In der Umgebung von $\bar{z}$ sind alle anderen Punkte von Σ_{gc} nichtdegenerierte kritische Punkte. Die charakteristischen Zahlen sind hier γ_j, δ_j, $j \in J_0(\bar{z})$, wobei $\gamma_j = sign(\bar{\mu}_j \cdot D_t \mathbf{L}(\bar{z}))$ und δ_j die Anzahl der negativen Zahlen der Form Δ_{kj}, $k \in J_0(\bar{z}) \setminus \{j\}$ sind. Hier ist $\mathbf{L}(z)$ genauso definiert wie in der vorigen Singularität von Typ 4. Die Struktur von Σ_{gc} wird im Bild 5 dargestellt. Es gibt $|J_0(\bar{z})|$ Zweige, die von $\bar{z}$ ausgehen, und das Indextupel (LI, LCI, QI, QCI) auf dem Zweig, wo g_k inaktiv ist, hat die Form $(\delta_k, |J_0(\bar{z})| - 1 - \delta_k, 0, 0)$.

Nach dieser detaillierten Definition der 5 spezifischen verallgemeinerten kritischen Punkte sind wir in der Lage, die Regularität im Sinne von Jongen, Jonker und Twilt (kurz JJT-Regularität) zu definieren. Wenn eine Abbildung $(f, H, G) \in C^\infty(\mathbb{R}^{n+1}, \mathbb{R}^{m+p+1})$ gegeben ist, wird auch ein einparametrisches Optimierungsproblem eindeutig festgelegt. Wir werden dieses parametrische Problem direkt mit dem Tupel (f, H, G) bezeichnen oder die Notation $P((f, H, G), t)$ benutzen. Falls eindeutig ist, zu welchem Tupel (f, H, G) ein einparametrisches Optimierungsproblem gehört, nutzen wir die



MFCQ ist erfüllt. MFCQ ist verletzt

Bild 5: Σ_{gc} in einer Umgebung von $\bar{z} \in \Sigma_{gc}^5$.

einfache Bezeichnung $P(t)$, ansonsten nutzen wir $P((f,H,G),t)$. Wir benutzen die Bezeichung $\Sigma_{gc}(f,H,G)$ für die Menge der verallgemeinerten kritischen Punkte des Problems (f,H,G). Falls eindeutig ist, welches Problem (f,H,G) behandelt wird, schreiben wir einfach Σ_{gc}. Mit Σ_{gc}^i, $i = 1,\ldots,5$ bezeichnen wir die Untermengen von Σ_{gc}, die aus den Punkten vom Typ i bestehen. Für die Einführung der JJT-Regularität definieren wir die Menge $\mathcal{F}$.

$$\mathcal{F} = \left\{ (f,H,G) \in C^\infty(\mathbb{R}^{n+1}, \mathbb{R}^{m+p+1}) \;\middle|\; \Sigma_{gc}(F,H,G) = \bigcup_{i=1}^{5} \Sigma_{gc}^i(f,H,G) \right\} \tag{11}$$

Definition 19
Ein durch die Daten $(f,H,G) \in C^\infty(\mathbb{R}^{n+1}, \mathbb{R}^{m+p+1})$ definiertes einparametrisches Optimierungsproblem heißt regulär im Sinne von Jongen, Jonker und Twilt, oder JJT-regulär, falls $(f,H,G) \in \mathcal{F}$.

Nützlich für die weiteren Untersuchungen ist die folgende leichte Umformulierung.

Definition 20
Ein einparametrisches Optimierungsproblem sei durch die Daten $(f,H,G) \in C^\infty(\mathbb{R}^{n+1}, \mathbb{R}^{m+p+1})$ gegeben, und S sei eine Teilmenge von $\mathbb{R}^{n_1+1}$.

Das Problem (f, H, G) (oder $P(t)$) ist JJT-regulär auf S, falls

$$\Sigma_{gc}(F, H, G) \cap S = \bigcup_{i=1}^{5} \left(\Sigma_{gc}^{i}(f, H, G) \cap S\right)$$

gilt. Das wird mit $P(t) \in \mathcal{F}|_S$ bezeichnet.

Hat S die Struktur $S_1 \times I\!R$, mit $S_1 \subset I\!R^n$ (oder $I\!R^n \times S_2$, mit $S_2 \subset I\!R$-), dann nutzen wir die Bezeichnung $P(t) \in \mathcal{F}|_{S_1}$ (oder $P(t) \in \mathcal{F}|_{S_2}$) statt $P(t) \in \mathcal{F}|_{S_1 \times I\!R}$ (oder $P(t) \in \mathcal{F}|_{I\!R^{n1} \times S_2}$).

Die JJT-Regularität ist eine stärkere Voraussetzung als die von Kojima und Hirabayashi (siehe [60]), andererseits ist eine genauere Unterscheidung der degenerierten Fälle vorhanden, und schärfere Einblicke in die Strukturen können dadurch gewonnen und verwendet werden. Im Rest dieses Abschnittes beschäftigen wir uns mit den wichtigsten Eigenschaften, die bezüglich dieser JJT-Regularität bekannt sind. Basierend auf der Untersuchung der lokalen Struktur von Σ_{gc} in den fünf Singularitäten und aufgrund ihrer Definition gibt uns die JJT-Regularität nähere Information über die globale Struktur der Menge Σ_{gc}, wie Satz 6 zeigt.

Satz 6 *(siehe [35])*
Es sei $(f, H, G) \in \mathcal{F}$, dann ist die Menge Σ_{gc}^1 offen und dicht in Σ_{gc}, und für $i = 2, \ldots, 5$ sind die Mengen Σ_{gc}^i diskret.

Für eine Kurvenverfolgungsmethode in der Menge Σ_{gc} ist diese letzte Aussage wichtig, denn Σ_{gc}^1 ist praktisch problemlos zu verfolgen, und die Schwierigkeiten, nämlich die Singularitäten, treten isoliert auf. Von natürlichem Interesse ist die Struktur der Menge der stationären Punkte unter der Voraussetzung $\mathcal{F}$. Dazu betrachten wir nicht nur die Menge

$$\Sigma_{stat} = \{(x, t) \,|x \text{ ist ein stationärer Punkt von } P(t)\}\,, \tag{12}$$

sondern auch ihre Abschließung $\overline{\Sigma_{stat}}$ (hier wird mit $\overline{A}$ die Abschließung der Menge A bezeichnet). Diese Menge $\overline{\Sigma_{stat}}$ stimmt mit der in [35] definierten Menge Σ_{KT} überein, und wir haben dann das in Satz 7 dargestellte Ergebnis.

Satz 7 *(siehe [35])*
Es sei $(f, H, G) \in \mathcal{F}$, dann ist die Menge $\overline{\Sigma_{stat}}$ eine eindimensionale (stückweise zweimal stetig differenzierbare) Mannigfaltigkeit mit Rand. Es gilt auch die folgende Äquivalenz: $\bar{z} \in \overline{\Sigma_{stat}}$ ist ein Randpunkt genau dann, wenn $J_0(\bar{z}) \neq \emptyset$ und MFCQ in $\bar{z}$ verletzt ist.

Für die Definition und Details der Mannigfaltigkeiten mit Rand kann man auf [26] oder [50] verweisen. In diesem Satz ist schon eine tiefgreifende Untersuchung der lokalen Struktur von Σ_{stat} in einer Umgebung der fünf Typen enthalten. Die Schwierigkeiten treten in den letzten zwei Singularitäten auf.

Der nächste Satz beleuchtet die MFCQ-Frage innerhalb von $\overline{\Sigma_{stat}}$ vollständig.

Satz 8 *(siehe [35])*
Es seien $(f, H, G) \in \mathcal{F}$ und $\bar{z} \in \overline{\Sigma_{stat}}$, dann ist die MFCQ in $\bar{z}$ genau dann verletzt, wenn eine der beiden folgenden Bedingungen erfüllt ist:

1. *$\bar{z}$ ist vom Typ 4.*
2. *$\bar{z}$ ist vom Typ 5 und alle $\bar{\mu}_j$, $j \in J_0(\bar{z})$ in der dritten Bedingung der Definition 18 des Typ 5 haben gleiches Vorzeichen.*

Eine andere interessante und wichtige verfügbare Information über die einzelnen Singularitäten ist die lokale Struktur der parametrischen zulässigen Menge

$$M(t) = \{x \in I\!R^n | H(x,t) = 0, G(x,t) \geq 0\} \tag{13}$$

in einer Umgebung der jeweiligen Typen (siehe z.B. [14]). Das liefert uns eine nützliche geometrische und topologische Sicht auf die lokale Struktur der betrachteten Optimierungsprobleme in einer Umgebung der Singularitäten.

Wenn man einen Kurvenverfolgungsprozeß betrachtet, kann man sich auf die vom Optimalitätsgesichtspunkt interessante Menge beschränken, nämlich auf Σ_{stat} oder eine Menge mit $QI = 0$. Es könnte aber auch von Interesse oder nützlich sein, so viel wie möglich der größeren Menge Σ_{gc} zu untersuchen, z.B. weil man alle möglichen stationären Punkte beschreiben will und dazu die möglichen Verbindungsstücke von $\Sigma_{gc} \setminus \Sigma_{stat}$ nutzt. In dieser Hinsicht ist wichtig zu wissen, ob eine von der Singularität ausgehende Kurve in beide Richtungen des Parameters t existiert oder nicht. Für die Punkte vom Typ 3 und Typ 4 kommt diese Eigenschaft nicht in Frage, wie ihre lokale Struktur zeigt. Beim Punkt vom Typ 2 ist die Fortsetzungsfrage innerhalb von Σ_{gc} immer positiv, aber nicht so innerhalb von Σ_{stat}.

Beim Punkt vom Typ 5 hängt diese Frage entscheidend von der Gültigkeit der MFCQ ab, sogar für die beiden Fälle Σ_{gc} und Σ_{stat}. Für Σ_{stat} ist das implizit in den Sätzen 7 und 8 beantwortet, denn Typ 5 wird als Randpunkt

untersucht. Andererseits hat die Menge Σ_{gc} immer Fortsetzungen in den Punkten vom Typ 5 unter der MFCQ, man braucht nur zu beachten, daß die charakteristischen Zahlen γ_j, die die t-Richtung der ausgehenden Zweige bestimmen, sowohl positive als auch negative Einträge besitzen. Bei der Verletzung der MFCQ kann man auch die Unmöglichkeit der Fortsetzung mittels der lokalen Struktur von $M(t)$ erkennen, denn in eine der t-Richtungen ist die zulässige Menge lokal leer (siehe [14] [33]).

Eine weitere nützliche Untersuchung der 4 Singularitäten in Bezug auf die Kurvenverfolgungsideen ist die Frage nach der Möglichkeit, in einer Singularität zu springen. Für Details kann man z.B. auf [30], [14], [18] verweisen, wir geben hier eine kurze Beschreibung der Ergebnisse.

Mit einem Kurvenverfolgungsverfahren kann man nur den Teil von Σ_{gc} oder $\overline{\Sigma_{stat}}$ beschreiben, der pfadzusammenhängend mit dem Anfangspunkt ist. Das ist eine Einschränkung, die z.B. bei der Benutzung des Einbettungsprinzips für die Lösung von Optimierungsproblemen von Bedeutung ist. Daher ist es von einem globalen Gesichtspunkt aus wichtig, wie man andere Zusammenhangskomponenten lokalisieren kann. Unter der Voraussetzung der JJT-Regularität kann man versuchen, die Information über die Singularitäten für einen Sprung auszunutzen. Ein Sprung von einer Singularität kann auch von Interesse sein, weil man keine Fortsetzung in $\overline{\Sigma_{stat}}$ oder Σ_{gc} besitzt.

Eine natürliche Möglichkeit ist die Benutzung von Abstiegsverfahren für ein festes t im Problem $P(t)$. Dazu ist notwendig, im Anfangspunkt eine zulässige Abstiegsrichtung zu erzeugen. Es sei bemerkt, daß uns die Realisierung eines solchen Abstiegsverfahrens nicht unbedingt zu einer neuen Zusammenhangskomponente bringt, sondern uns auch nur einen anderen Teil einer schon bekannten liefern kann. Man kann die Frage stellen, in welchen der verschiedenen Singularitäten der JJT-Regularität man eine solche Richtung relativ einfach und unter Benutzung der durch die Kurvenverfolgung erhaltenen Information bestimmen kann.

Bemerkung 3

In Punkten vom Typ 1 mit $LI + QCI > 0$ existiert immer eine Abstiegsrichtung, für ihre Bestimmung spielt aber der erzeugte Pfad der Kurvenverfolgung keine Rolle. Bei den Punkten vom Typ 5 ohne MFCQ ist die Möglichkeit zum Springen ausgeschlossen, denn die zulässige Menge reduziert sich lokal auf einen einzigen Punkt. Punkte vom Typ 5 mit der MFCQ sind keine Rück-

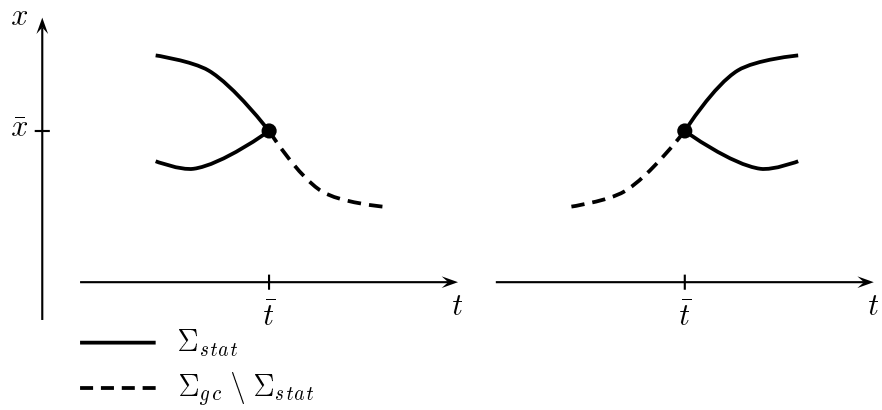


Bild 6: Ruckkehrpunkte vom Σ_{stat} um $\bar{z} \in \Sigma_{gc}^2$.

kehrpunkte, weder in $\overline{\Sigma_{stat}}$ noch in Σ_{gc}, und es ist deswegen die Sprungfrage weniger von Interesse.

Bei Punkten vom Typ 2 ist die Frage nach Sprungmöglichkeiten interessant, falls ein Rückkehrpunkt in $\overline{\Sigma_{stat}}$ vorliegt, wie im Bild 6 gezeigt.

Satz 9 *(siehe [14], [17], [18])*
Es seien

- *$\bar{z}$ ein solcher verallgemeinerter kritischer Punkt von Typ 2, daß die Menge $\overline{\Sigma_{stat}}$ um $\bar{z}$ einen Rückkehrpunkt besitzt (ohne Beschränkung der Allgemeinheit existiere $\overline{\Sigma_{stat}}$ lokal nur für $t \leq \bar{t}$),*
- *$x^0(t)$, $x^1(t) \subset \overline{\Sigma_{stat}}$ seien für $t \leq \bar{t}$ die beiden Zweige von stationären Punkten, wobei $x^1(t)$ den Zweig mit mehreren aktiven Indizes bezeichnet,*
- *$\dot{x}^0$,$\dot{x}^1$ seien die Ableitungen beider Kurven $x^0(t)$ und $x^1(t)$ im Punkt $\bar{t}$.*

Dann ist der Vektor $\vec{x}$, definiert durch

$$\vec{x} = \lim_{t \nearrow \bar{t}} \frac{x^0(t) - x^1(t)}{\|x^0(t) - x^1(t)\|} = \frac{\dot{x}^0 - \dot{x}^1}{\|\dot{x}^0 - \dot{x}^1\|},$$

eine zulässige Abstiegsrichtung von $P(\bar{t})$.

Bemerkung 4
Um einen verallgemeinerten kritischen Punkt $\bar{z}$ vom Typ 3 oder Typ 4 kann die Menge Σ_{gc} parametrisiert werden. Es sei $z(s) = (x(s), t(s))$ eine solche differenzierbare Parametrisierung, definiert auf einem offenen Intervall $(-\varepsilon, \varepsilon)$ und mit $z(0) = \bar{z}$. Der Vektor $\dot{x}(0)$, definiert durch den Limes

$$\dot{x}(0) = \lim_{s \to 0} \frac{x(s) - \bar{x}}{s},$$

ist verschieden von Null. Eine andere Möglichkeit zur Berechnung der Normalisierung von $\dot{x}(0)$ bietet sich durch den Limes

$$\vec{x} = \frac{\dot{x}(0)}{\|\dot{x}(0)\|} = \lim_{s \nearrow 0} \frac{x(s) - x(-s)}{\|x(s) - x(-s)\|}$$

an.

Andererseits gilt nach Definition, daß Σ_{gc} lokal nur auf einer der Halbebenen $\{t \leq \bar{t}\}$ oder $\{t \geq \bar{t}\}$ liegt, ohne Beschränkung der Allgemeinheit auf $\{t \leq \bar{t}\}$. Für $t < \bar{t}$ hinreichend nahe bei $\bar{t}$ gibt es genau zwei Funktionen, $s^+(t) > 0$ und $s^-(t) < 0$, durch die sich beide Zweige von Σ_{gc} mit $x^+(t) = x(s^+(t))$ und $x^-(t) = x(s^-(t))$ beschreiben lassen. Letztlich kann $\vec{x}$ auch durch den Limes

$$\vec{x} = \lim_{t \nearrow \bar{t}} \frac{x^+(t) - x^-(t)}{\|x^+(t) - x^-(t)\|}$$

berechnet werden.

Satz 10 *(siehe [30],[14], [17], [18])*
Es sei $\bar{z}$ ein verallgemeinerter kritischer Punkt vom Typ 3 (ohne Beschränkung der Allgemeinheit existiere Σ_{gc} lokal nur für $t \leq \bar{t}$). Dann ist entweder der Vektor $\vec{x}$, wie in der vorigen Bemerkung 4 definiert, oder der negative Vektor $-\vec{x}$ eine zulässige Abstiegsrichtung für $P(\bar{t})$.

Bemerkung 5
Es sei eine Parametrisierung $z(s)$ gemäß Bemerkung 4 von Σ_{gc} um einen verallgemeinerten kritischen Punkt $\bar{z}$ vom Typ 4 gewählt. Dann gilt (siehe z.B. [60]) $\vec{x}^T D_x f(\bar{z}) \neq 0$. Möglich ist, durch eine eventuelle Änderung der Richtung, in der die Kurve beschrieben wird, die Parametrisierung $z(s)$ so zu wählen, daß $\vec{x}^T D_x f(\bar{z}) < 0$. Nach Definition sind alle Punkte $z(s)$ mit $s \neq 0$ nichtdegenerierte kritische Punkte.

Vor der Untersuchung der Existenz von Sprungrichtungen bei Punkten vom Typ 4 wollen wir noch die Menge

$$\Sigma_{loc-\min} = \{(x,t)\,|x \text{ ist ein lokaler Minimalpunkt von } P(t)\} \qquad (14)$$

definieren.

Satz 11 *(siehe [30], [14], [17], [18])*
Es seien $\bar{z}$ ein verallgemeinerter kritischer Punkt vom Typ 4 (ohne Beschränkung der Allgemeinheit existiere Σ_{gc} lokal nur für $t \leq \bar{t}$) und $z(s)$ eine lokale Parametrisierung von Σ_{gc} mit $\vec{x}^T D_x f(\bar{z}) < 0$, wobei $\vec{x}$ wie in Bemerkung 4 definiert ist. Sind die Punkte des Zweiges $z(s)$, $s > 0$ nahe bei 0, nicht in $\Sigma_{loc-\min}$, dann existiert eine zulässige Abstiegsrichtung für diesen Punkt bezüglich $P(t(s))$, die uns einen Sprung ermöglicht. Andererseits ist keine Sprungmöglichkeit vorhanden, wenn diese Punkte lokale Minimalpunkte sind.

Bemerkung 6
Außer im Fall $|I| = m = n$ ist die Existenz eines Zweiges von $z(s)$, der nicht aus lokalen Minimalpunkten besteht, immer gesichert. Ob ein solcher Zweig derjenige mit den kleinsten Werten der Zielfunktion ist, entscheidet die Möglichkeit, die Abstiegsrichtungen als Sprung zu nutzen. Daraus folgt, daß für Punkte $z \in \Sigma_{gc}^4 \setminus \overline{\Sigma_{stat}}$ immer eine Sprungrichtung gefunden werden kann. Unter Berücksichtigung des Satzes 8 ist dann die MFCQ in allen Punkten vom Typ 4 ohne Sprungrichtungen verletzt.

Als weitere Eigenschaft der JJT-Regularität ist ihr generischer Charakter von Bedeutung.

Satz 12 *(siehe [35])*
$\mathcal{F}$ ist eine offene und dichte Menge in $C^3(\mathbb{R}^{n+1}, \mathbb{R}^{m+p+1})$, versehen mit der C^3-strong Topologie.

Für entscheidende Hinweise zum Beweis dieses Satzes siehe [35] und [34]. Für einen kompletten, auf dem Thomschen jet-Transversalitäts-Satz basierenden Beweis kann man auf die soeben erschienene Arbeit [63] verweisen.

Ein direkter Zugang zum Beweis dieses Satzes kann im Buch [13] gefunden werden. In diesem Zugang wird die direkte Benutzung der Existenz der

Zerlegung der Einheit bzw. der Existenz glatter Stützfunktionen gewählt (siehe z.B. [2], [31]).

Unter Benutzung dieses Ansatzes kann man in einer sogenannten lokal-globalen Prozedur (siehe z.B. [32], [2]) die Offenheit und Dichtheit der Menge $\mathcal{F}$ im Raum $C^3(I\!R^{n+1}, I\!R^{m+p+1})$ zeigen, sobald man die beiden folgenden Eigenschaften bewiesen hat.

Die erste Eigenschaft bezieht sich auf die Möglichkeit, ausgehend von beliebigen Daten $(f, H, G) \in C^\infty(I\!R^{n+1}, I\!R^{m+p+1})$ mit beliebig kleinen Störungen die Menge $\mathcal{F}$ erreichen zu können. Die Parameter der Störungen der Zielfunktion sind eine symmetrische $(n \times n)$-Matrix $A \in I\!R^{\frac{1}{2}n(n+1)}$ und ein Vektor $b \in I\!R^n$. Beide Parameter seien im Vektor $\mathcal{A} = (A, b) \in I\!R^{\frac{1}{2}n(n+1)+n}$ zusammengefaßt. Als Störungsparameter der Restriktionfunktionen nutzen wir eine $((m+p)\times n)$-Matrix C und einen Vektor $d \in I\!R^{m+p}$, zusammengefaßt in $\mathcal{C} = (C, d) \in d \in I\!R^{(n+1)(m+p)}$. Wenn ein Parametervektor $(\mathcal{A}, \mathcal{C})$ gegeben ist, definieren wir das zugehörige gestörte einparametrische Problem mit den Daten $\big(f(z) + \frac{1}{2}x^T Ax + b^T x, (H, G)(z) + Cx + d\big)$. Dieses gestörte Problem bezeichnen wir mit $P(t, (\mathcal{A}, \mathcal{C})) = (f, H, G, (\mathcal{A}, \mathcal{C})) \in C^\infty(I\!R^{n+1}, I\!R^{m+p+1})$.

Satz 13 *(Störungssatz, siehe [61])*
Es sei $(f, H, G) \in C^\infty(I\!R^{n+1}, I\!R^{m+p+1})$, dann hat jede meßbare Untermenge von

$$\Big\{(\mathcal{A}, \mathcal{C}) \in I\!R^{\frac{1}{2}n(n+1)+n+(n+1)(m+p)} \,|(f, H, G, (\mathcal{A}, \mathcal{C})) \notin \mathcal{F}\Big\} \tag{15}$$

das Lebesgue-Maß Null.

Hier ist das Lebesgue-Maß im Raum $I\!R^{\frac{1}{2}n(n+1)+n+(n+1)(m+p)}$ zu verstehen. Es ist wichtig zu bemerken, daß dieser Satz nicht die Dichtheit der Menge $\mathcal{F}$ impliziert, denn die gestörten Probleme bleiben nicht unbedingt in einer Umgebung, die von einer stetigen, positiven, im Unendlichen verschwindenden Funktion $\varepsilon(x, t)$ definiert ist. Andererseits impliziert dieser Satz die Dichtheit des Komplements von (15) und daher die Möglichkeit, die JJT-Regularität mit beliebig kleinen Störungen und ausgehend von einem beliebigen Problem erreichen zu können.

Die andere Eigenschaft der fünf verallgemeinerten kritischen Punkte der Klasse $\mathcal{F}$, die für den oben genannten Zugang zum Beweis wichtig ist, bezieht sich auf die lokale Stabilität. Gegeben seien eine feste Umgebung U, eine positive Zahl r und eine positive ganze Zahl k. Die zugehörige C^k_s-Umgebung

$V^k_{(s_1,s_2)}(U,r)$ der Abbildung $P \in C^k(I\!R^{s_1}, I\!R^{s_2})$ ist als die folgende Untermenge von $C^k(I\!R^{s_1}, I\!R^{s_2})$ definiert:

$$V^k(P,U,r) = \left\{ T \in C^k(I\!R^{s_1}, I\!R^{s_2}) \;\middle|\; \begin{array}{c} |D_\alpha P_j(x) - D_\alpha T_j(x)| < r, \forall |\alpha| \leq k \\ \forall j \in \{1,\ldots,s_2\}\,, \text{ und } \forall x \in U \end{array} \right\}.$$

Hier wird die Notation D_α ebenso wie in der vorigen Formel (3) verwendet.

Satz 14 *(siehe [13] für Typ 1 und 3, oder auch Satz 19)*
Es seien $(\bar{f}, \bar{H}, \bar{G}) \in C^3(I\!R^{n+1}, I\!R^{m+p+1})$ und $\bar{z} = (\bar{x}, \bar{t})$ fixiert. Es sei $\bar{z} \in \Sigma^i_{gc}(\bar{f}, \bar{H}, \bar{G})$ (bzw. $\bar{z} \notin \Sigma_{gc}(\bar{f}, \bar{H}, \bar{G})$) mit $i \in \{1,\ldots,5\}$. Dann existieren eine offene Umgebung $U_{\bar{z}}$ von $\bar{z}$ und eine positive Zahl $r_{\bar{z}}$, so daß für alle Abbildungen $(\tilde{f}, \tilde{H}, \tilde{G}) \in V^3((\bar{f}, \bar{H}, \bar{G}), U_{\bar{z}}, r_{\bar{z}})$ folgendes gilt: $\Sigma_{gc}(\tilde{f}, \tilde{H}, \tilde{G}) \cap U_{\bar{z}} \subset \Sigma^i_{gc}(\tilde{f}, \tilde{H}, \tilde{G}) \cup \Sigma^1_{gc}(\tilde{f}, \tilde{H}, \tilde{G})$ (bzw. $\Sigma_{gc}(\tilde{f}, \tilde{H}, \tilde{G}) \cap U_{\bar{z}} = \emptyset$).

Dieser Satz ergibt sich aus dem folgenden Lemma, das mit Stetigkeitsargumenten und einer genauen Fallunterscheidung der die fünf Typen definierenden Ungleichungen hergeleitet werden kann:

Lemma 2 *(siehe auch Lemma 6)*
Es seien $z_n \in I\!R^{n+1}$ und $(f_n, H_n, G_n) \in C^3(I\!R^{n+1}, I\!R^{m+p+1})$ Folgen mit den Konvergenzbedingungen

- $z_n \to \bar{z}$,
- $\forall |\alpha| \leq 3, D_\alpha(f_n, H_n, G_n)(z_n) \to D_\alpha(\bar{f}, \bar{H}, \bar{G})(\bar{z})$.

Es sei $\bar{z} \in \Sigma^i_{gc}(\bar{f}, \bar{H}, \bar{G})$, mit $i \in \{1,\ldots,5\}$ (bzw. $\bar{z} \notin \Sigma_{gc}(\bar{f}, \bar{H}, \bar{G})$). Dann gilt

$$\{z_n\} \cap \Sigma_{gc}(f_n, H_n, G_n) \in \Sigma^i_{gc}(f_n, H_n, G_n) \cup \Sigma^1_{gc}(f_n, H_n, G_n)$$

(bzw $z_n \notin \Sigma_{gc}(f_n, H_n, G_n)$) für hinreichend große n.

Die generischen Singularitäten und Sprünge sind für den Fall der einparametrischen semi-infiniten Optimierung untersucht worden (siehe [25], [38], [39], [63]).

Basiert auf dieser Theorie der parametrischen Optimierung gibt es auch Algorithmen und Programme für die numerische Verfolgung der Lösungskurven. Das Programpaket **PAFO**, entwickelt an der Humboldt Universität zu

Berlin (siehe [12] und [66]), basiert sich auf einer Kurvenverfolgungsmethode (genannt auch PATH III in [14]) mit Sprüngen (genannt auch JUMPI und JUMP II in [14]).

PATH III berechnet eine numerische Beschreibung einer kompakten Zusammenhangskomponent von Σ_{gc}. Dabei werden mittels einen aktive Indexmengen Strategie und einer sogenannten Prädiktor-Korrektor Prozedur eine Diskretisierung eines Intervalles $[t_A, t_B]$ und zugehörige verallgemeinerte kritische Punkte gefunden. Die Hauptschwierigkeit besteht darin, die neuen Indexmengen zu erkennen. PAFO löst dieses Problem. Keine Schwierigkeiten solcher Art treten bei Punkten vom Typ 3 oder 4 auf. In diesem Fällen ist die Indexmenge lokal konstant und die Schwierigkeiten, gegeben durch das Rückkehrphenomem, werden durch Benutzung einer anderer Variablen, die die Kurve lokal parametrisiert, beseitigt.

Es ist bekannt, daß die numerische Beschreibung einer Zusamnnehangskomponente nicht unbedingt zum Ziel führen muß, sondern daß man Sprünge zu anderen Zusammenhangskomponenten benötigt. Der Algorithmus JUMP I benutzt, falls nötig ist, alle Sprungmöglichkeiten in den Singularitäten, die in diesem Kapitel beschrieben wurden. Die Grenzen dieses Verfahrens in seinem Erfolg stehen bei den Punkten von Typ 4 und 5, bei denen eine Zusammenhangskomponente der zulässige Menge sich auf einen Punkt reduziert und dann verschwindet.

Jump II enthält auch Sprünge in gewissen verallgemeinerten kritischen Punkten von Typ 1.

Es sei bemerkt, daß auch JUMP II nicht in allen Fällen erfolgreich ist.

4 Anwendung bei Optimierungsverfahren.

Das Ziel dieses Kapitels ist die Untersuchung von verschiedenen spezifischen einparametrischen Optimierungsproblemen unter dem Gesichtspunkt der im vorigen Kapitel eingeführten JJT-Regularität. Diese Untersuchung umfaßt natürliche Fragestellungen für die verschiedenen Einbettungen, wie z.B. :

- Welche qualitativen Vorteile liefert die JJT-Regularität für die verschiedenen Einbettungen in bezug auf Konvergenzaussagen der Kurvenverfolgungsverfahren mit Sprüngen?
- Bleibt die JJT-Regularität eine generische Voraussetzung innerhalb der einparametrischen Probleme mit der Struktur der behandelten Einbettungen?

Andere Fragestellungen sind speziell für bestimmte Einbettungen interessant und werden detailliert in den folgenden Abschnitten untersucht.

4.1 Eine innere Einbettung.

In diesem Abschnitt untersuchen wir eine bestimmte Einbettung unter dem Gesichtpunkt der im vorigen Kapitel eingeführten Theorie zu einparametrischen Optimierungsproblemen. Unter einer Einbettung verstehen wir ein einparametrisches Optimierungsproblem, das zwei Optimierungsprobleme verschiedenen Schwierigkeitsgrades verbindet. Es kann z.B. ein einparametrisches Problem sein, das für $t = 0$ ein einfach lösbares Problem darstellt und für $t = 1$ ein zu lösendes Problem. Hinter dieser Herangehensweise steht die Benutzung der Kurvenverfolgung für die Berechnung von Minimalpunkten für $P(1)$. Die Kurvenverfolgung ist eine alte Technik, die auf der numerischen Lösung einparametrischer Gleichungssysteme basiert (siehe z.B. [1]).Einbettungen können manchmal Verfahren der nichtlinearen Optimierung darstellen (siehe z.B. [3], [4], [11], [15]).

Sehr wichtig bei Einbettungen als Lösungsansatz ist die Erreichbarkeit durch Lösungskurven des festgelegten interessanten Parameterwertes (z.B. für uns $t = 1$). Die Existenz einer verbindenden Lösungskurve von $t = 0$

zu $t = 1$ ist also für die Kurvenverfolgungsverfahren von großer Bedeutung. Bedingungen für die Existenz solcher Kurven sind im folgenden Satz gegeben.

Satz 15 *(siehe [10])*
Es sei $(f, H, G) \in C^3(\mathbb{R}^{n+1}, \mathbb{R}^{m+p+1})$ ein im Sinne von Kojima und Hirabayashi reguläres einparametrisches Optimierungsproblem mit folgenden zusätzlichen Eingenschaften:

C1 *Es existiert eine kompakte Menge $C \subset \mathbb{R}^n$, so daß $M(t) \subset C$ für alle $t \in [0, 1]$ gilt, wobei die parametrische zulässige Menge $M(t)$ in (13) erklärt ist.*

C2 *Das Problem $P(0)$ besitzt einen einzigen bekannten stationären Punkt $(x_0, 0)$, der ein nichtdegenerierter Minimalpunkt ist.*

C3 *Für alle $t \in [0, 1]$ und alle zulässigen Punkte in $M(t)$ gilt die MFCQ.*

Dann existiert eine stetige Kurve in $\mathbb{R}^n \times [0, 1]$, die den Punkt $(x_0, 0)$ mit einem Punkt $(x_1, 1)$ verbindet, so daß für alle Paare (x, t) auf der Kurve der Punkt x ein stationärer Punkt des Probems $P(t)$ ist.

In der wichtigen Arbeit [16] wird für kompakte Mengen, gegeben durch Restriktionen (H, G), die Äquivalenz zwischen der globalen Gültigkeit der MFCQ und der Stabilität vom Homöomorphie-Typ untersucht. Als starke Folgerung für das Einbettungsprinzip ergibt sich, daß unter den Voraussetzungen C1 und C3 aus Satz 15 die Mengen $M(0)$ und $M(1)$ den gleichen Homöomorphie-Typ besitzen. Damit wird klar, welche starke Beschränkung die für die Erfüllung der Bedingung C2 oft genutzte Voraussetzung der Konvexität von $M(0)$ darstellt.

Die Existenz der verbindenden Kurve, gesichert durch Satz 15, liefert uns die Konvergenz einer Kurvenverfolgungsprozedur ohne die Notwendigkeit von Sprüngen. Ausgehend von den Bedingungen in diesem Satz, versuchen wir in diesem Abschnitt andere hinreichende Bedingungen zu formulieren, die den Erfolg von Kurvenverfolgungsverfahren mit Sprüngen gewährleisten. Dabei achten wir darauf (siehe die Sätze im vorhergehenden Kapitel), daß das Erfülltsein der Voraussetzung C3 die Präsenz von Singularitäten verhindert, die keine Sprungmöglichkeiten bieten. Wir werden nicht eine solche

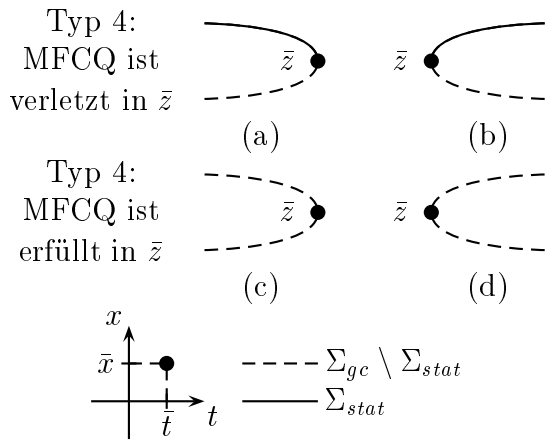


Bild 7: Σ_{gc} und Σ_{stat} in einer Umgebung vom $\bar{z} \in \Sigma_{gc}^4$

strenge Bedingung voraussetzen, sondern eine andere fordern, die die Richtung solcher unangenehmen Rückkehrpunkte festlegt. Dabei soll eine erfolgreiche Kurvenverfolgung mit Sprüngen für wachsende Werte des Parameters t theoretisch gesichert werden. Eine Einbettung mit solchen Eigenschaften wird präsentiert. Für die Benutzung der Sprünge benötigen wir die JJT-Regularität, so daß eine Untersuchung dieser Voraussetzung in Hinsicht auf ihre generische Gültigkeit erforderlich ist.

Für die Formulierung der inneren Einbettung betrachten wir Optimierungsprobleme ohne Gleichungsrestriktionen der Form $(f(x,t), G(x,t)) \in \mathcal{F} \subset C^3(\mathbb{R}^{n+1}, \mathbb{R}^{p+1})$. Da wir mit solchen Singularitäten arbeiten, in denen keine Sprungmöglichkeit vorhanden ist, müssen wir uns entsprechend der Bemerkung 3 und den Sätze 10 und 11 auf die Punkte vom Typ 4 und vom Typ 5 konzentrieren.

In unserem Fall ist $I = \emptyset$, und nach Satz 7 ist die Struktur von Σ_{gc} und Σ_{stat} in einer Umgebung eines Punktes vom Typ 4 wie in Bild 7 gezeigt.

Der Satz 11 im Zusammenhang mit der Bemerkung 6 liefert uns die genaue Charakterisierung des Falles der Punkte vom Typ 4, die keine Sprungmöglichkeit besitzen, wie in der folgenden Bemerkung beleuchtet wird.

Bemerkung 7

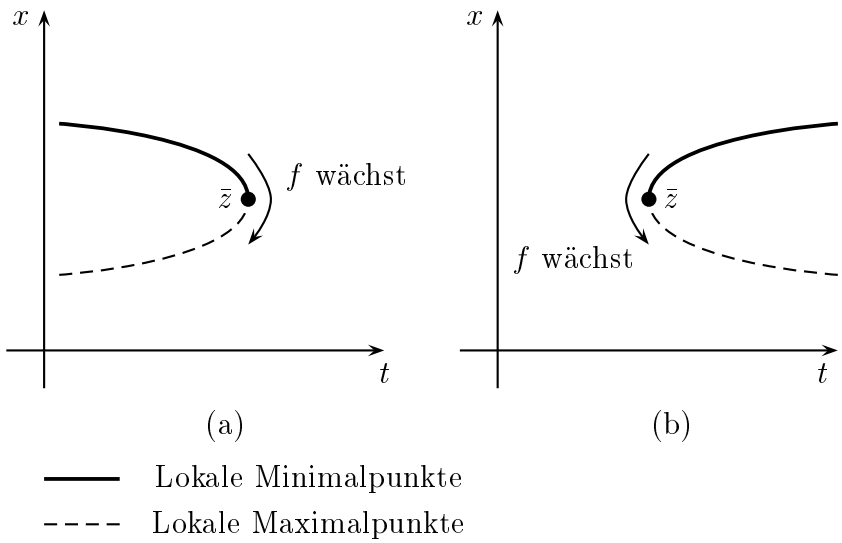


Bild 8: Punkte vom Typ 4 ohne Sprungrichtung.

Es seien $(f(x,t), G(x,t)) \in \mathcal{F}$. *In einem Punkt* $\bar{z} \in \Sigma_{gc}^4 \cap \overline{\Sigma_{loc-\min}}$ *ist genau dann keine Sprungrichtung vorhanden, wenn die* **Bedingung A** *erfüllt ist.*

Bedingung A $\exists V_{\bar{z}} \subset I\!R^{n+1}$, *so daß* $\forall (x_1, t_1) \in \Sigma_{stat} \cap V_{\bar{z}}$ *und* $\forall (x_2, t_2) \in (\Sigma_{gc} \setminus \Sigma_{stat}) \cap V_{\bar{z}}$ *die Ungleichung* $f(x_1, t_1) < f(x_2, t_2)$ *gilt.*

Mit $\overline{\Sigma_{loc-\min}}$ wird die Abschließung von $\Sigma_{loc-\min}$ bezeichnet. Solche Fälle werden im Bild 8 gezeigt:

Es sei bemerkt, daß solche Punkte vom Typ 4 ohne Sprungrichtung immer in $\overline{\Sigma_{stat}}$ enthalten sind, und laut Satz 7 ist die MFCQ verletzt.

Nach Satz 7 und Bemerkung 3 ist die lokale Struktur von Σ_{gc} und Σ_{stat} in einem Punkt vom Typ 5 wie im folgenden Bild 9 gezeigt, und interessant ist nur die Gültigkeit der MFCQ.

Die Fälle a) und b) in den Bildern 8 und 9 sind von verschiedener Bedeutung für Kurvenverfolgungsmethoden, die mit wachsenden t-Parametern arbeiten. Wir unterscheiden diese beiden Arten von Rückkehrpunkten in folgender Definition.

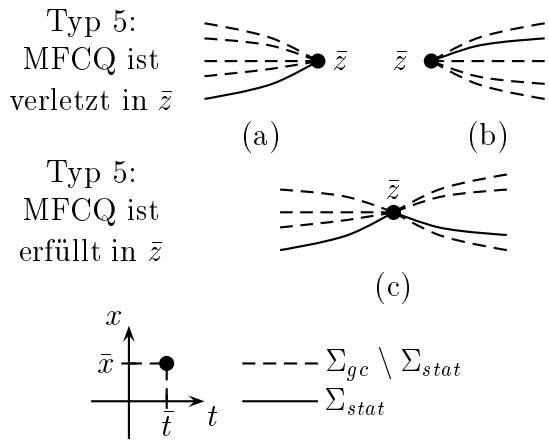


Bild 9: Σ_{gc} und Σ_{stat} in einer Umgebung vom $\bar{z} \in \Sigma_{gc}^5$.

Definition 21
Es seien $(f(z), G(z)) \in \mathcal{F}$ und $\bar{z} \in \Sigma_{gc}$. Der Punkt $\bar{z}$ heißt ein Rückkehrpunkt in negativer (bzw. positiver) Richtung, falls eine Umgebung V von $\bar{z}$ existiert mit folgender Eigenschaft:

$$\forall (x,t) \in \Sigma_{gc} \cup V \text{ gilt } t \leq \bar{t} \ (t \geq \bar{t}).$$

Die Beispiele (a) (bzw. (b)) in den Bildern 8 und 9 stehen für Rückkehrpunkte in negativer (bzw. positiver) Richtung.

Bemerkung 8
Laut Definition kann die MFCQ nur für Punkte vom Typ 4 oder Typ 5 verletzt sein. In einem Punkt $\bar{z} \in \Sigma_{gc}^4 \cup \Sigma_{gc}^5$ existiert nach den Definitionen 17 und 18 ein bis auf eine multiplikative Konstante eindeutig bestimmter Vektor $\bar{u} \in I\!R^{|J_0(\bar{z})|}$ mit $\bar{u}_j \neq 0, \ \forall j \in J_0(\bar{z})$ und

$$\sum_{j \in J_0(\bar{z})} \bar{u}_j D_x g_j(\bar{z}) = 0. \tag{16}$$

Wegen des Farkas-Lemmas (siehe z.B.[24]) ist die Verletzung der MFCQ in $\bar{z}$ äquivalent ($I = \emptyset$) zur Bedingung, daß alle $\bar{u}_j$, $j \in J_0(\bar{z})$, gleiches Vorzeichen haben.

Nach Bedingung 4 aus Definition 17 für Punkte vom Typ 4 und nach Bedingung 2 aus Definition 18 für Punkte vom Typ 5 gilt $\sum_{j\in J_0(\bar{z})} \bar{u}_j D_t g_j(\bar{z}) \neq 0$.

Definieren wir die folgende **Bedingung B** für JJT-reguläre einparametrische Probleme $(f(z), G(z))$.

Bedingung B *In jedem Punkt* $\bar{z} \in \Sigma_{gc}$, *in dem die MFCQ verletzt ist, gilt die Ungleichung*

$$\sum_{j\in J_0(\bar{z})} \bar{u}_j D_t g_j(\bar{z}) > 0,$$

wobei $\bar{u} \in I\!R^{|J_0(\bar{z})|}$, $\bar{u}_j > 0$, $\forall j \in J_0(\bar{z})$ *die Beziehung (16) realisiert.*

Kommen wir zur tatsächlichen Beziehung zwischen den Sprungmöglichkeiten, der Richtung der jeweiligen Rückkehrpunkte und den gegebenen Bedingungen.

Satz 16
Es seien $(f(z), G(z)) \in \mathcal{F}$ *und Bedingung B erfüllt. Dann sind alle Rückkehrpunkte ohne Sprungrichtungen in positiver Richtung.*

<u>Beweis:</u>
Der Beweis für Punkte vom Typ 5 ist nichts anderes als die geeignete Interpretation der charakteristischen Zahlen γ_j, $j \in J_0$. Es sei $\bar{z} \in \Sigma^5_{gc}$, in dem die MFCQ verletzt ist. Nehmen wir der Einfachheit der Bezeichnungen halber an, daß $J_0(\bar{z}) = \{1, \ldots, \bar{p}\}$ ist. Es sei $\tilde{j}$ ein fester Index in $J_0(\bar{z})$, z.B.$\tilde{j} = \bar{p}$. Die Bedingungen 1 bis 3 der Definition 18 implizieren die lineare Unabhängigkeit der Vektoren $\{D_x g_j(\bar{z}), j = 1, \ldots, \bar{p}-1\}$. Dann ist eine eindeutige Lösung $\bar{u}^{\bar{p}} = (\bar{u}^{\bar{p}}_1, \ldots, \bar{u}^{\bar{p}}_{\bar{p}-1})$ des linearen Gleichungsystems

$$D_x f(\bar{z}) - \sum_{j=1}^{\bar{p}-1} \bar{u}^{\bar{p}}_j D_x g_j(\bar{z}) = 0$$

definiert, die außerdem die vierte Bedingung derselben Definition $\bar{u}^{\bar{p}}_j \neq 0$ für $j = 1, \ldots, \bar{p}-1$ erfüllt.

Betrachten wir das folgende Gleichungssytem.

$$\begin{bmatrix} D_x f(x,t) - \sum_{j=1}^{\bar{p}-1} u_j^{\bar{p}} D_x g_j(x,t) \\ g_1(x,t) \\ \vdots \\ g_{\bar{p}-1}(x,t) \end{bmatrix} = 0 \tag{17}$$

Aus dem Satz über implizite Funktionen bekommen wir ein Interval der Form $(\bar{t}-\epsilon, \bar{t}+\epsilon)$ und eindeutig bestimmte Abbildungen

$$(x^{\bar{p}}(t), u^{\bar{p}}(t)) : (\bar{t}-\epsilon, \bar{t}+\epsilon) \to I\!R^{n+\bar{p}-1|}$$

mit den Eigenschaften $x^{\bar{p}}(\bar{t}) = \bar{x}$, $u^{\bar{p}}(\bar{t}) = \bar{u}^{\bar{p}}$, und für jedes $t \in (\bar{t}-\epsilon, \bar{t}+\epsilon)$ ist der Vektor $(x^{\bar{p}}(t), u^{\bar{p}}(t), t)$ eine Lösung von (17).

Wenn die Zulässigkeit $g_{\bar{p}}(x^{\bar{p}}(t), t) \geq 0$ erfüllt ist, dann ist $(x^{\bar{p}}(t), t)$ ein verallgemeinerter kritischer Punkt. Aus der linearen Unabhängigkeit der Vektoren $\{Dg_j(\bar{z}), j = 1, \ldots, \bar{p}\}$ folgt die Ungleichung

$$\frac{d}{dt} g_{\bar{p}}(x^{\bar{p}}(\bar{t}), \bar{t}) \neq 0.$$

Das heißt aber nichts anderes, als daß die Kurve $(x^{\bar{p}}(t), t)$ in der Menge Σ_{gc} entweder für $t \in (\bar{t}, \bar{t}+\epsilon)$ oder für $t \in (\bar{t}-\epsilon, \bar{t})$ enthalten ist. Die Menge Σ_{gc} um $\bar{z}$ besteht aus den zulässigen Teilen der Kurven $(x^{\tilde{j}}(t), t)$ für $\tilde{j} \in J_0(\bar{z})$. Also ist es hinreichend die Beziehung

$$\frac{d}{dt} g_{\tilde{j}}(x^{\tilde{j}}(\bar{t}), \bar{t}) > 0$$

für alle $\forall\ \tilde{j} \in J_0(\bar{z})$ zu zeigen. Zeigen wir diese Ungleichung für $\tilde{j} = \bar{p}$.

Dazu berechnen wir diese Ableitung:

$$\frac{d}{dt} g_{\bar{p}}(x^{\bar{p}}(\bar{t}), \bar{t}) = D_x g_{\bar{p}}(\bar{x}, \bar{t})\ \dot{x}^{\bar{p}}(\bar{t}) + D_t g_{\bar{p}}(\bar{x}, \bar{t}). \tag{18}$$

Nach Definition von $\dot{x}^{\bar{p}}(\bar{t})$ folgt folgende Gleichung für alle $j = 1, \ldots, \bar{p}-1$

$$D_x g_j(\bar{z}) \dot{x}^{\bar{p}}(\bar{t}) = -D_t g_j(\bar{z}). \tag{19}$$

Die Existenz von $\bar{u}_j > 0$, $j \in J_0(\bar{z})$ mit

$$\sum_{j=1}^{\bar{p}} \bar{u}_j D_x g_j(\bar{z}) = 0. \tag{20}$$

ergibt sich nach Bemerkung 8 und wegen der Verletzung der MFCQ.

Daraus folgt durch Multplikation mit $\dot{x}^{\bar{p}}(\bar{t})$ und unter Berücksichtigung der Beziehung (19):

$$-\sum_{j=1}^{\bar{p}-1} \bar{u}_j D_t g_j(\bar{z}) + \bar{u}_{\bar{p}} D_x g_{\bar{p}}(\bar{z}) \dot{x}^{\bar{p}}(\bar{t}) = 0.$$

Das Einsetzen dieser Beziehung in die Gleichung (18) ergibt

$$\frac{d}{dt} g_{\bar{p}}(x^{\bar{p}}(\bar{t}), \bar{t}) = \bar{u}_{\bar{p}} \left(\sum_{j \in J_0(\bar{z})} \bar{u}_j D_t g_j(\bar{z}) \right).$$

Bedingung B und die Ungleichungen ($\bar{u}_{\bar{p}} > 0$) implizieren dann die gesuchte Ungleichung

$$\frac{d}{dt} g_{\bar{p}}(x^{\bar{p}}(\bar{t}), \bar{t}) > 0.$$

Kommen wir jetzt zu den Punkten vom Typ 4.

Es sei $\bar{z} \in \Sigma_{gc}^4$ ohne Sprungrichtung. Bemerkung 6 impliziert die Verletzung der MFCQ in $\bar{x} \in M(\bar{t})$. Es sei wieder ein Vektor $\bar{u} \in I\!R^{\bar{p}}$, wie in der Bemerkung 8, der die Beziehung 16 realisiert, mit $\bar{u}_j > 0$, $\forall j \in J_0(\bar{z}) = \{1, \ldots, \bar{p}\}$ ausgewählt. Wie im vorigen Teil des Beweises sei $u^{\bar{p}} = (u_1, \ldots, u_{\bar{p}-1})$. Nach Bemerkung 2 gilt für $\bar{u}$ folgende Bedingung: Der Vektor $(\bar{z}, \bar{\theta}) = (\bar{x}, \bar{t}, 0, \bar{u}^{\bar{p}})$ ist ein nicht degenerierter kritischer Punkt des Problems

$$\widetilde{\mathcal{P}} \quad \min \quad \tilde{\Lambda}(z, \theta) = t$$

$$\widetilde{\mathcal{X}} = \{(z, \theta) \in I\!R^{n+\bar{p}+1} | \Gamma(z, \theta) = 0\},$$

wobei

$$\mathcal{L}(z, \theta) = \theta_0 f(z) + \sum_{j=1}^{\bar{p}-1} \theta_j D_x g_j(z) + \bar{u}_p D_x g_p(z),$$

$$\Gamma(z, \theta) = \begin{pmatrix} \mathcal{L}(z, \theta) \\ G_{J_0(\bar{z})}(z) \end{pmatrix}.$$

Das impliziert die drei folgenden Bedingungen:

1.

$$rank\left(D\Gamma(\bar{z},\bar{\theta})\right) = n + \bar{p}.$$

2. $\exists \lambda_j,\ j = 1, \ldots, n+\bar{p}$, so daß $\lambda = (\lambda_1, \ldots, \lambda_{n+\bar{p}}) \neq 0$ und

$$D\tilde{\Lambda}(\bar{z},\bar{\theta}) = \lambda^T D\Gamma(\bar{z},\bar{\theta}).$$

3. Die Zahl

$$D^2\tilde{L}(\bar{z},\bar{\theta})|_{T_{(\bar{z},\bar{\theta})}\widetilde{\mathcal{X}}} \tag{21}$$

ist verschieden von Null, wobei $\tilde{L}$ die Lagrange-Funktion bezüglich $\widetilde{\mathcal{P}}$ bezeichnet.

Die obige Bedingung 1 impliziert, daß die Menge $\widetilde{\mathcal{X}}$ eine eindimensionale differenzierbare Mannigfaltigkeit um den Punkt $(\bar{z},\bar{\theta})$ ist.Die Punkte vom Typ 4 sind Rückkehrpunkte, deren Richtung vom Vorzeichen der Zahl (21) abhängt. Ist diese Zahl positiv (bzw. negativ), so hat $\bar{z}$ eine positive (bzw. negative) Stellung. Es soll also gezeigt werden, daß (21) positiv ist. Dazu schreiben wir die Jacobi-Matrix von Γ in $(\bar{z},\bar{\theta})$.

$$D\Gamma(\bar{z},\bar{\theta}) = \begin{pmatrix} D_x^2\mathcal{L}(\bar{z},\bar{\theta}) & D_{xt}^2\mathcal{L}(\bar{z},\bar{\theta}) & D_x^T f(\bar{z}) & -D_x^T g_1(\bar{z}) \cdots - D_x^T g_{\bar{p}-1}(\bar{z}) \\ D_x g_1(\bar{z}) & D_t g_1(\bar{z}) & & \\ \vdots & \vdots & 0 & 0 \\ D_x g_{\bar{p}}(\bar{z}) & D_t g_{\bar{p}}(\bar{z}) & & \end{pmatrix}$$

Aus **Bedingung B** wissen wir, daß

$$\sum_{j=1}^{\bar{p}} \bar{u}_j D_t g_j(\bar{z}) = -D_t\mathcal{L}(\bar{z},\bar{\theta}) > 0 \tag{22}$$

gilt. Andererseits hat der Gradient von $\tilde{\Lambda}$ die Struktur

$$\begin{aligned} D_t\tilde{\Lambda}(\bar{z},\bar{\theta}) &= 1 \\ D_{(x,\theta)}\tilde{\Lambda}(\bar{z},\bar{\theta}) &= 0_{n+\bar{p}}. \end{aligned}$$

Der Lagrange-Vektor kann dann berechnet werden, und aus der Beziehung

$$D\tilde{\Lambda}(\bar{z},\bar{\theta}) = \frac{1}{-D_t\mathcal{L}(\bar{z},\bar{\theta})}(0_n,\bar{u})D\Gamma(\bar{z},\bar{\theta})$$

folgt, daß der eindeutig bestimmte Lagrange-Vektor gleich $\lambda = (0_n, \bar{u})$ ist. Die Lagrange-Funktion $\hat{L}$ bezüglich $\widetilde{\mathcal{P}}$ im Punkt $(\bar{z},\bar{\theta})$ ergibt sich jetzt zu

$$\begin{aligned} \tilde{L}(z,\theta) &= t + \frac{1}{D_t\mathcal{L}(\bar{z},\bar{\theta})}\sum_{j=1}^{\bar{p}} \bar{u}_j g_j(z) \\ &= t - \frac{1}{D_t\mathcal{L}(\bar{z},\bar{\theta})}\mathcal{L}(\bar{z},\bar{\theta}). \end{aligned} \tag{23}$$

Mit $(w_x, w_t, w_{u_0}, w_{u^{\bar{p}}}) \in I\!R^{n+\bar{p}+1}$ bezeichnen wir einen Vektor, der den Tangentialraum vom $\widetilde{\mathcal{X}}$ an dem Punkt $(\bar{z},\bar{\theta})$ linear erzeugt. Nach Definition gilt dann:

$$D\Gamma(\bar{z},\bar{\theta})\begin{pmatrix} w_x \\ w_t \\ w_{u_0} \\ w_{u^{\bar{p}}} \end{pmatrix} = 0_{n+\bar{p}}. \tag{24}$$

Aus der Form der Ableitung $D\Gamma(\bar{z},\bar{\theta})$ erkennt man sofort, daß $w_t = 0$ und $w_{u_0} \neq 0$ sind. Multiplizieren wir die Gleichung (24) mit dem Vektor $(w_x, 0_{\bar{p}})$, so erhalten wir die folgende Beziehung

$$w_x^T\ D_x^2\mathcal{L}(\bar{z},\bar{\theta})\ w_x + w_x^T\ D_x^T f(\bar{z}) w_{u_0} = 0. \tag{25}$$

Aus (23) bemerkt man leicht, daß die Lagrange-Funktion $\tilde{L}(z,\theta)$ bezüglich $\widetilde{\mathcal{P}}$ im Punkt $(\bar{z},\bar{\theta})$ unabhängig von den Variablen $u^{\bar{p}} \in I\!R^{\bar{p}-1}$ und u_0 ist. Daraus folgt, unter Berücksichtigung der Gleichung $w_t = 0$, die folgende Darstellung für (21):

$$(w_x, w_t, w_{u_0}, w_{u^{\bar{p}}})^T D^2\tilde{L}(\bar{z},\bar{\theta})\begin{pmatrix} w_x \\ w_t \\ w_{u_0} \\ w_{u^{\bar{p}}} \end{pmatrix} = -\frac{w_x^T D_x^2\mathcal{L}(\bar{z},\bar{\theta}) w_x}{D_t\mathcal{L}(\bar{z},\bar{\theta})}.$$

Durch Einsetzen von (25) in die obige Formel folgt dann:

$$D^2\tilde{L}(\bar{z},\bar{\theta})|_{T_{(\bar{z},\bar{\theta})}\widetilde{\mathcal{X}}} = \frac{w_x^T D_x^T f(\bar{z}) w_{u_0}}{D_t\mathcal{L}(\bar{z},\bar{\theta})}. \tag{26}$$

Nach Definition von $(w_x, w_t, w_{u_0}, w_{u^{\bar{p}}})$ als Element des Tangentialraumes folgt die Existenz einer Abbildung $(x(\tau), t(\tau), u_0(\tau), u^{\bar{p}}(\tau))$, in einer Umgebung $(-\epsilon, \epsilon) \subset I\!R$ definiert, mit folgenden Eigenschaften:

- $\Gamma(x(\tau), t(\tau), u_0(\tau), u^{\bar{p}}(\tau)) = 0$, for all $\tau \in (-\epsilon, \epsilon)$,
- $(x(0), t(0), u_0(0), u^{\bar{p}}(0)) = (\bar{x}, \bar{t}, 0, \bar{u}^{\bar{p}}) = (\bar{z}, \bar{\theta})$,
- $(\dot{x}(0), \dot{t}(0), \dot{u}_0(0), \dot{u}^{\bar{p}}(0)) = (w_x, w_t, w_{u_0}, w_{u^{\bar{p}}})$.

Wir wissen, daß $w_{u_0} \neq 0$. Nach (26) folgt, daß $w_x^T D_x^T f(\bar{z})$ auch ungleich Null ist. Es bleibt zu zeigen: $w_x^T D_x^T f(\bar{z}) w_{u_0} < 0$. Nehmen wir an, daß $w_{u_0} = -1$. Der andere Fall ist ähnlich zu diesem. Folgende Gleichung gilt für alle $\tau \in (-\epsilon, \epsilon)$:

$$u_0(\tau) D_x^T f(x(\tau), t(\tau)) - \sum_{j=1}^{\bar{p}-1} u_j(\tau) D_x^T g_j(x(\tau), t(\tau)) - \bar{u}_{\bar{p}} D_x^T g_{\bar{p}}(x(\tau), t(\tau)) = 0. \tag{27}$$

Da $w_{u_0} = -1$ und $\bar{u}_j > 0,\ j = 1, \ldots, \bar{p}$ sind, gelten folgenden Ungleichungen für $\tau < 0$.

$$\begin{aligned} u_0(\tau) &> 0 \\ \frac{u_j(\tau)}{u_0(\tau)} &> 0,\ j =, \ldots, \bar{p} - 1 \\ \frac{u_{\bar{p}}}{u_0(\tau)} &> 0 \end{aligned}$$

Diese Beziehungen und die Gleichung (27) implizieren, daß die Punkte $(x(\tau), t(\tau))$ stationär bezüglich $P(t(\tau))$ für $\tau < 0$ sind. Da $(\bar{x}, \bar{t})$ ein Punkt vom Typ 4 ohne Sprungrichtungen ist, sind für $\tau < 0$, die Punkte $(x(\tau), t(\tau))$ lokale Minimalpunkte (stationäre Punkte) und, für $\tau > 0$, lokale Maximalpunkte von $(P(t(\tau))$.

Bedingung A impliziert, daß

$$D_t f(x(0), t(0)) = w_x^T D_x^T f(\bar{z}) > 0.$$

Nach (22) und (26) kommt man dann zur Beziehung $D^2 \tilde{L}(\bar{z}, \bar{\theta})|_{T_{(\bar{z}, \bar{\theta})} \tilde{\mathcal{X}}} > 0$.
□

Eine Einbettung, die die Bedingung B erfüllt, schlagen wir im folgenden vor. Die zulässige einparametrische Menge $M(t)$ dieser Einbettung liegt für $t < 1$ im Inneren der zulässigen Menge des zu lösenden Problems $P(1)$. Für die Feststellung der Eigenschaften dieser Einbettung nutzen wir den vorigen Satz 16.

Es seien die abbildungen $(f, G) = (f, g_1, \ldots, g_p) \in C^3(I\!R^n, I\!R^{p+1})$ fixiert, die das Problem

$$P(f, g) \quad \min \{f \mid g_j(x) \geq 0, j \in J\} \tag{28}$$

mit $J = \{1, \ldots, p\}$ definieren.

Ein Punkt $x \in I\!R^n$ mit $g_j(x) > 0$ für alle $j \in J$ heißt ein innerer Punkt des Problems $P(f, g)$. Es sei x^0 ein solcher innerer Punkt von (28) und $q \in I\!R^p$ sei so fixiert, daß $g_j(x^0) > q_j > 0$ für alle $j \in J$ gilt.

Definieren wir das folgende einparametrische Optimierungsproblem, das von den vorigen Daten abhängig ist.

$$P^{int}_{(x^0,q)}(f, G, t) \quad \min \left\{ f^{int}(f, x, t) \mid g^{int}_j(g, x, t) \geq 0, j \in J \right\}, \tag{29}$$

Dabei stelle

$$f^{int}(f, x, t) = tf(x) + (1-t)\|x - x^0\|^2 \tag{30}$$

$$g^{int}_j(g, x, t) = g_j(x) - (1-t)q_j \tag{31}$$

dar.

Für jede $j \in J$ gilt die Ungleichung

$$D_t g^{int}_j(g, x, t) = q_j > 0.$$

Die Bedingung B ergibt sich dann für dieses parametrische Problem. Für das Problem $P^{int}_{(x^0,q)}(f, G, 0)$ ist der Punkt x^0 ein nichtdegenerierter kritischer Punkt, da die folgenden drei Bedingungen erfüllt sind.

$$g^{int}_j(g, x^0, 0) > 0, \ \forall j \in J \tag{32}$$

$$D_x f^{int}(f, x^0, 0) = 0 \tag{33}$$

$$D^2_x f^{int}(f, x^0, 0) = 0 \tag{34}$$

Der Punkt $(x^0, 0)$ kann dann als Anfangspunkt einer Kurvenverfolgungsprozedur für das Problem $P^{int}_{(x^0,q)}(f, G, t)$ benutzt werden. Unter der Regularität $P^{int}_{(x^0,q)}(f, G, t) \in \mathcal{F}$ impliziert die direkte Benutzung von Satz 16 die Bemerkung 9.

Bemerkung 9
Eine Kurvenverfolgungsprozedur mit Sprüngen, ausgehend von $(x^0, 0)$ und angewandt auf das Problem $P^{int}_{(x^0,q)}(f, G, t)$, würde nicht den Wert $t = 1$ erreichen können, falls ein Sprung, d.h. eine existierende zulässige Abstiegsrichtung, bei einem Abstiegsverfahren nicht erfolgreich wäre. Wenn z.B. der zulässige Bereich des Problems $P(f, g)$ kompakt ist, wäre eine Kurvenverfolgungsprozedur mit Sprüngen immer erfolgreich.

Nach dieser Bemerkung bringen wir die grundlegende Voraussetzung

$$P^{int}_{(x^0,q)}(f, G, t) \in \mathcal{F}$$

zur Diskussion. In verschiedenen Arbeiten über bestimmte Einbettungen (siehe z.B. [3], [4], [11], [15]) hat man auch die JJT-Regularität vorausgesetzt. Diese Annahme wurde auch dort analysiert und Störungssätze, ausgehend von den durch die Einbettung entstandenen einparametrischen Problemen, wurden vorgestellt. Solche Ergebnisse geben eine Anwort auf die Frage: Wie kann man ein einparametrisches Problem (gegeben durch eine Einbettung) so stören, daß die JJT-Regularität zustande kommt? Eine direkte Anwort auf diese Frage kann man aus Satz 13 erhalten.

Für die Einbettung $P^{int}_{(x^0,q)}(f, G, t)$ sind wir an der Fragestellung interessiert: Wie geeignet bezüglich der Daten $(f, G) \in C^3(\mathbb{R}^n, \mathbb{R}^{p+1})$ ist die Vorausstezung, daß das zugehörige einparametrisches Problem $P^{int}_{(x^0,q)}(f, G, t)$ zur Klasse $\mathcal{F}$ gehört? Die Eigenschaften der Menge $\mathcal{F}$, offen und dicht zu sein, bringen keine Antwort auf diese Problematik, denn die spezifische Art und Weise, in der die Daten $(f, G) \in C^3(\mathbb{R}^n, \mathbb{R}^{p+1})$ von der Einbettung für die Konstruktion des einparametrischen Problems benutzt werden, spielt jetzt eine wichtige Rolle.

Mit unserer Einbettung ist eine Abbildung vom Raum $C^3(\mathbb{R}^n, \mathbb{R}^{p+1})$ in den Raum $C^3(\mathbb{R}^{n+1}, \mathbb{R}^{p+1})$ assoziiert, die von den Vektoren x_0 und q abhängt und mit $T_{(x^0,q)}$ bezeichnet wird. Sie ist als

$$T_{(x^0,q)} : C^3(\mathbb{R}^n, \mathbb{R}^{p+1}) \mapsto C^3(\mathbb{R}^{n+1}, \mathbb{R}^{p+1})$$

mit

$$T_{(x^0,q)}(f, g_1, \ldots, g_p) := \begin{pmatrix} tf + (1-t)\|x - x^0\|^2 \\ g_1 - (1-t)q_1 \\ \vdots \\ g_p - (1-t)q_p \end{pmatrix}$$

definiert.

Das einparametrische Problem, definiert mit den Daten $T_{(x^0,q)}(f,G)$, ist nichts anderes als die Einbettung $P^{int}_{(x^0,q)}(f,G,t)$. Interessant ist dann die Untersuchung der Menge $T^{-1}_{(x^0,q)}(\mathcal{F})$. Es gilt Satz 17.

Satz 17
Es sei $(x^0,q) \in I\!R^n \times I\!R^q$ fest. Dann ist die Menge $T^{-1}_{(x^0,q)}(\mathcal{F})$ eine generische Untermenge von $C^3(I\!R^n, I\!R^{p+1})$, versehen mit der C^3_s-(or Whitney-) Topologie.

<u>Beweis:</u>
Nach Definition 20 kann die Klasse $\mathcal{F}$ als ein abzählbarer Durchschnitt wie folgt geschrieben werden:

$$\mathcal{F} = \cap_{n=1}^{\infty} \mathcal{F}|_{[-n,n]}.$$

Wir zeigen, daß die Menge $T^{-1}_{(x^0,q)}(\mathcal{F}|_{[-n,n]}) \subset C^3(I\!R^n, I\!R^{p+1})$ offen und dicht für jede $n = 1, 2, ...$ ist. Das Ergebnis des Satzes folgt dann aus den Beziehungen

$$\cap_{n=1}^{\infty} T^{-1}_{(x^0,q)}(\mathcal{F}|_{[-n,n]}) = T^{-1}_{(x^0,q)}(\cap_{n=1}^{\infty} \mathcal{F}|_{[-n,n]}) = T^{-1}_{(x^0,q)}(\mathcal{F}).$$

Es sei $n \in \{1, 2, ...\}$ fixiert. Wie im vorigen Kapitel bemerkt, kann man, um die Offenheit und Dichtheit der Menge $T^{-1}_{(x^0,q)}(\mathcal{F}|_{[-n,n]})$ zu zeigen, geeignete Versionen der Sätze 13 und 14 beweisen. Ähnlich wie im Fall der Klasse $\mathcal{F}$ liefert die Standard-lokal-global-Technik die genannten Eigenschaften von $T^{-1}_{(x^0,q)}(\mathcal{F}|_{[-n,n]})$. Wir haben dann zwei Behauptungen zu zeigen.

Behauptung 1 *(Störungssatz)*
Es sei $(\bar{f}, \bar{G}) \in C^3(I\!R^n, I\!R^{p+1})$ fest, dann hat jede meßbare Untermenge der Menge

$$\left\{ (\mathcal{A}, \mathcal{C}) \in I\!R^{\frac{1}{2}n(n+1)+n+p(n+1)} \left| T_{(x^0,q)} \left[(\bar{f}, \bar{G}, (\mathcal{A}, \mathcal{C})) \right] \notin \mathcal{F}|_{[-n,n]} \right. \right\} \tag{35}$$

das Lebesgue-Maß Null.

Dabei sind $(\mathcal{A}, \mathcal{C}) = ((A,b),(C,d))$ und $(\bar{f}, \bar{G}, (\mathcal{A}, \mathcal{C})) = (\bar{f} + 0.5x^T Ax + b^T x, \bar{G} + Cx + d)$.

Behauptung 2
Es seien $(\bar{f}, \bar{G}) \in C^3(I\!R^n, I\!R^{p+1})$ *fest und* $\bar{x}$ *ein Punkt mit* $T_{(x^0,q)}(\bar{f}, \bar{G}) \in \mathcal{F}|_{\{\bar{x}\}\times[-n,n]}$. *Dann existieren eine offene Umgebung* $U_{\bar{x}}$ *von* $\bar{x}$ *und eine positive Zahl* $r_{\bar{x}}$, *so daß* $T_{(x^0,q)}(\tilde{f}, \tilde{G}) \in \mathcal{F}|_{U_{\bar{x}}\times[-n,n]}$ *für alle Abbildungen* $(\tilde{f}, \tilde{G}) \in V^3((\bar{f}, \bar{G}), U_{\bar{x}}, r_{\bar{x}})$.

Beweis zur Behauptung 1:
Das vom Störungsparameter $(\mathcal{A}, \mathcal{C})$ abhängige Problem $T_{(x^0,q)}\left[(\bar{f}, \bar{G}, (\mathcal{A}, \mathcal{C}))\right]$ hat die Gestalt

$$\begin{array}{ll} P_{(\mathcal{A},\mathcal{C})}(t) & \\ \quad \min & t\bar{f} + \|x - x^0\|^2 + t(0.5x^T Ax + b^T x) \\ \quad bzgl. \ : & \bar{g}_j - (1-t)q_j + c_j^T x + d_j \geq 0, \ \forall j \in J \ . \end{array}$$

Die Ähnlichkeit dieses Problems mit dem von Satz 13 läßt uns mit praktisch demselben Beweis das folgende Ergebnis sichern:
Jede meßbare Untermenge der Menge

$$\left\{(\mathcal{A}, \mathcal{C}) \in I\!R^{\frac{1}{2}n(n+1)+n+p(n+1)} \,\middle|\, T_{(x^0,q)}\left[(\bar{f}, \bar{G}, (\mathcal{A}, \mathcal{C}))\right] \notin \mathcal{F}|_{I\!R\setminus\{0\}}\right\} \tag{36}$$

hat das Lebesgue-Maß Null.

Der Fall $t = 0$ muß herausgenommen werden, da in der Zielfunktion der Störungsteil mit t multipliziert wird. Für die Untersuchung dieses Falles $t = 0$ beweisen wir, daß die Menge

$$\left\{\mathcal{C} \in I\!R^{p(n+1)} \,\middle|\, (\|x - x^0\|, \bar{G} - q + Cx + d) \notin \mathcal{F}_{OP}\right\} \tag{37}$$

das Lebesgue-Maß Null hat.

Das heißt, daß für fast alle $\mathcal{C}$ folgende Bedingungen bezüglich des Problems (nicht einparametrisch) $(\|x - x^0\|, \bar{G} - q + Cx + d)$ gelten:

1. In allen zulässigen Punkten gilt die LICQ.

2. Alle kritischen Punkte sind nichtdegeneriert.

Es seien $J_0 \subset J$ und $J_1 \subset J_0$ zwei Indexmengen. Betrachten wir die Abbildungen

$$K_1(x, \mathcal{C}) = \bar{G}_{J_0}(x) - q_{J_0} + C_{J_0}x + d_{J_0}$$

und

$$K_2(x,u,\mathcal{C}) = \begin{pmatrix} 2(x-x^0) - \sum_{j\in J_0} u_j(D_x\bar{g}_j(x) - c_j^T) \\ K_1(x,\mathcal{C}) \\ u_{J_1} \end{pmatrix},$$

wobei $u \in I\!R^{|J_0|}$.

Die partiellen Ableitungen nach d_{J_0} von $K_1(x,\mathcal{C})$ bilden eine Einheitsmatrix, und daraus folgt, daß 0 ein regulärer Wert von K_1 ist. Nach dem parametrisierten Satz von Sard (Satz 4-2) gilt diese Eigenschaft für fast alle $\mathcal{C}$. Es sei bemerkt, daß diese Regularität von 0 für die Abbildungen $K_1(.,\mathcal{C})$ für fast alle $\mathcal{C}$ die Gültigkeit der LICQ in solchen zulässigen Punkten mit der aktiven Indexmenge J_0 impliziert. Da nur endlich viele mögliche Indexmengen der Form J_0 gewählt werden können, folgt die vorhergehende Bedingung 1 für das Problem mit den Daten $(\|x-x^0\|, \bar{G} - q + Cx + d)$.

Nimmt man die partiellen Ableitungen nach d_{J_0}, u_{J_1} und nach c_k (bzw. nach x), falls $u \neq 0$ und z.B. $u_k \neq 0$, mit $k \in J_0$ (bzw. falls $u = 0$), erhält man eine reguläre Matrix. Mit denselben Argumenten wie für K_1 ist abzuleiten, daß für fast alle $\mathcal{C}$ jede mögliche Wahl von $J_1 \subset J_0 \subset J$ eine Abbildung $K_2(.,.,\mathcal{C})$ mit 0 als regulärem Wert liefert. Das heißt aber einerseits, $J_1 = \emptyset$ (ND1), und andererseits, die Jacobi-Matrix von $K_2(.,.,\mathcal{C})$ ist regulär (also ND2). Daraus folgt die zweite Bedingung für das Problem mit den Daten $(\|x-x^0\|, \bar{G} - q + Cx + d)$.

Es sei bemerkt, daß die Menge (37) meßbar mit dem Lebesgue-Maß Null ist. Daher ist die Menge

$$\left\{(\mathcal{A},\mathcal{C}) \in I\!R^{\frac{1}{2}n(n+1)+n+p(n+1)} \,\middle|\, T_{(x^0,q)}\left[(\bar{f},\bar{G},(\mathcal{A},\mathcal{C}))\right] \notin \mathcal{F}|_{\{0\}}\right\} \tag{38}$$

auch meßbar und mit verschwindendem Lebesgue-Maß.

Jede meßbare Teilmenge von (35) kann man dann in zwei meßbare Teilmengen zerlegen. Der eine Teil ist ihr Durchschnitt mit (38), der andere Teil ist der Rest. Der Durchschnitt mit (38) hat das Lebesgue-Maß Null, da die Menge (38) selbst das Lebesgue-Maß Null hat. Der andere Teil ist in der Menge (36) enthalten, und hat dann wegen seiner Meßbarkeit das Lebesgue-Maß Null.

Beweis zur Behauptung 2:

Es sei die Behauptung 2 für das Tupel $(\bar{f},\bar{G}) \in C^3(I\!R^n, I\!R^{p+1})$, und der Punkt $\bar{x} \in I\!R^n$, mit $T_{(x^0,q)}(\bar{f},\bar{G}) \in \mathcal{F}|_{\{\bar{x}\}\times[-n,n]}$, nicht erfüllt. Dann existieren Folgen $x_k \in I\!R^n$ und $(f^k, G^k) \in C^3(I\!R^n, I\!R^{p+1})$ mit

- $x_k \to \bar{x}$,
- $\forall |\alpha| \leq 3, D_\alpha(f^k, G^k)(z_k) \to D_\alpha(\bar{f}, \bar{G})(\bar{z})$,
- $T_{(x^0,q)}(f^k, G^k) \notin \mathcal{F}|_{\{x_k\}\times\{[-n,n]\}}$.

Das impliziert die Existenz einer Folge $t_k \in [-n, n]$ mit $T_{(x^0,q)}(f^k, G^k) \notin \mathcal{F}|_{\{x_k\}\times\{t_k\}}$.

Aus der Kompaktheit von $[-n, n]$ können wir annehmen, daß die Folge t_k gegen $\bar{t} \in [-n, n]$ konvergiert. Die Definition der Funktionen $f^{int}(f^k, x, t) \in C^3(I\!R^{n+1}, I\!R)$ und $g_j^{int}(g_1^k, x, t) \in C^3(I\!R^{n+1}, I\!R)$ für $j \in J$ und die Konvergenz von t_k zu $\bar{t}$ ergeben zwei Bedingungen.

- $z_k \to \bar{z}$
- $\forall |\alpha| \leq 3, D_\alpha(f^{int}(f^k, z), G^{int}(G^k, z))(z_k) \to D_\alpha(f^{int}(\bar{f}, z), G^{int}(\bar{G}, z))(\bar{z})$

Diese Bedingungen, zusammen mit den beiden Voraussetzungen

$$\begin{aligned} T_{(x^0,q)}(\bar{f}, \bar{G}) &\in \mathcal{F}|_{\{\bar{x}\}\times[-n,n]} \text{ und} \\ T_{(x^0,q)}(f^k, G^k) &\notin \mathcal{F}|_{\{x_k\}\times\{t_k\}} \end{aligned}$$

widersprechen dem Satz 2, weil die beiden folgenden Aussagen erfüllt sind.

- Die Annahme $T_{(x^0,q)}(\bar{f}, \bar{G}) \in \mathcal{F}|_{\{\bar{x}\}\times[-n,n]}$ impliziert folgendes: Falls $(\bar{x}, \bar{t})$ ein verallgemeinerter kritischer Punkt von $P^{int}_{(x^0,q)}(\bar{f}, \bar{G}, t)$ ist, so ist er vom Typ 1,2,3,4 oder 5.
- Die Bedingung $T_{(x^0,q)}(f^k, G^k) \notin \mathcal{F}|_{\{x_k\}\times\{t_k\}}$ ergibt, daß der Punkt (x_k, t_k) ein verallgemeinerter kritischer Punkt von $P^{int}_{(x^0,q)}(f^k, G^k, t)$ und kein Punkt der Typen 1,2,3,4 oder 5 ist.

Damit ist der Beweis zur Behauptung 2 und auch des Satzes beendet. □

Abschließend vergleichen wir kurz die vorgestellten Einbettungen und ihre Eigenschaften mit anderen ähnlichen Zugängen für die Benutzung der Kurvenverfolgungsverfahren für die nichtlineare Optimierung. In [19] wird eine differenzierbare Homotopie vorgestellt. Für $t = 0$ wird das KKT-System erhalten, außer der Positivität der Multiplikatoren. Dies wird gewonnen, indem

man mit einem Paar (x^0, w^0) beginnt und die strikte positive Komplementarität im ganzen Pfad erzwingt, wobei x^0 ein innerer Punkt des zu lösenden Problems ist und w^0 die Rolle von positiven Multiplikatoren spielt.

Das einparametrische System besteht aus einer Kombination linearer Homotopie im Gradienten-Teil des KKT-Systems und einer Newton-Homotopie aus der inneren Methode im Komplementaritätsteil.

Ausgangspunkt ist eine Menge der Form

$$\Omega = \{x \in I\!R^n | G(x) > 0\}$$

und ein Optimierungsproblem mit dem zulässigen Bereich

$$\bar{\Omega} = \{x \in I\!R^n | G(x) \geq 0\}\,.$$

Uns interessieren hier die Konvergenzbedingungen der Homotopie im Zusammenhang mit unserer Einbettung. Die in [19] benutzte Regularität, damit die kritischen Mengen eine vernünftige Gestalt für die Kurvenverfolgung besitzen, ist schwächer als die JJT-Regularität. Andererseits werden die folgenden drei Bedingungen für die Konvergenz vorausgesetzt.

D1 Die Menge Ω soll nicht leer und beschränkt sein.

D2 LICQ ist in jedem zulässigen Punkt $x \in \bar{\Omega}$ erfüllt.

D3 Für alle Randpunkte $x \in \partial\Omega = \bar{\Omega} \setminus \Omega$ haben der Normalen-Kegel in x und Ω einen leeren Durchschnitt, d.h. $\forall x \in \partial\Omega$

$$\left\{ x - \sum_{j \in J_0(x)} \mu_j Dg_j(x) \,\middle|\, \sum_{j \in J_0(x)} \mu_j > 0 \text{ und } \mu_j \geq 0 \right\} \cap \Omega = \emptyset.$$

Verzichten wir, auf die ersten beiden Bedingungen einzugehen, da der wesentliche Unterschied durch die Voraussetzung **D3** markiert wird. Speziell folgt aus der Gültigkeit von **D3**, daß die Funktion $g_0^{x^0} = \frac{1}{2}\|x - x^0\|$ für jedes $x^0 \in \Omega$ einen einzigen stationären Punkt auf dem gesamten zulässigen Bereich $\bar{\Omega}$ hat, und dieser ist x^0. Außerdem ist x^0 ein nichtdegenerierter stationärer Punkt mit quadratischem Index $QI = 0$ und quadratischem Ko-Index $QCI = n$.

Andererseits gilt wegen der LICQ und unter Benutzung des parametrisierten Satz vom Sard, daß für fast alle x^0 in der offenen Menge Ω die Funktion $g_0^{x^0}$ regulär auf der Menge $\bar{\Omega}$ ist, d.h. alle kritische Punkte auf $\bar{\Omega}$ sind nicht-degeneriert.

Damit sind der Satz 5.2.2 und die Bemerkung 5.2.2 von [31] in diesem Fall anwendbar. Es folgt, daß der zusammenhängende Teil Ω nulldimensionale q-Homologieräume für $q > 0$ und einen eindimensionalen 0-Homologieraum hat. Die mit den Bedingungen **D1** bis **D3** verbundenen topologischen Einschränkungen, die für unsere Einbettung nicht erforderlich sind, werden dann erklärt.

Als letzten Vergleichspunkt machen wir folgende Bemerkung. Falls die Bedingungen **D1** bis **D3** erfüllt sind, können wir $x^0 \in \Omega$ nehmen und eine Version unserer Einbettung so betrachten, daß die zulässige Menge konstant zwischen 0 und 1 bleibt. Unter Regularität kann die Kurve der stationären Punkte nicht nach $t = 0$ zurückkehren, damit ist die Konvergenz zu einem stationären Punkt ohne Notwendigkeit von Sprüngen gesichert.

4.2 Generische quadratische Penalty-Einbettungen.

Die Benutzung glatter Penalty-Funktionen für die Lösung von Optimierungsproblemen mit Nebenbedingungen ist oftmals numerisch schwierig. Die Hauptschwierigkeit liegt in der unvermeidbaren schlechten Konditionierung, wenn der Parameter nahe dem vereinbarten Wert ist (siehe z.B. [27]).

Eine Möglichkeit zur Beseitigung dieser Problematik ist die Umformulierung der Optimalitätsbedingungen des restriktionsfreien Problems als nichtlineares einparametrisches Gleichungssystem und die Benutzung von Kurvenverfolgungsansätzen dafür. Das kann neue Einsichten in die Schwierigkeiten des Verfahrens und seiner Konvergenz mit sich bringen, wenn z.B. die möglichen Singularitäten erkannt und untersucht werden (siehe. z.B. [54], [55], [56] und für die dort benutzte Singularitätenanalysis die vorhergehenden Arbeiten [57] und [58]).

Mit diesem Ansatz wurden verschiedene Verfahren der Optimierung mit Nebenbedingungen unter der Annahme der JJT-Regularität untersucht (siehe [3], [4], [11], [15]). Zum Vergleich der Ansätze und für die Konvergenzuntersuchungen ist die JJT-Regularität die Schlüsselvoraussetzung, weil man vor allem aus der lokalen Struktur und den im vorhergehenden Kapitel vorgestellten Eigenschaften profitiert.

Die JJT-Regularität ist keine strenge Voraussetzung (generischer Charakter). Das aber kann sich ändern, wenn man sich auf einparametrische Probleme spezifischer Struktur beschränkt. Zu dieser Problematik sind in letzter Zeit mehrere Arbeiten (siehe [23], [65], [53]) erschienen, die für verschiedene Klassen eine spezielle Untersuchung der generischen Singularitäten enthalten.

Wir untersuchen in diesem Abschnitt den generischen Charakter der JJT-Regularität für die Probleme, die durch eine die quadratische Penalty-Methode darstellende Einbettung erhalten werden. Damit kommen wir zu (nicht einparametrischen) Problemen der Gestalt

$$P \quad \min\{f(x)|\ x \in M\} \tag{39}$$

zurück, wobei

$$M := \{x \in I\!R^n|\ h_i(x) = 0,\ i \in I, \quad g_j(x) \geq 0,\ j \in J\}$$

und $I = \{1, \ldots, m\}$, $J = \{1, \ldots, p\}$.

Der quadratische Penalty-Ansatz zur Lösung dieses Problems erfolgt durch die Betrachtung einer Folge von Optimierungsproblemen ohne Nebenbedingungen, z.B. folgender Form.

$$P(t) \quad \min\{f(x) + \left(\frac{t}{(1-t)}\right)^2 \left[\sum_{i \in I} h_i^2(x) + \sum_{j \in J} (\min\{g_j(x), 0\})^2\right] \mid x \in I\!R^n\} \tag{40}$$

Hier bewegt sich der Parameter t wachsend gegen $t = 1$, und dadurch gewinnt der quadratische Penalty-Teil an Gewicht.

In [11] ist die enge Beziehung zwischen den Lösungen von (40) und ein vorgeschlagenes einparametrisches Optimierungsproblem mit Nebenbedingungen zu finden. Das vorgeschlagene einparametrische Problem ist eine Einbettung von (39), die nach einigen zu besseren Eigenschaften des parametrischen Problems führenden Transformationen z.B. nach der Kompaktifizierung, die folgende Gestalt

$$\widehat{P}(t) \quad \min\{\hat{f}(x, v, w, t)|(x, v, w) \in \hat{M}(t)\} \tag{41}$$

besitzt, wobei

$$\hat{f}(x, v, w, t) := tf(x) + (1-t)\|x - x^0\|^2 + \|v - v^0\|^2 + \|w - w^0\|^2$$

und

$$\hat{M}(t) := \left\{ (x,v,w) \in I\!R^{n+m+p} \;\middle|\; \begin{array}{c} th_i(x) + (1-t)(v_i - v_i^0) = 0,\ i \in I \\ tg_j(x) + (1-t)(w_j - w_j^1) \geq 0,\ j \in J \\ q - \|x - x^1\|^2 - \|v - v^0\|^2 - \|w - w^1\|^2 \geq 0 \end{array} \right\}$$

gelten.

Die Vektoren $(x^0, v^0, w^0) \in I\!R^{n+m+p}$, $w^1 \in I\!R^p$, $x^1 \in I\!R^n$, und die positive Zahl $q \in I\!R$ sind alle fest und besitzen folgende Eigenschaften.

$$\begin{aligned} \|x^0 - x^1\|^2 + \|w^0 - w^1\|^2 &< p \\ w_j^1 - w_j^0 &< 0,\ \forall j \in J \end{aligned} \tag{42}$$

Wie in den vorhergehenden Kapiteln bezeichnet d_K einen Teilvektor von $d \in I\!R^s$, gegeben durch die Komponenten d_i, $i \in K$, falls $K \subset \{1, \ldots, s\}$ ist.

Das einparametrische Problem (41) wird ausgehend von den Daten

$$(f, H, G) := (f(x), h_i(x), i \in I, g_j(x), j \in J) \in C^3(I\!R^n, I\!R^{m+p+1})$$

konstruiert.

Diese Konstruktion kann mittels der Abbildung $\Phi : C^3(I\!R^n, I\!R^{m+p+1}) \longmapsto C^3(I\!R^{n+m+p+1}, I\!R^{m+p+2})$, definiert als

$$\Phi(f, H, G) = \begin{pmatrix} tf(x) + (1-t)\|x - x^0\|^2 + \|v - v^0\|^2 - \|w - w^0\|^2 \\ th_1(x) + (1-t)(v_1 - v_1^0) \\ \vdots \\ th_m(x) + (1-t)(v_m - v_m^0) \\ tg_1(x) + (1-t)(w_1 - w_1^1) \\ \vdots \\ tg_p(x) + (1-t)(w_p - w_p^1) \\ q - \|x - x^1\|^2 - \|v - v^0\|^2 - \|w - w^1\|^2 \end{pmatrix} \tag{43}$$

dargestellt werden.

Wir wollen die Voraussetzung $\Phi(f, H, G) \in \mathcal{F}$ untersuchen. Für die Benutzung der Einbettung ist sie jedoch nicht auf der ganzen reellen Achse notwendig. Wir konzentrieren uns deswegen auf die schwächere Voraussetzung $\Phi(f, H, G) \in \mathcal{F}|_{[0,1]}$, was nichts anderes heißt, als die Untersuchung der Menge $\Phi^{-1}(\mathcal{F}|_{[0,1]})$. Das Ergebnis bezüglich dieser Problematik wird mit dem Satz 18 bereitgestellt.

Satz 18
Die Menge $\Phi^{-1}(\mathcal{F}|_{[0,1]}) \subset C^3(I\!R^n, I\!R^{m+p+1})$ *ist offen und dicht bezüglich der starken Whitney-Topologie.*

Wie im vorigen Abschnitt und im Fall der Klasse $\mathcal{F}$ kann dieses Ergebnis mittels einer lokal$\longrightarrow$global Prozedur erhalten werden, sofern man zwei Fakten bewiesen hat.

Erstens die lokale Stabilität der Voraussetzung $\Phi(f, H, G) \in \mathcal{F}|_{[0,1]}$ und zweitens die Möglichkeit, diese Voraussetzung mittels beliebig kleiner Störungen erreichen zu können.

Nach dem Muster des Beweises der Behauptung 2 kann man die lokale Stabilität mit Stetigkeits- und Kompaktheitsargumenten auf den Satz 2 über die Stabilität der 5 Typen der JJT-Regularität reduzieren.

Das zu zeigende Störungsergebnis lautet wie folgt.

Proposition 1 *(Störungssatz)*
Es seien $(\bar{f}, \bar{H}, \bar{G}) \in C^3(I\!R^n, I\!R^{m+s+1})$ *fest, dann hat jede meßbare Untermenge der Menge*

$$\left\{ \mathcal{P} \in I\!R^{\frac{1}{2}n(n+1)+n+(n+1)(m+p)} \left| \Phi\left[(\bar{f}, \bar{H}, \bar{G}, \mathcal{P})\right] \notin \mathcal{F}|_{[0,1]} \right. \right\} \tag{44}$$

das Lebesgue-Maß Null.

Dabei sind die Parametervektoren $\mathcal{P} = (\mathcal{A}, \mathcal{C})$, $\mathcal{A} = (A, b) \in I\!R^{\frac{1}{2}n(n+1)+n}$ und $\mathcal{C} = (C, d) \in I\!R^{(n+1)(m+p)}$. Das gestörte Problem

$$(\bar{f}, \bar{H}, \bar{G}, \mathcal{P}) \in C^3(I\!R^n, I\!R^{m+p+1})$$

ist das mit den Daten $\left(\bar{f}(x) + \frac{1}{2}x^T Ax + b^T x, (\bar{H}, \bar{G})(x) + Cx + d\right)$.

Wir benutzen die Bezeichnungen $\Phi(\bar{f}, \overline{H}, \overline{G}, \mathcal{P})$ oder $\overline{\Phi}(x, v, w, t, \mathcal{P})$ für das Problem mit den Daten

$$\Phi \begin{pmatrix} \bar{f} + 0.5x^T Ax + b^T x \\ \bar{h}_1 + c_1^T x + d_1 \\ \vdots \\ \bar{h}_m + c_m^T x + d_m \\ \bar{g}_1 + c_{m+1}^T x + d_{m+1} \\ \vdots \\ \bar{g}_s + c_{m+p}^T x + d_{m+p} \end{pmatrix}$$

und $\hat{M}(\overline{H}, \overline{G}, \mathcal{C}, t)$ für seinen parametrischen zulässigen Bereich.

Der Beweis dieser Behauptung verläuft analog dem Beweis des entsprechenden Satzes für die Klasse $\mathcal{F}$ (Satz 13). Einen vollständigen Beweis vom Satz 13 findet man in [61] und [60]. Der Beweis reduziert sich auf die folgenden drei Behauptungen.

Behauptung 3
Für fast alle $\mathcal{P} \in \mathbb{R}^{\frac{1}{2}n(n+1)+n+(n+1)(m+p)}$ gilt, daß jeder verallgemeinerte kritische Punkt bezüglich des Problems $\Phi(\bar{f}, \overline{H}, \overline{G}, \mathcal{P})$, in dem die LICQ erfüllt ist, ein Punkt der Typen 1, 2 oder 3 ist.

Behauptung 4
Für fast alle $\mathcal{C} \in \mathbb{R}^{(n+1)(m+p)}$ sind folgende Bedingungen erfüllt:

1. *Die Teilmenge*

$$\left\{(x, v, w, t) \in \hat{M}(\overline{H}, \overline{G}, \mathcal{C}, t) | \textit{die LICQ ist verletzt}\right\} \tag{45}$$

   ist eine null-dimensionale Mannigfaltigkeit.

2. *In jedem Punkt $(\overline{x}, \overline{v}, \overline{w}, \overline{t})$ der Menge (45) gelten folgende Bedingungen:*

   (a) *Die Gradienten der aktiven Restriktionen (sowohl Gleichungen als auch Ungleichungen) sind linear unabhängig.*

   (b) *Jede verschwindende lineare Kombination der partiellen Ableitungen der aktiven Restriktionen (bzgl. der Variablen x, v und w) hat alle Koeffizienten, die zu Ungleichungen gehören, verschieden von Null.*

Behauptung 5
Ist $\mathcal{C} \in \mathbb{R}^{n(m+p)+m+p}$ ein fester Parameter, so daß die Bedingungen aus Behauptung 4 erfüllt sind, so gilt für fast alle $\mathcal{A} \in \mathbb{R}^{\frac{1}{2}n(n+1)+n}$, daß jeder Punkt der Menge (45) vom Typ 4 ist.

Wieder ist der parametrisierte Satz von Sard das fundamentale Werkzeug für die Beweise. Der erste Teil besteht darin, Abbildungen zu definieren, die vom Parameter der Störungen (nämlich $\mathcal{P}$) und von anderen Variablen (diese letzte Variablengruppe enthält diejenigen des parametrischen Problems, also (x, v, w, t), und die Multiplikatoren) abhängen und folgende Bedingungen erfüllen:

1. Null ist ein regulärer Wert der Abbildung.

2. Fixiert man die Parameterwerte und betrachtet die dadurch entstandene Abbildung, die nur von der eben genannten zweiten Variablengruppe abhängt, dann impliziert die Regularität von Null als Wert der neuen Abbildung Bedingungen, die in der Definition der 5 Singularitäten enthalten sind.

Der entscheidende Unterschied zwischen den vorhergehenden Behauptungen und den entsprechenden Aussagen in [61] liegt in der kleineren Menge verfügbarer Parameter für die Störung im einparametrischen Optimierungsproblem. Es sei z.B. bemerkt, daß keine Störungen bezüglich der Variablen v und w vorhanden sind, und daß die Kompaktifizierungsungleichung auch keinerlei Störungen unterliegt. In diesem Sinne ist der entscheidende Teil die Überprüfung, daß Null ein Regulärwert der entsprechenden Abbildungen (erste der vorhergenannten Bedingungen) bleibt.

Zum Beweis der Behauptungen legen wir die Abbildungen $(\bar{f}, \bar{H}, \bar{G}) \in C^3(I\!R^n, I\!R^{m+p+1})$ fest und führen folgende nützlichen Bezeichnungen ein: Die Variable y ist definiert als $y = (x, v, w) \in I\!R^{n+m+p}$, und $u \in I\!R^{m+s+1}$ wird für die Bezeichnung der Multiplikatoren des Problems $\overline{\Phi}(y, t, \mathcal{P})$ benutzt. Die $m + p + 2$ definierenden Funktionen des Problems $\overline{\Phi}(y, t, \mathcal{P})$ seien wie folgt indiziert:

$$\begin{aligned}
\overline{\Phi}_0(y,t,\mathcal{A}) &= t\overline{f}(x) + (1-t)\|x - x^0\|^2 + \|v - v^0\|^2 - \\
&\quad -\|w - w^0\|^2 + t(0.5x^T Ax + b^T x), \\
\overline{\Phi}_i(y,t,\mathcal{C}) &= t\overline{h}_i(x) + (1-t)(v_i - v_i^0) + t(c_i^T x + d_i),\ i \in I, \\
\overline{\Phi}_{m+j}(y,t,\mathcal{C}) &= t\overline{g}_j(x) + (1-t)(w_j - w_j^1) + t(c_{m+j}^T x + d_{m+j}),\ j \in J, \\
\overline{\Phi}_{m+p+1}(y,t,\mathcal{C}) &= q - \|x - x^1\|^2 - \|v - v^0\|^2 - \|w - w^1\|^2.
\end{aligned}$$

Mit $\overline{\Phi}_K(x, v, w, t, \mathcal{A}) \in C^3(I\!R^{n+m+p+1}, I\!R^{|K|})$ bezeichnen wir den Teil der Abbildung,der durch eine Indexmenge $K \subseteq \{0, \ldots, m + p + 1\}$ gegeben ist.

Beweis zur Behauptung 3:

Wir definieren eine Abbildung, die von drei Indexmengen abhängt, nämlich:

- $K_1 \subseteq \{m + 1, \ldots, m + p + 1\}$, die als die Menge der aktiven Ungleichungsrestriktionen des Problems $\overline{\Phi}(y, t, \mathcal{A})$ interpretiert wird. Wir betrachten den interesanten Fall $\{m + p + 1\} \subset K_1$.

- $K_2 \subseteq \{1, \ldots, \gamma\}$, wobei $\gamma = n + m + p + m + |K_1|$.
- $K_3 \subseteq K_1$, die als Menge der verschwindenden, zu den aktiven Ungleichungsrestriktionen assoziierten Multiplikatoren gedacht ist.

Mit $\overline{K_1}$ bezeichnen wir die Indexmenge, die durch Hinzufügen der Indizes der Gleichungsrestriktionen zu K_1 entsteht, nämlich $\overline{K_1} = \{1, \ldots, m\} \cup K_1$. Die dieser Indexmenge und dem Problem $\overline{\Phi}_K(x, v, w, t, \mathcal{A})$ entsprechende Kojima-Abbildung wird jetzt wie folgt gebildet: .

$$\mathcal{K}(y, t, u, \mathcal{P}) = \begin{pmatrix} D_y \overline{\Phi}_0(y, t, \mathcal{A}) - \sum_{k \in \overline{K_1}} u_k D_y \overline{\Phi}_k(y, t, \mathcal{C}) \\ \\ \overline{\Phi}_k(y, t, \mathcal{C}), \quad k \in \overline{K_1} \\ \\ u_k + \overline{\Phi}_k(y, t, \mathcal{C}), \quad k \in \{1, \ldots, m + s + 1\} \setminus \overline{K_1} \end{pmatrix}.$$

Die symmetrischen Matrizen mit l Zeilen und Spalten bezeichnen wir mit $M^s(l)$ und identifizieren diese Matrizen mit Vektoren des Raumes $I\!R^{\frac{1}{2}p(p+1)}$. Wir definieren dann die folgende Abbildung mit Werten im Raum $M^s(\gamma)$:

$$M(y, t, u, \mathcal{P}) = \begin{pmatrix} D_y^2[\overline{\Phi}_0(y,t,\mathcal{A}) - \sum_{k \in \overline{K_1}} u_k \overline{\Phi}_k(y,t,\mathcal{C})] & D_y^T \overline{\Phi}_{\overline{K_1}}(y,t,\mathcal{C}) \\ \\ D_y \overline{\Phi}_{\overline{K_1}}(y,t,\mathcal{C}) & 0 \end{pmatrix}.$$

Es sei $\mathcal{M}(K_2) \subset M^s(\gamma)$ die Menge solcher symmetrischer Matrizen mit Rang $|K_2|$, deren Spalten mit Indizes in K_2 unabhängig sind. Die Menge $\mathcal{M}(K_2)$ ist eine differenzierbare Mannigfaltigkeit $\mathcal{M}(K_2)$ mit Kodimension $\frac{1}{2}(\gamma - |K_2|)(\gamma - |K_2| + 1)$. Es sei mit $\Theta \in C^3(I\!R^{\frac{1}{2}\gamma(\gamma+1)}, I\!R^{\frac{1}{2}(\gamma - |K_2|)(\gamma - |K_2| + 1)})$ eine differenzierbare Abbildung bezeichnet, die $\mathcal{M}(K_2)$ lokal definiert, und deren Ableitungen bezüglich solcher Elemente der Matrizen außerhalb der Spalten mit Indizes in K_2 eine Einheitsmatrix bilden (siehe z.B. [61]).

Die entscheidende Abbildung Ψ_1 dieser Behauptung sei jetzt als

$$\Psi_1(y, t, u, l, \mathcal{P}) = \begin{pmatrix} \mathcal{K}(y, t, u, \mathcal{P}) \\ M(y, t, u, \mathcal{P}) - l \\ \Theta(l) \\ u_{K_3} \end{pmatrix}$$

definiert.

Die Variable l bewegt sich im Raum $I\!R^{\frac{1}{2}\gamma(\gamma+1)}$. Wie erwähnt, besteht der wichtigste Schritt dieses Beweises in der Untersuchung der Jakobi-Matrix von Ψ_1.

$$
\begin{array}{c}
 \\ \mathcal{K} \\ \\ \\ M-l \\ \Theta(l) \\ u_{K_3}
\end{array}
\begin{array}{c}
\begin{array}{cccccc} D_{(y,t)} & D_u & D_l & D_A & D_b & D_d \end{array} \\
\begin{bmatrix}
 & & & \otimes & {\scriptstyle tI_n} & 0 \\
 & & & 0 & 0 & 0 \\
\otimes & \otimes & 0 & 0 & 0 & {\scriptstyle tI_{m+s}} \\
 & & & 0 & 0 & 0 \\
\otimes & \otimes & -I_{\frac{1}{2}\gamma(\gamma+1)} & \begin{matrix} {\scriptstyle tI_{\frac{1}{2}n(n+1)}} \\ 0 \end{matrix} & 0 & 0 \\
0 & 0 & {\scriptstyle I_{\frac{1}{2}(\gamma-|K_2|)(\gamma-|K_2|+1)}|\times} & 0 & 0 & 0 \\
0 & {\scriptstyle I_{|K_3|}|0} & 0 & 0 & 0 & 0
\end{bmatrix}
\end{array}
\tag{46}
$$

Hier und in den folgenden Matrizen bezeichnet $\otimes$ eine Teilmatrix, die für die Analyse uninteressant ist.

Wir zeigen, daß diese Matrix in allen Nullstellen$\Psi_1(y,t,u,l,\mathcal{P}) = 0$ linear unabhängige Zeilen hat, falls folgende Bedingungen in der Nullstelle gelten:

1. t ist verschieden von 0 und 1.

2. Die Matrix $D_y\overline{\Phi}_{\overline{K_1}}(y,t,\mathcal{C})$ hat linear unabhängige Zeilen, was dem Erfülltsein der LICQ entspricht.

Aus der Struktur von (46) folgt im Fall $\{n+1,\ldots,\gamma\} \subset K_2$, daß $\frac{1}{2}n(n+1) \geq \frac{1}{2}(\gamma-|K_2|)(\gamma-|K_2|+1)$ ist. Daher kann man die gesuchte Vollrangbedingung auf eine ähnliche Vollrangbedingung reduzieren, die nur für einige der Komponenten von $\mathcal{K}(y,t,u,\mathcal{C})$ gestellt ist.

Da $\Psi_1(y,t,u,l,\mathcal{P}) = 0$ gilt, reduziert sich die Untersuchung der Bedingung $\{n+1,\ldots,\gamma\} \subset K_2$ auf die Analyse der letzten $2m+p+|K_1|$ Spalten

der Matrix $M(y,t,u,\mathcal{P})$, dargestellt in (47).

$$\begin{array}{cccc} D_x & D_{(v,w)} & k \in \overline{K_1} \setminus \{m+p+1\} & m+p+1 \end{array}$$
$$\begin{bmatrix} \bigotimes & 0 & D_x^T \overline{\Phi}_k & -2(x-x^1) \\ 0 & 2I_{m+s} & \begin{matrix}(1-t)I \\ 0\end{matrix} & \begin{matrix}-2(v-v^0) \\ -2(w-w^1)\end{matrix} \\ D_x \overline{\Phi}_k & (1-t)I \quad | \quad 0 & 0 & 0 \\ -2(x-x^1)^T & -2(v,w)^T - 2(v^0,w^1)^T & 0 & 0 \end{bmatrix} \tag{47}$$

Sind im Vektor $(w-w^1)$ alle Komponenten mit Indizes aus K_1 gleich Null, so folgt die lineare Unabhängigkeit der letzten $2m+p+|K_1|$ Spalten aus der LICQ (Vollrang von $D_y \overline{\Phi}_{\overline{K_1}}(y,t,\mathcal{C})$) unter Benutzung des passenden Teils der Matrix $2I_{m+s}$. Im anderen Fall folgt die genannte lineare Unabhängigkeit aus der Regularität der Teilmatrix (48).

$$\begin{pmatrix} 2I_{m+p} & \begin{matrix}(1-t)I \\ 0\end{matrix} & \begin{matrix}-2(v-v^0) \\ -2(w-w^1)\end{matrix} \\ (1-t)I \quad | \quad 0 & 0 & \\ -2(v,w)^T - 2(v^0,w^1)^T & 0 & \end{pmatrix} \tag{48}$$

Die letzten $2m+p+|K_1|$ Spalten von $M(y,t,u,\mathcal{P})$ sind also linear unabhängig. Kommen wir zur Analyse der Matrix (46). Es seien $\vartheta = (\vartheta^{\mathcal{K}}, \vartheta^M, \vartheta^\Theta, \vartheta^{K_3})$ die Koeffizienten einer verschwindenden Linearkombination der Spalten von (46). Wir zeigen $\vartheta = 0$.

Die Koeffizienten, mit der Teilabbildung $\mathcal{K}$ azossiiert, teilen wir wie folgt ein: $\vartheta^{\mathcal{K}} = (\vartheta^x, \vartheta^{(v,w)}, \vartheta^{\{1,\ldots,m+p\}}, \vartheta^{m+p+1})$, wobei ϑ^{m+p+1} zur Kompaktifizierungsungleichung und $\vartheta^{\{1,\ldots,m+p\}}$ zu den anderen aktiven Restriktionen gehören.

Die Einheitsmatrix tI_n (bzw. tI_{m+p}), zu finden in (46) in der Spalte der partiellen Ableitungen nach den Parametern b (bzw. d), impliziert, daß $\vartheta^x = 0$ (bzw. $\vartheta^{\{1,\ldots,m+p\}} = 0$). Das, zusammen mit dem eben bewiesenen Fakt über die letzten Zeilen von $M(y,t,u,\mathcal{A})$, liefert aus der Struktur von (46) die Gleichungen $\vartheta^M = 0$ und $\vartheta^\Theta = 0$.

Zum Beweis der Beziehungen $(\vartheta^{(v,w)}, \vartheta^{m+p+1}, \vartheta^{K_3}) = 0$ analysieren wir den in (49) dargestellten Rest der Jacobi-Matrix (46).

$$
\begin{array}{cccc}
D_x & D_{(v,w)} & \{1,\ldots,m\} & K_1 \\
\end{array}
$$

$$
\left[\begin{array}{cccc}
 & & (1-t)I_m & 0 \quad -2(v-v^0) \\
0 & 2I_{m+p} & & (1-t)I_{K_1-1} \\
 & & 0 & 0 \quad -2(w-w^1) \\
-2(x-x^1)^T & -2(v-v^0)^T|-2(w-w^1)^T & 0 & 0 \\
0 & 0 & 0 & 0 \mid I_{|K_3|} \mid 0
\end{array}\right] \tag{49}
$$

Es reicht, $\vartheta^{m+p+1} = 0$ zu zeigen. Das folgt aber einfach aus der Struktur der mit D_v und $\{1,\ldots,m\}$ assoziierten Spalten, falls $(x,v) \neq (x^1,v^0)$ gilt. Es sei dann $(x,v) = (x^1,v^0)$. Da die Kompaktifizierungsungleichung als aktiv angenommen ist, gilt dann die Beziehung $\|w-w^1\|^2 = q \neq 0$.

Gehört der Index $m+p+1$ nicht zur Menge K_3(d.h. die letzte Spalte von (49) hat nur Nullen in den Zeilen, assoziiert mit ϑ^{K_3}), dann impliziert die Regularität der Teilmatrix

$$
\left(\begin{array}{cc}
2I_{m+p} & \begin{array}{c} -2(v-v^0) \\ -2(w-w^1) \end{array} \\
 & \\
-2(v-v^0)^T| - 2(w-w^1)^T & 0
\end{array}\right),
$$

daß $(\vartheta^{(v,w)}, \vartheta^{m+p+1}) = 0$, und wir sind fertig.

Falls $\{m+p+1\} \in K_3$ $(u_{m+p+1} = 0)$, reduziert sich die Gleichung

$$
D_y\overline{\Phi}_0(y,t,\mathcal{A}) - \sum_{k\in\overline{K_1}} u_k D_y\overline{\Phi}_k(y,t,\mathcal{C}) = 0
$$

(enthalten in $\mathcal{K}(y,t,u,\mathcal{P}) = 0$), betrachtet nur für die partiellen Ableitungen D_w, auf die folgenden Gleichungen für alle $j \in J$.

$$
2(w_j - w_j^0) = \begin{cases} (1-t)u_{m+j}, & m+j \in K_1, \\ 0, & m+j \notin K_1 \end{cases} \tag{50}
$$

Andererseits implizieren die Ungleichungen

$$
\begin{array}{rcl}
\|w-w^1\|^2 & = & q \\
\|w^0-w^1\|^2 & < & q
\end{array}
$$

die Existenz eines Index (o.B.d.A. gleich 1) in J, so daß w_1 verschieden von w_1^1 und von w_1^0 ist. Aus (50) folgt $m+1 \in K_1$, $u_{m+1} \notin K_3$ und $w_1 - w_1^0 \neq 0$. Aus der Gestalt der Matrix (49) ergibt sich, daß aus $u_{m+1} \notin K_3$ die Gleichung $\vartheta^{w_1} = 0$ folgt, und dann $\vartheta^{m+p+1}(w_1 - w_1^1) = 0$, was $\vartheta^{m+p+1} = 0$ impliziert.

Wir haben bewiesen, daß in jeder Nullstelle der Abbildung $\Psi_1(y,t,u,z,\mathcal{A})$, bei der t verschieden von 0 und 1 ist und $D_y\overline{\Phi}_{\overline{K_1}}(y,t,\mathcal{C})$ vollen Rang hat, die Jacobi-Matrix von Ψ_1 auch linear unabhängige Spalten hat. Diese Bedingung entspricht der Transversalitätsbedingung in [61], und die Folgerungen bis zum Erhalt der drei möglichen Fälle vom Typ 1 bis 3 können mit denselben dort benutzten Argumenten gezogen werden. Damit ist die erste Behauptung bewiesen. □

Beweis zur Behauptung 4 :

Wieder machen wir den Beweis für den Fall, daß $t \neq 1$ und $t \neq 0$. Daraus folgt, daß die Kompaktifizierungsungleichung in den Punkten, in denen die LICQ verletzt ist, aktiv sein muß, denn die Gradienten der anderen Restriktionen sind linear unabhängig.

Die Abbildung, definiert in diesem Teil, hängt wie im vorigen Beweis von zwei Indexmengen ab. Es seien K_1 und K_3 wie vorher festgelegt: Da die Kompaktifizierungsungleichung aktiv sein muß, und der Einfachheit der Bezeichnungen halber, vereinbaren wir $\{m+p+1\} \notin K_1$.

In diesem Beweis benötigen wir eine Variable ξ mit Indizes in $\overline{K_1}$ für die Koeffizienten einer Linearkombination der aktiven Restriktionen außer die Kompaktifizierung. Es ist dann $\xi \in I\!R^{|\overline{K_1}|}$. Es sei die folgende Abbildung

$$\Psi_2(y,t,\xi,\mathcal{C}) = \begin{pmatrix} \sum_{k\in\overline{K_1}} \xi_k D_y\overline{\Phi}_k(y,t,\mathcal{C}) + D_y\overline{\Phi}_{m+p+1}(y,t,\mathcal{C}) \\ \overline{\Phi}_k(y,t,\mathcal{C}), \quad k \in \overline{K_1} \\ \overline{\Phi}_{m+p+1}(y,t,\mathcal{C}) \\ \xi_{K_3} \end{pmatrix}$$

definiert.

Da $D_y\overline{\Phi}_{m+p+1}(y,t,\mathcal{C}) = -2(x-x^1, v-v^0, w-w^1) \neq 0$ in einer Nullstelle $(y,t,\xi,\mathcal{C})$ von Ψ_2 ist, muß der Vektor $\xi \in I\!R^{|\overline{K_1}|}$ verschieden von Null sein. Es sei o.B.d.A. $\xi_1 \neq 0$, was $(v_1 - v_1^0) \neq 0$ und $1 \notin K_3$ impliziert. Betrachten wir die Jacobi-Matrix (51) dieser Abbildung in solchen Punkten mit

$\Psi_2(y,t,u,\mathcal{C}) = 0$, $t \neq 1$ und $t \neq 0$.

$$\begin{array}{cccccc} D_x & D_{(v,w)} & D_{\xi_1} & D_{\xi_{K_3}} & D_{c_1} & D_{d_{\overline{K_1}}} \end{array}$$
$$\begin{bmatrix} \bigotimes & 0 & D_x^T\overline{\Phi}_1 & \bigotimes & \xi_1 t I_n & 0 \\ 0 & -2I_{m+p} & \begin{matrix}(1-t)\\0\end{matrix} & 0 & 0 & 0 \\ D_x\overline{\Phi}_{\overline{K_1}} & (1-t)I_{\overline{K_1}}|0 & 0 & 0 & \bigotimes & tI_{K_1} \\ -2(x-x^1)^T & -2(v-v^0)^T|-2(w-w^1)^T & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_{|K_3|} & 0 & 0 \end{bmatrix} \tag{51}$$

Es seien $\vartheta = (\vartheta^x, \vartheta^{(v,w)}, \vartheta^{\overline{K_1}}, \vartheta^{m+p+1}, \vartheta^{K_3})$ die Koeffizienten einer verschwindenden Linearkombination der Zeilen von (51), wobei die Einteilung von ϑ mit den verschiedenen Gruppen von Zeilen übereinstimmt. Wir wollen zeigen, daß $\vartheta = 0$.

Da t und ξ_1 verschieden von Null sind, ergeben die Matrizen tI_{K_1}, $\xi_1 t I_n$ und $I_{|K_3|}$ die Beziehungen $\vartheta^x = 0$, $\vartheta^{\overline{K_1}} = 0$ und $\vartheta^{K_3} = 0$. Der von Null verschiedene Faktor $(1-t)$ in der Spalte assoziiert zu den partiellen Ableitungen bzgl. ξ_1, impliziert, daß der Koeffizient ϑ^{v_1} auch Null ist. Eine genaue Betrachtung der Spalte, die mit den partiellen Ableitungen bzgl. v_1 assoziiert ist, und die Berücksichtigung der Bedingung $(v_1 - v_1^0) \neq 0$, ergeben $\vartheta^{m+p+1} = 0$. Dasselbe kann für die anderen Komponenten von $\vartheta^{(v,w)}$ (für ϑ^{v_1} ist das schon bekannt) unter Benutzung der geeigneten Teile der Matrix $-2I_{m+p}$ erhalten werden.

Unter Benutzung des parametrisierten Satzes von Sard gilt für fast alle $\mathcal{C}$, daß die durch die Festlegung von $\mathcal{C}$ erhaltene Abbildung auch Null als Regulärwert hat. Das liefert die Behauptung mit ähnlichen Argumenten wie für den Beweis von Lemma 1 in [61] (z.B. die Transversalitätsbedingung dort entspricht der Vollrangbedingung hier). Damit ist die Behauptung bewiesen. □

Beweis zur Behauptung 5 :

In diesem Beweis ist ein Parameterwert $\mathcal{C}$ festgelegt (o.B.d.A. $\mathcal{C} = 0$ und aus den Bezeichnungen weggelassen), so daß die Ergebnisse der Behauptung 4 erfüllt sind. Wir wollen zeigen, daß für fast alle $\mathcal{A}$ alle verallgemeinerten kritischen Punkte in (45) vom Typ 4 sind (Typ 5 ist wegen der Anzahl der Restriktionen ausgeschlossen). Da die Menge (45) abzählbar ist, reicht es, die

Aussage nur für einen festen Punkt $(\overline{y}, \overline{t})$ aus (45) zu zeigen. Wir beschränken uns im Beweis auf den Fall mit $t \neq 0$ und $t \neq 1$. Speziell folgt daraus, daß die Kompaktifizierungsungleichung aktiv in $(\overline{y}, \overline{t})$ sein muß. Hier haben K_1 und $\bar{K}_1$ dieselbe Bedeutung wie im vorigen Beweis der Behauptung 4, aber nur auf den Punkt $(\overline{y}, \overline{t})$ bezogen. Der Vektor $\bar{\xi}$ sei durch die folgende Gleichung festgelegt.

$$\sum_{k \in \overline{K_1}} \overline{\xi}_k D_y \overline{\Phi}_k(\overline{y}, \overline{t}) + D_y \overline{\Phi}_{m+p+1}(\overline{y}, \overline{t}) = 0$$

Aus den Ergebnissen der Behauptung 4 folgt:

- Der Vektor $\bar{\xi} \in I\!R^{|\overline{K_1}|}$ ist eindeutig bestimmt und alle Komponenten sind verschieden von Null.
- $D_t \mathbf{L}(\overline{y}, \overline{t}) \neq 0$ und die Matrix $W^T D_y^2 \mathbf{L}(\overline{y}, \overline{t}) W$ ist regulär, wobei, wie in Definition 17

  $$\mathbf{L}(y, t) := \sum_{k \in \overline{K_1}} \overline{\xi}_k \overline{\Phi}_k(y, t) + \overline{\Phi}_{m+p+1}(y, t)$$

  und die Spalten von W eine Basis des Unterraumes

  $$T = \left\{ \eta \in I\!R^{n+m+p} | D_y \overline{\Phi}_k(\overline{y}, \overline{t}) \eta = 0, k \in \overline{K_1} \right\}$$

  bilden.

Das zu zeigende Ergebnis kann nach Definition 17 und Bemerkung 2 auf die Gültigkeit der beiden Ungleichungen

$$D_y \overline{\Phi}_0(\overline{y}, \overline{t}, \mathcal{A}) W \neq 0 \tag{52}$$

$$D_y \overline{\Phi}_0(\overline{y}, \overline{t}, \mathcal{A}) W (W^T D_y^2 \mathbf{L}(\overline{y}, \overline{t}) W)^{-1} W^T D_y^T \overline{\Phi}_0(\overline{y}, \overline{t}, \mathcal{A}) \neq 0 \tag{53}$$

für fast alle $\mathcal{A}$ zurückgeführt werden.

Beginnen wir mit der Berechnung einer Matrix W. Da die Spalten von W eine Basis des orthogonalen Unterraums der Matrix

$$\begin{array}{ccc} D_x & D_v & D_w \\ \end{array}$$
$$\left[\begin{array}{ccc} D_x \overline{\Phi}_I(\overline{y}, \overline{t}) & (1-t) I_m & 0 \\ D_x \overline{\Phi}_{K_1}(\overline{y}, \overline{t}) & 0 & (1-t) I_{|K_1|} \,|\, 0 \end{array} \right]$$

bilden müssen, kann W wie folgt

$$W = \begin{pmatrix} -(1-t)I_n & 0 \\ D_x\overline{\Phi}_{\overline{K_1}}(\overline{y},\overline{t}) & 0 \\ 0 & I_{p-|K_1|} \end{pmatrix}$$

gewählt werden. Die Ungleichung (52) kann dann als

$$D_y\overline{\Phi}_0(\overline{y},\overline{t},(A,b))W = \begin{pmatrix} t(t-1)(\overline{x}^TA+b^T)+\chi_1^T & \chi_2^T \end{pmatrix} \neq 0$$

geschrieben werden, wobei $\chi_1 \in I\!R^n$ und $\chi_2 \in I\!R^{s-|K_1|}$ vom Parameter $\mathcal{A}$ unabhängige Vektoren sind. Da χ_1 fest ist, gilt für fast alle $\mathcal{A} = (A,b)$, daß der Teil $t(t-1)(\overline{x}^TA+b^T)+\chi_1^T$ von $D_y\overline{\Phi}_0(\overline{y},\overline{t},(A,b))W$ ungleich Null ist.

Die Ungleichung (53) bedarf eines Ausdrucks für $(W^TD_y^2\mathbf{L}(\overline{y},\overline{t})W)^{-1}$. Diese Matrix ergibt sich nach Definition mit Hilfe der Multiplikation.

$$W^T \begin{pmatrix} \Delta_1 & 0 & 0 \\ 0 & -2I_{m+|K_1|} & 0 \\ 0 & 0 & -2I_{p-|K_1|} \end{pmatrix} W$$

Diese führt zu den Ausdruck

$$(W^TD_y^2\mathbf{L}(\overline{y},\overline{t})W)^{-1} = \begin{pmatrix} \Delta_2 & 0 \\ 0 & -2I_{p-|K_1|} \end{pmatrix},$$

wobei Δ für Teilmatrizen steht, die uninteressant für die Analyse sind, speziell hängen sie nicht vom Parameter $\mathcal{A}$ ab. Der Ausdruck Δ_2 ist eine reguläre $n \times n$-Matrix. Die Ungleichung (53) bekommt jetzt die folgende Gestalt.

$$(t(t-1)(\overline{x}^TA+b^T)+\chi_1^T)\Delta_2(t(t-1)(A\overline{x}+b)+\chi_1) - 2||\chi_2||^2 \neq 0$$

Die Regularität von Δ_2 und die Unabhängigkeit der Ausdrücke Δ_2, ξ_1 und ξ_2 von $\mathcal{A} = (A,b)$ implizieren die Gültigkeit von (53) für fast alle $\mathcal{A}$. Damit ist dieser Beweis beendet. $\square$

Alle diese Behauptungen sind für $t \neq 0$ und $t \neq 1$ bewiesen, es sollen noch zwei Fälle bedacht werden. Für $t = 0$ haben die Parameter keinen Einfluß auf das Problem. Mit dem Lemma 3 wird gezeigt, daß die Regularität in $t = 0$ durch die spezielle Selektion der Vektoren (x^0, w^0) und (x^1, w^1) gewährleistet wird (siehe (4.2)).

Lemma 3
Jeder verallgemeinerte kritische Punkt des Problems

$$\begin{array}{ll} \min & ||x-x^0||^2+||v-v^0||^2+||w-w^0||^2 \\ \text{s.t.} & \\ & v-v^0=0, \\ & w-w^1\geq 0, \\ & -||x-x^1||^2-||v-v^0||^2-||w-w^1||^2\geq -q \end{array} \tag{54}$$

ist nichtdegeneriert.

<u>Beweis</u>:

Die LICQ ist offensichtlich in allen zulässigen Punkten von (54) erfüllt. Einfachheitshalber nehmen wir an, daß keine Gleichungsrestriktionen vorhanden sind. Wir zeigen ND1.

Ist eine lineare Restriktion aktiv, dann verschwindet die entsprechende partielle Ableitung der quadratischen Restriktion $(-2(w_j-w_j^1))$. Die Ungleichung $w_j^0\neq w_j^1$ impliziert dann, daß der zugehörige Multiplikator ungleich Null ist.

Ist der Multiplikator der quadratischen Ungleichung Null, dann folgt $x=x^0$ und außerdem, für alle $j\in J$, ist w_j entweder gleich w_j^0 oder gleich w_j^1 . Das, zusammen mit den Ungleichungen

$$||x^0-x^1||^2+||w-w^1||^2\leq ||x^0-x^1||^2+||w^0-w^1||^2<q$$

spricht dagegen, daß die Kompaktifizierung aktiv ist. Damit ist ND1 klar.

ND2 wird bewiesen, indem die strikte Definitheit der Hesse-Matrix der Lagrange-Funktion gezeigt wird. Da die linearen Restriktionen keinen Beitrag zur genannten Hesse-Matrix liefern, reicht es, nur die Hesse-Matrizen der Zielfunktion und der quadratischen Restriktion zu analysieren. Beide sind aber Einheitsmatrizen, nur mit verschiedenem Vorzeichen.

Die strikte Definitheit folgt dann, wenn der zur Kompaktifizierungsrestriktion assoziierte Multiplikator verschieden von -1 ist. Das ist aber eine Konsequenz der Ungleichung $x^0\neq x^1$. $\square$

Für $t=1$ betrachten wir folgendes Ergebnis.

Lemma 4
Für fast alle $\mathcal{P}$ sind alle verallgemeinerten kritischen Punkte des Problems

$$\begin{array}{ll} \min & \overline{f}(x) + 0.5x^T Ax + b^T x + ||v - v^0||^2 + ||w - w^0||^2 \\ \text{s.t.} & \\ & \overline{h}_i(x) + c_i^T x + d_i = 0, \ i \in I, \\ & \overline{g}_j(x) + c_{m+j}^T x + d_{m+j} \geq 0, \quad j \in J, \\ & -||x - x^1||^2 - ||v - v^0||^2 - ||w - w^1||^2 \geq -q \end{array} \tag{55}$$

vom Typ 1.

<u>Beweis</u>:

Man kann leicht unter Benutzung des parametrisierten Satzes von Sard und der Gültigkeit der LICQ bezüglich der Kompaktifizierungsrestriktion zeigen, daß für fast alle $\mathcal{P}$ die LICQ in jedem zulässigen Punkt vom Problem (55) erfüllt ist.

Es reicht dann festzustellen, daß für fast alle $\mathcal{P}$ jeder kritische Punkt von (55) nichtdegeneriert ist.

Betrachten wir zunächst den Fall $x \neq x^1$. Es seien Z_1 eine aktive Indexmenge mit $\{1, \ldots, m\} \subset Z_1 \subset \{1, \ldots, m+p+1\}$, eine Teilmenge davon Z_2 und

$$\Upsilon(y, \xi, \mathcal{P}) = \begin{pmatrix} D_y \overline{\Phi}_0(y, 1, \mathcal{A}) - \sum_{k \in Z_1} \xi_k D_y \overline{\Phi}_k(y, 1, \mathcal{C}) \\ \overline{\Phi}_k(y, 1, \mathcal{C}), \quad k \in Z_1 \\ \xi_{Z_2} \end{pmatrix}$$

die zugehörige Abbildung.

Zur Benutzung des parametrisierten Satzes von Sard zeigen wir, daß Null ein Regulärwert von $\Upsilon(y, \xi, \mathcal{P})$ ist, eingeschränkt auf die offene Menge $\{(y, \xi, \mathcal{P}) | x \neq x^1\}$.

Betrachten wir den folgenden Teil der Jacobi-Matrix von $\Upsilon(y, \xi, \mathcal{P})$.

$$\begin{array}{ccccc} D_x & D_{(v,w)} & D_{\xi_{Z_3}} & D_b & D_{d_{Z_1 \setminus \{m+p+1\}}} \\ \left[\begin{array}{c} \otimes \\ 0 \\ D_x \overline{\Phi}_{Z_1 \setminus \{m+p+1\}} \\ -2(x-x^1)^T \\ 0 \end{array}\right. & \begin{array}{c} 0 \\ 2(1+\xi_{m+p+1}) I_{m+p} \\ 0 \\ -2(v-v^0)^T | -2(w-w^1)^T \\ 0 \end{array} & \begin{array}{c} \otimes \\ 0 \\ 0 \\ 0 \\ I_{|Z_3|} \end{array} & \begin{array}{c} I_n \\ 0 \\ 0 \\ 0 \\ 0 \end{array} & \left.\begin{array}{c} 0 \\ 0 \\ I_{Z_1 \setminus \{m+p+1\}} \\ 0 \\ 0 \end{array}\right] \end{array} \tag{56}$$

Wenn $m+p+1 \notin Z_1$, sind die zugehörigen Einträge von (56) nicht vorhanden, insbesondere nicht die vorletzte Gruppe von Zeilen (die Ableitung der Kompaktifizierung). Man betrachtet dann die zu den Ableitungen nach b, v, w, d_{Z_1} und ξ_{Z_3} gehörenden Spalten und findet eine reguläre Matrix.

Falls $m+p+1 \in Z_1$, kommt man zur selben Analyse, muß aber beachten, daß $\xi_{m+p+1} \neq -1$, denn $w^0 \neq w^1$ und $x \neq x^1$.

Aus dem parametrisierten Satz von Sard folgt jetzt, daß die Abbildung $\Upsilon_{\mathcal{P}}(y,\xi)$, eingeschränkt auf $\{(y,\xi)|x \neq x^1\}$, für fast alle $\mathcal{P}$ Null als Regulärwert hat. Daraus folgt, wenn $\Upsilon_{\mathcal{P}}^{-1}(0) \neq \emptyset$, daß $Z_3 = \emptyset$ (entspricht ND1) und daß die Jacobi-Matrix von $\Upsilon_{\mathcal{P}}$ regulär auf $\{(y,\xi)|x \neq x^1\}$ ist, was ND2 entspricht.

Die Menge solcher Parameterwerte $\mathcal{P}$, für die das Problem (55) einen kritischen Punkt mit $x = x^1$ enthält, ist vom Maß Null. Es sei bemerkt, daß die folgende Beziehung für eine beliebige Indexmenge $J_0 \subset J$ erfüllt sein muß.

$$\begin{aligned} D\bar{f}(x^1) + Ax^1 + b \quad \in \quad & span\{D\bar{h}_i(x^1) + c_i, i \in I, D\bar{g}_j(x^1) + c_j, j \in J_0\}, \\ & \overline{h}_i(x^1) + c_i^T x^1 + d_i = 0, \ i \in I, \\ & \overline{g}_j(x^1) + c_{m+j}^T x^1 + d_{m+j} = 0, \quad j \in J_0 \end{aligned}$$

Betrachtet man die verschiedenen Teilmengen nach der Selektion von J_0, stellt man fest, daß jede einzelne eine Mannigfaltigkeit niedrigerer Dimension als der Parameter $\mathcal{P}$ ist und deswegen das Maß Null haben muß. Z.B.: Für $I \cup J_0 \neq \emptyset$ betrachtet man nur die Zulässigkeit sowie die Ableitungen nach d. Für $I \cup J_0 = \emptyset$ wird die Inklusion eine Gleichung, so daß die Betrachtung der Ableitungen nach b ausreicht.

Damit ist der Fall $x = x^1$ abgeschlossen und der Beweis dieses Lemmas beendet. □

Als abschließende Bemerkung des Beweises und dieses Abschnittes achten wir darauf, daß das Teilergebnis für den Fall $t = 1$ eine meßbaren Menge ergibt. Damit kann man, ähnlich wie im vorigen Abschnitt, die Behauptung der Proposition 1 in zwei Fällen betrachten (nämlich $t = 1$ und $t \in [0,1)$) ohne Schwierigkeiten mit den jeweiligen Aussagen zu bekommen.

Damit betrachten wir den Beweis der Generizität beendet.

Bemerkung 10

Die Einbettung gegeben in (41) wurde in [11] sowohl theoretisch als numerisch

untersucht. Der Satz 18 liefert eine theoretische Rechtfertigung der Regularität in Sinne Jongen, Jonker und Twilt dieser speziellen Einbettung. In diesem Sinne liefert dieser Satz eine neue und wichtige Eigenschaft dieser Einbettung.

Der Beweis folgt Zügen des Beweises des allgemeinen Störungssatzes bewiesen von Rückmann und Tammer in [61]. Er ist aber nicht eine direkte Folgerung. Es mußten verschiedene Aspekte neu betrachtet werden. Einerseits ist wegen der wenigen für die Störung verfügbaren Parameter der Beweis des Schrittes 1 schwieriger, insbesondere der Beweis der Voraussetzungen des parametrisierten Satzes von Sard. Dies betrifft auch den Punkt vom Typ 4 (Schritt 2). Andererseits ist die getrennte Betrachtung der Punkte in $t = 0$ *und* $t = 1$ *und die daraus resultierenden Ergebnisse in der oben genannte Arbeit nicht vorhanden (weil sie dort nicht notwendig ist).*

5 Regularität einparametrischer Variationsungleichungen.

In diesem Abschnitt beschäftigen wir uns mit einem Problem, das eng mit Optimierungsproblemen zusammenhängt, aber auch von eigenständigem Interesse ist, die Lösung von Variationsungleichungen. Wir definieren Lösungen dieser Ungleichungen und führen Bedingungen zur Charakterisierung der Lösungen an, wobei auch Lösungsansätze und äquivalente Formulierungen des Problemes präsentiert werden.

Unser Ziel in diesem Kapitel ist die Betrachtung einparametrischer Variationsungleichungen und einer Regularität in dieser Klasse, die der JJT-Regularität von einparametrischen Optimierungsproblemen ähnlich ist. Dazu gehört natürlich eine eventuelle Neudefinition der Singularitäten und die Untersuchung der Struktur und der Interpretation solcher Ergebnisse für die zugrundeliegende Aufgabe. Wegen der verschiedenen Natur beider Aufgaben sind solche Schritte keine direkte Wiederholung des Falls von Optimierungsproblemen. Es sind dabei an verschiedenen Stellen neue Argumente und Blickpunkte erforderlich. Die Betrachtung solcher Regularität im Rahmen der Variationsungleichungen ist ein neuer Beitrag zu der bisherigen Entwicklung dieses Problems.

Nachdem diese Regularität inklusive ihrer Generizität definiert und untersucht worden ist, kommen wir zu Aussagen, die die Benutzung von Kurvenverfolgungsprozeduren gewährleisten. Diese Aussagen basieren auf der Benutzung von Eigenschaften der eingeführten Regularität.

5.1 Variationsungleichungen.

Definition 22
Es seien eine abgeschlossene Menge $M \subset \mathbb{R}^n$ und eine Abbildung $F : M \to \mathbb{R}^n$ gegeben. Das Problem, einen Punkt $x \in M$ zu bestimmen, in welchem

$$F(x)(y - x) \geq 0, \quad \forall y \in M \tag{57}$$

gilt, wird Variationsungleichung genannt.

Dieses Problem hat zahlreiche Anwendungen und modelliert mehrere interessante und wichtige Probleme sowohl in der Wirtschaft als auch in der Technik. Einige ausgewählte sind Nasch Equilibria von nichtkooperativen Spielen, traffic assignment oder Netzwerk Equilibria. Zu grundlegenden Gedanken und zahlreichen Literaturangaben bezüglich dieses Problems verweisen wir auf die Übersichtsarbeit [20] und das Buch [48].

In Abhängigkeit von der Gestalt und den Eigenschaften der Menge M oder der Abbildung F entspricht die Variationsungleichung verschiedenen wichtigen bekannten Problemen. Der wichtigste Fall ist, wenn die Menge M konvex ist. Dann ist das in (57) gesuchte globale Verhalten relativ einfach zu beschreiben, insbesondere durch den Normalenkegel. Ist die Menge M konvex und F eine Gradient-Abbildung, dann entsprechen die Lösungen der Variationsungleichung stationären Punkten. Andere wichtige Fälle sind z.B. $M = I\!R^n$ (woraus die Gleichung $F(x) = 0$ resultiert) und $M = I\!R^n_+$ (wobei die im allgemeinen nichtlinearen Komplementaritätsprobleme entstehen).

Wir konzentrieren uns in dieser Arbeit auf den Fall, daß sich die Menge M, genau wie für Optimierungsprobleme, durch endlich viele Gleichungen und Ungleichungen, bestimmt durch die Abbildungen $(H, G) : I\!R^n \to I\!R^{m+p}$, in der Form

$$M = \{y \in I\!R^n | H(x) = 0, G(x) \geq 0\}$$

schreiben läßt. Dabei sind die Indexmengen $I = \{1, \ldots, m\}$ und $J = \{1, \ldots, p\}$. Damit sind alle mit den zulässigen Mengen verbundenen Begriffe der Optimierungsprobleme, wie z.B. Constraint qualifications LICQ und MFCQ oder aktive Indexmenge, ohne Mißverständnisse auch für Variationsungleichungen zu benutzen.

Eine Variationsungleichung wird durch ihre Daten definiert, die nichts anderes als die Abbildungen $(F, H, G) : I\!R^n \to I\!R^{n+m+p}$ sind.

Definition 23
Es sei eine Variationsungleichung $V(F, H, G)$ gegeben. Eine lokale Lösung dieser Variationsungleichung ist ein Punkt $x \in M$, der die Beziehung (57) zumindest auf eine Menge $M \cap U$ erfüllt, wobei U eine Umgebung von x ist.

Da die Beziehung (57) einer Optimalitätseigenschaft des Punktes x bezüglich der linearen Funktion $g(y) = F(x) \cdot y$ auf der zulässigen Menge M entspricht, sind die Optimalitätsbedingungen hier erfüllt. Das heißt aber,

daß jede Lösung ein Fritz-John-Punkt ist und daß unter einer Constraint qualification auch ein stationärer Punkt entsteht.

Definition 24
Es sei eine Variationsungleichung $(F, H, G) \in C^1(\mathbb{R}^n, \mathbb{R}^{n+m+p})$ *gegeben. Ein zulässiger Punkt* $x \in M$ *heißt stationär, falls die folgende Beziehung*

$$F(x) - \sum_{i=1}^{m} \lambda_i Dh_i(x) - \sum_{j \in J_0(x)} \mu_j Dg_j(x) = 0 \tag{58}$$

mit $\mu_j \geq 0$, $\forall j \in J_0(x)$ *gilt..*

Ist ein Punkt stationär, so heißt das noch nicht, daß er auch eine Lösung ist, auch nicht lokal. Der komplizierteste Punkt bei der Lösung ist die Gültigkeit der Ungleichung für die ganze Menge. Eine andere wichtige Angelegenheit ist die komplette Darstellung der Normalkegel, was sich aber unter Constraints Qualifications vereinfachen läßt.

Eine Nullstelle von F (auf M) ist immer eine Lösung der Variationsungleichung. Ist aber ein stationärer Punkt x keine Nullstelle, so sind Bedingungen zweiter Ordnung, wie im ersten Kapitel, nötig. Da sich diese Bedingungen auf die Variable der Minimierung, also y, beziehen, hängen sie nicht von der Abbildung F ab.

Ist die LICQ erfüllt und $x \in M$ ein stationär Punkt, so gelten für die symmetrische Matrix

$$-\sum_{i=1}^{m} \lambda_i D^2 h_i(x) - \sum_{j \in J_0(x)} \mu_j D^2 g_j(x) \tag{59}$$

zwei Bedingungen:

- Ist x eine lokale Lösung der Variationsungleichung, so ist die Matrix (59) positiv semidefinit in dem Tangentialraum.
- Ist die Matrix (59) positiv definit, dann ist x zumindest eine lokale Lösung der Variationsungleichung.

Wenn die Menge M konvex ist, dann wird die ganze Überlegung einfacher. Wir nehmen in diesem konvexen Fall an, daß die Funktionen g_j konkav sind

und die h_i linear affin. Einerseits sind alle lokalen Lösungen auch globale. Andererseits sind unter der LICQ alle Matrizen $D^2 g_j(x)$ für $j \in J_0(x)$ in dem Tangentialraum negativ semidefinit.

Für den Fall, daß M konvex ist und eine Constraints qualification gilt, z.B. MFCQ, besteht eine direkte Beziehung zwischen den Lösungen der Variationsungleichung und den Lösungen des folgenden KKT-Systems.

$$\begin{array}{lcl} F(x) - \sum_{i=1}^{m} \lambda_i Dh_i(x) - \sum_{j=1}^{p} \mu_j Dg_j(x) & = & 0 \\ H(x) & = & 0 \\ \mu^T G(x) & = & 0 \\ G(x) & \geq & 0 \\ \mu & \geq & 0 \end{array} \tag{60}$$

Dieser konvexe Fall ist eingehend untersucht worden und verschiedene Umformulierungen der Variationsungleichung, die für verschiedene Lösungsansätze geeignet sind, liegen vor. Eine große Klasse solcher Umformulierungen bezieht sich direkt auf die Darstellung des Systems (60) als Gleichungssystem, das im allgemein nicht differenzierbar ist. Dafür sind dann die modernen Methoden der Lösung nichtglatter Gleichungssysteme mittels verallgemeinerter Newton-Methoden einsetzbar. Für einen Überblick in dieser Richtung und Referenzen verweisen wir auf [48].

Eine andere mögliche Umformulierung ist die Darstellung von Variationsungleichungen mit Hilfe von Optimierungsproblemen. Man kann die Variationsungleichung mittels einer gap-Funktion oder einer merit-Funktion in ein sowohl restringiertes, als auch nicht restringiertes globales Optimierungsproblem überführen. Dabei sind im allgemeinen die vorhandenen Funktionen oder ihre Ableitungen nicht differenzierbar, und wir sind wieder auf die Effizienz der Newton-Verfahren für nichtglatte Gleichungssysteme angewiesen. Dazu verweisen wir auf [48], [20], [21], [22], [9], [8], [5], [28], [49].

5.2 Einparametrische Variationsungleichungen.

In diesem Abschnitt definieren wir die Singularitäten der JJT-Regularität für den Fall der einparametrischen Variationsungleichungen und die Klasse der regulären Probleme. Ausgangspunkt dafür ist der Begriff eines stationären Punktes (siehe Definition 24) und seine natürliche Verallgemeinerung zum kritischen Punkt und zum verallgemeinerten kritischen Punkt.

Definition 25
Es sei eine Variationsungleichung $(F, H, G) \in C^1(I\!R^n, I\!R^{n+m+p})$ *gegeben. Ein zulässiger Punkt* $x \in M$ *heißt kritischer Punkt, falls die Beziehung (58) gilt (nicht unbedingt mit positiven* μ*).*

Andererseits heißt x *verallgemeinerter kritischer Punkt, falls die Vektoren*

$$\{F(x), Dh_i(x), Dg_j(x), i \in I, j \in J_0(x)\}$$

linear abhängig sind.

Betrachten wir Variationsungleichungen, die von einem zusätzlichen eindimensionalen reellen Parameter abhängen. Darunter verstehen wir die in den Variablen $z = (x, t) \in I\!R^{n+1}$ definierten Abbildungen (F, H, G), so daß für jeden festen Paramerter $t \in I\!R$ die durch die Daten $(F(., t), H(., t), G(., t))$ gegebeneVariationsungleichung entsteht. Damit sind die für Variationsungleichungen eingeführten Begriffe eines stationären, kritischen oder verallgemeinerten kritischen Punktes auch in Bezug auf einparametrische Variationsungleichungen verwendbar. In diesem Zusammenhang sind auch für Variationsungleichungen die Mengen $M(t)$, Σ_{gc} und Σ_{stat} ohne Mißverständnisse erklärt.

Die Definition regulärer einparametrischer Variationsungleichungen verläuft über die endliche und generische Klassifizierung der lokalen Struktur der Menge Σ_{gc} aller verallgemeinerten kritischen Punkte.

Definition 26
Ein verallgemeinerter kritischer Punkt $\bar{z}$ *einer Variationsungleichung* (F, H, G) *heißt nichtdegeneriert, falls die LICQ und zwei weitere Bedingungen erfüllt sind.*

VU-ND1 *: Für die eindeutige Lösung* $(\bar{\lambda}, \bar{\mu}) \in I\!R^{m+|J_0(z)|}$ *des Systems (58) gilt* $\mu_j \neq 0, \ \forall j \in J_0(z)$.

VU-ND2 *: Die Matrix* $D_x\mathbb{L}(\bar{z}, \bar{\lambda}, \bar{\mu})|_{T_{\bar{x}}M(\bar{t})}$ *ist nicht singulär.*

Dabei wird die Abbildung $\mathbb{L}_{J_0} : I\!R^{n+1+m+|J_0(z)|} \to I\!R^n$ als

$$\mathbb{L}_{J_0}(z, \lambda, \mu) = F(z) - \sum_{i=1}^{m} \lambda_i Dh_i(z) - \sum_{j \in J_0(z)} \mu_j Dg_j(z)$$

definiert und der lineare Unterraum $T_{\bar{x}}M(\bar{t})$ als der Tangentialraum an die Menge $M(\bar{t})$ in $\bar{x}$ verstanden.

Die Matrix $D_x\mathbb{L}_{J_0}(\bar{z},\bar{\lambda},\bar{\mu})$ ist im Unterschied zu Optimierungsproblemen im allgemeinen nicht symmetrisch. Hier wird $D_x\mathbb{L}_{J_0}(\bar{z},\bar{\lambda},\bar{\mu})|_{T_{\bar{x}}M(\bar{t})}$ als eine Matrix der Form $V^T D_x\mathbb{L}_{J_0}(\bar{z},\bar{\lambda},\bar{\mu})V$ verstanden, wobei die Spalten von V eine Basis von $T_{\bar{x}}M(\bar{t})$ bilden. Sind V_1 und V_2 zwei solche Basen, dann gilt die Beziehung $V_1 = V_2\left[V_2^T V_2\right]^{-1} V_2^T V_1$. Damit entsteht die Gleichung:

$$V_1^T D_x\mathbb{L}(\bar{z},\bar{\lambda},\bar{\mu})V_1 = V_1^T V_2\left[V_2^T V_2\right]^{-1} V_2^T D_x\mathbb{L}_{J_0}(\bar{z},\bar{\lambda},\bar{\mu})V_2\left[V_2^T V_2\right]^{-1} V_2^T V_1.$$

Da die Matrix $V_2^T V_1$ regulär ist (ihre Spalten sind Basen desselben linearen Teilraumes), sind die Regularität von $V^T D_x\mathbb{L}_{J_0}(\bar{z},\bar{\lambda},\bar{\mu})V$ und auch das Vorzeichen seiner Determinanten nicht von der eventuellen Selektion der Matrix V abhängig.

Für die Untersuchung der lokalen Struktur der kritischen Mengen spielt für die Variationsungleichungen die Abbildung

$$\mathbb{H}_{J_0(z)}(z,\lambda,\mu) = \begin{pmatrix} \mathbb{L}_{J_0}(z,\lambda,\mu) \\ H(z) \\ G_{J_0(z)}(z) \end{pmatrix}$$

eine wichtige Rolle, denn ihre Nullstellen einschließlich der Multiplikatoren repräsentieren lokal die Menge derjenigen kritischen Punkte, die zu einer aktiven Indexmenge gehören.

Bemerkung 11
In Analogie mit Optimierungsproblemen ist die Bedingung **VU-ND2** *unter der LICQ auch äquivalent zur Regularität von $D_{(x,\lambda,\mu)}\mathbb{H}_{J_0(z)}(z,\lambda,\mu)$, der Unterschied besteht in ihrer Symmetrie im Optimierungsfall. Daher wird die Menge der kritischen Punkte unter der* **VU-ND1** *lokal durch die Nullstellen von $\mathbb{H}_{J_0(z)}$ repräsentiert, und sie bildet eine glatte, nach dem Parameter t parametrisierbare Mannigfaltikeit.*

Die Nichtsymmetrie der Matrix $D_x\mathbb{L}_{J_0}$ verhindert die Definition der quadratischen Indizes mittels der Vorzeichen der reellen Eigenwerte. Andererseits kann man bei nichtdegenerierten kritischen Punkten die linearen Indizes LI und LCI und das Vorzeichen von $\det(D_x\mathbb{L}_{J_0}(\bar{z},\bar{\lambda},\bar{\mu})|_{T_{\bar{x}}M(\bar{t})})$ eindeutig bestimmen. Falls der Tangentialraum der lineare Teilraum $\{0\}$ ist, betrachten

wir nach Definition $D_x\mathbb{L}_{J_0}(\bar{z},\bar{\lambda},\bar{\mu})|_{T_{\bar{x}}M(\bar{t})}$ als regulär und ihre Vorzeichen als $+1$. Im konvexen Fall ist eine Lösung durch die Eigenschaft $LI = 0$ charakterisiert.

5.3 Singularitäten und lokale Eigenschaften

Für die Definition der Singularitäten der einparametrischen Variationsungleichungen gehen wir von der Definition der 5 Typen der Menge $\mathcal{F}$ (Definitionen 14 bis 18) aus und betonen die entsprechenden Unterschiede bei Verletzung der Nichtdegeneriertheit.

Definition 27 *(vergleiche Definition 14)*
Ein verallgemeinerter kritischer Punkt $\bar{z}$ einer Variationsungleichung ist vom Typ 1, falls er nichtdegeneriert im Sinne der Definition 26 ist.

Bemerkung 12
Die lokale Struktur von Σ_{gc} ist genau dieselbe wie im Fall einparametrischer Optimierungsprobleme und wird im Bild 1 dargestellt. In der bezüglich Σ_{gc} relativen Umgebung von $\bar{z}$ bleibt die Nichtdegeneriertheit erfüllt, und die linearen Indizes LI, LCI sowie das Vorzeichen von $\det(D_x\mathbb{L}_{J_0}(\bar{z},\bar{\lambda},\bar{\mu},\bar{t})|_{T_{\bar{x}}M(\bar{t})})$ *sind konstant.*

Definition 28 *(vergleiche Definition 15)*
Ein verallgemeinerter kritischer Punkt $\bar{z} = (\bar{x},\bar{t})$ einer Variationsungleichung $(F,H,G) \in C^2(\mathbb{R}^{n+1},\mathbb{R}^{n+m+p})$ ist vom Typ 2, falls die Bedingungen **1.** *und* **2.** *wie in der Definition 15 gelten und folgende Umformulierungen der anderen drei Bedingungen erfüllt sind:*

Typ2: 3-VU *: $D_x\mathbb{L}_{J_0}(\bar{z},\bar{\lambda},\bar{\mu})|_{T_{\bar{x}}M(\bar{t})}$ ist regulär.*

Typ2: 4-VU *: $D_x\mathbb{L}_{J_0}(\bar{z},\bar{\lambda},\bar{\mu})|_{T^+_{\bar{x}}M(\bar{t})}$ ist regulär, wobei $T^+_{\bar{x}}M(\bar{t})$ genauso wie in Bedingung* **3.** *der Definition 15 verstanden wird.*

Typ2: 5-VU *: Bedingung* **5.** *der genannten Definition ist erfüllt, aber mit β wie folgt definiert:*

$$\beta \quad = \quad -W(W^T \cdot D_x\mathbb{L}_{J_0} \cdot W)^{-1}W^T \cdot [D_x\mathbb{L}_{J_0} \cdot \alpha + D_t\mathbb{L}_{J_0}]\,,$$

W und α genommen wie in Definition 15.

Bemerkung 13
Bei Betrachtung der Nullstellenmenge von $\mathbb{H}_{J_0^+}$ und $\mathbb{H}_{J_0}$, der LICQ und der Regularität in den Bedingungen **Typ2: 3-VU** *und* **Typ2: 4-VU** *ergibt sich ähnlich wie im Fall der Optimierungsprobleme sind um $\bar{z}$ zwei Kurven $x^0(t)$ (bzw. $x^1(t)$) definiert, die lokal Σ_{gc} beschreiben. Wegen der Bedingung* **Typ2: 5-VU** *haben sie keine parallelen Tangentialräume. Die lokale Struktur der Menge Σ_{gc} ist daher ähnlich der der Optimierungsprobleme.*

Zur Beschreibung der Änderung der linearen Indizes und des Vorzeichens von $\det(D_x\mathbb{L}_{J_0}(\bar{z},\bar{\lambda},\bar{\mu},\bar{t})|_{T_{\bar{x}}M(\bar{t})})$ entlang der Kurven verwenden wir die Zahlen $sign(\gamma)$ und $sign(\delta)$, wobei

$$\gamma = Dg_p(x^0(\bar{t}),\bar{t}) = D_x g_p(\bar{z}) \cdot D_t x^0(\bar{t}) + D_t g_p(\bar{z}) \neq 0 \tag{61}$$

und

$$\delta = \det(D_x\mathbb{L}_{J_0}(\bar{z},\bar{\lambda},\bar{\mu})|_{T_{\bar{x}}M(\bar{t})}) \cdot \det(D_x\mathbb{L}_{J_0}(\bar{z},\bar{\lambda},\bar{\mu})|_{T^+_{\bar{x}}M(\bar{t})}) \tag{62}$$

sind.

Die Änderung der genannten Größen entlang der beiden Kurven ist nicht willkürlich und erfüllt die Bedingung.

Behauptung 6
Es sei $\bar{z}$ ein verallgemeinerter kritischer Punkt einer Variationsungleichung vom Typ 2. Mit $(x^1(t),\mu(t),t)$ sei eine Parametrisierung der Nullstellen von $\mathbb{H}_{J_0}$ nach t um den Punkt $(\bar{x},\bar{\mu},\bar{t})$ bezeichnet. $D\mu_p(\bar{t})$ sei die Ableitung des verschwindenden Multiplikators. Dann gilt:

$$sign(\gamma).sign(\delta) = -sign(D\mu_p(\bar{t}))\ .$$

Zum Beweis dieser Behauptung benötigen wir das folgende Lemma über das Schurkomplement einer Matrix.

Definition 29
Eine Matrix A sei zerlegt in

$$A = \begin{pmatrix} B & C \\ D & E \end{pmatrix}, \tag{63}$$

wobei B und E quadratisch sind. Außerdem sei B regulär. Die Matrix

$$S(A|B) = E - DB^{-1}C$$

heißt das Schurkomplement von B in A.

Lemma 5 *(siehe [52])*
Es sei A eine Matrix mit den Eigenschaften aus Definition 29. Dann gelten folgende Beziehungen:

1. $\det(A) = \det(E) \cdot \det(S(A|B))$.
2. *Wenn A symmetrisch ist, so gilt:* $In(A) = In(B) + In(S(A|B))$.
3. *Wenn A regulär ist, so gilt:*

$$A^{-1} = \begin{pmatrix} B^{-1} + B^{-1} \cdot C \cdot S(A|B)^{-1} \cdot D \cdot B^{-1} & -B^{-1} \cdot C \cdot S(A|B)^{-1} \\ -S(A|B)^{-1} \cdot D \cdot B^{-1} & S(A|B)^{-1} \end{pmatrix}.$$

Beweis zur Behauptung 6:

Es sei V_1 eine Basis von $T_{\bar{x}}M(\bar{t})$ und die Matrix $V_0 = \left[V_1 \vdots v_1\right]$ eine Basis von $T_{\bar{x}}^+M(\bar{t})$. Das ist möglich, weil $T_{\bar{x}}M(\bar{t}) \subset T_{\bar{x}}^+M(\bar{t})$. Dann gilt:

$$D_x\mathbb{L}_{J_0}|_{T_{\bar{x}}^+M(\bar{t})} = \begin{pmatrix} V_1^T D_x\mathbb{L}_{J_0} V_1 & V_1^T D_x\mathbb{L}_{J_0} v_1 \\ v_1^T D_x\mathbb{L}_{J_0} V_1 & v_1^T D_x\mathbb{L}_{J_0} v_1 \end{pmatrix}.$$

Nach Bedingung **Typ2: 3-VU** ist $V_1^T D_x\mathbb{L}_{J_0} V_1$ regulär, und aus 1. aus Lemma 5 folgt dann:

$$\det(D_x\mathbb{L}_{J_0}|_{T_{\bar{x}}^+M(\bar{t})}) = \det(D_x\mathbb{L}_{J_0}|_{T_{\bar{x}}M(\bar{t})}) \det(S(D_x\mathbb{L}_{J_0}|_{T_{\bar{x}}^+M(\bar{t})}|D_x\mathbb{L}_{J_0}|_{T_{\bar{x}}M(\bar{t})})).$$

Dabei wollen wir $S(D_x\mathbb{L}_{J_0}|_{T_{\bar{x}}^+M(\bar{t})}|D_x\mathbb{L}_{J_0}|_{T_{\bar{x}}M(\bar{t})})$ als

$$\begin{aligned} S(D_x\mathbb{L}_{J_0}|_{T_{\bar{x}}^+M(\bar{t})}|D_x\mathbb{L}_{J_0}|_{T_{\bar{x}}M(\bar{t})}) &= v_1^T D_x\mathbb{L}_{J_0} v_1 - \\ & \quad v_1^T D_x\mathbb{L}_{J_0} V_1 \left(V_1^T D_x\mathbb{L}_{J_0} V_1\right)^{-1} V_1^T D_x\mathbb{L}_{J_0} v_1 \end{aligned} \tag{64}$$

verstehen. Daher ergibt sich folgende Beziehung.

$$sign(\delta) = sign\left(S(D_x\mathbb{L}_{J_0}|_{T_{\bar{x}}^+M(\bar{t})}|D_x\mathbb{L}_{J_0}|_{T_{\bar{x}}M(\bar{t})})\right) \tag{65}$$

Aus der Gleichheit der Ableitungen der Komponenten (H, G_{J_0}) von $\mathbb{H}_{J_0^+}$ und $\mathbb{H}_{J_0}$ erhält man die Beziehung $D_t x^0(\bar{t}) - D_t x^1(\bar{t}) \in T_{\bar{x}}^+M(\bar{t})$, und daher $D_t x^0(\bar{t}) - D_t x^1(\bar{t}) = V_0 w_0$. Durch Multiplikation der Ableitungen der Komponente $\mathbb{L}_{J_0}$ von links mit der Matrix V_0 folgt die Beziehung:

$$V_0^T D_x\mathbb{L}_{J_0} V_0 w_0 + D_t\mu_p(\bar{t}) V_0^T D_x g_p(\bar{z}) = 0.$$

Daher wird dann die folgende Gleichung erhalten.

$$D_t x^0(\bar{t}) - D_t x^1(\bar{t}) = -D_t \mu_p(\bar{t}) \cdot V_0 \left(V_0^T D_x \mathbb{L}_{J_0} V_0\right)^{-1} V_0^T D_x^T g_p(\bar{z}) \qquad (66)$$

Andererseits gilt:

$$\gamma = D_x g_p(\bar{z}) \cdot D_t x^0(\bar{t}) + D_t g_p(\bar{z}) = D_x g_p(\bar{z}) \left(D_t x^0(\bar{t}) - D_t x^1(\bar{t})\right),$$

denn $D_x g_p(\bar{z}) \cdot D_t x^1(\bar{t}) + D_t g_p(\bar{z}) = 0$. Nach Einsetzen der Formel (66) ergibt sich:

$$\gamma = -D_t \mu_p(\bar{t}) \cdot D_x g_p(\bar{z}) V_0 \left(V_0^T D_x \mathbb{L}_{J_0} V_0\right)^{-1} V_0^T D_x^T g_p(\bar{z}).$$

Da $V_0 = \left[V_1 \vdots v_1\right]$, folgt dann:

$$\gamma = -D_t \mu_p(\bar{t}) \cdot \left(0 \vdots D_x g_p(\bar{z}) v_1\right)^T \begin{pmatrix} V_1^T D_x \mathbb{L}_{J_0} V_1 & V_1^T D_x \mathbb{L}_{J_0} v_1 \\ v_1^T D_x \mathbb{L}_{J_0} V_1 & v_1^T D_x \mathbb{L}_{J_0} v_1 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ v_1^T D_x^T g_p(\bar{z}) \end{pmatrix}.$$

Wegen Lemma 5 (3.) und (65) kommt man zu:

$$\gamma = -D_t \mu_p(\bar{t}) \cdot \left(D_x g_p(\bar{z}) v_1^T\right)^2 \left(D_x \mathbb{L}_{J_0}|_{T_{\bar{x}}^+ M(\bar{t})} | D_x \mathbb{L}_{J_0}|_{T_{\bar{x}} M(\bar{t})}\right)^{-1}.$$

Unmittelbar folgt:

$$sign(\gamma) \cdot sign\left(S(D_x \mathbb{L}_{J_0}|_{T_{\bar{x}}^+ M(\bar{t})} | D_x \mathbb{L}_{J_0}|_{T_{\bar{x}} M(\bar{t})})\right) = -sign(D_t \mu_p(\bar{t})).$$

Die Beziehung (65) liefert dann die Behauptung. □

Aus Behauptung 6 folgt die Möglichkeit, das Verhalten der Tripel

$$(LI, LCI, sign(\det(D_x \mathbb{L}_{J_0}(\bar{z}, \bar{\lambda}, \bar{\mu}, \bar{t})|_{T_{\bar{x}} M(\bar{t})})))$$

darzustellen (Bild 10).

Bemerkung 14

Solche Punkte vom Typ 2, bei denen ein Rückkehrphänomen der Kurve der stationären Punkte auftritt, bleiben im Fall konvexer zulässiger Mengen Lösungen der Variationsungleichung. Für einparametrische Optimierungsprobleme ist das wegen der Existenz einer Sprungrichtung nicht der Fall.

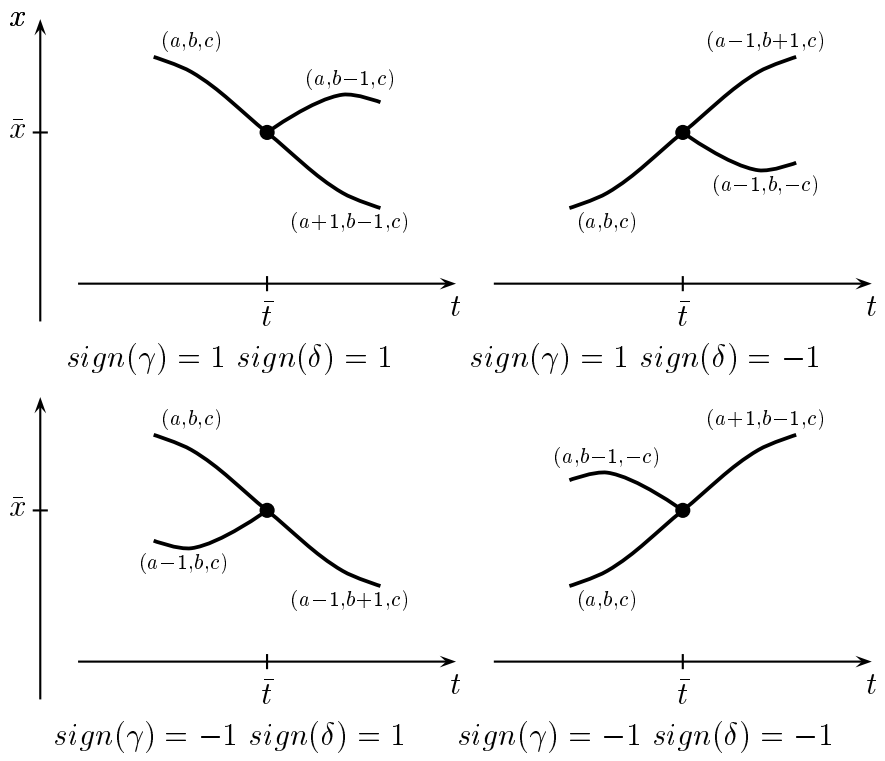


Bild 10: Typ 2 bei Variationsungleichungen

Definition 30 *(vergleiche Definition 16)*
Ein verallgemeinerter kritischer Punkt $\bar{z} = (\bar{x}, \bar{t})$ einer Variationsungleichung $(F, H, G) \in C^3(I\!R^{n+1}, I\!R^{n+m+p})$ ist vom Typ 3, falls die Bedingungen **1.** *und* **2.** *der Definition 16 gelten und folgende Umformulierungen der anderen drei Bedingungen erfüllt sind:*

Typ3: 3-VU *: $D_x \mathbb{L}_{J_0}(\bar{z}, \bar{\lambda}, \bar{\mu})|_{T_{\bar{x}}M(\bar{t})}$ hat den Rang $n - m - |J_0(\bar{z})| - 1$.*

Typ3: 4-VU *: Es seien*

- *$\Pi : I\!R^{n+1} \to I\!R^{m+|J_0(\bar{z})|}$ definiert als $\Pi = (H, G_{J_0(\bar{z})})$,*
- *W eine Matrix, deren Spalten eine lineare Basis des Unterraumes $T_{\bar{x}}M$ bilden,*
- *$w_1 \neq 0$ ein Vektor mit $W^T D_x \mathbb{L}_{J_0}(\bar{x}, \bar{\lambda}, \bar{\mu}, \bar{t}) W \cdot w_1 = 0$ und $v_1 = W \cdot w_1$ und*
- *$w_2 \neq 0$ ein Vektor mit $W^T D_x^T \mathbb{L}_{J_0}(\bar{x}, \bar{\lambda}, \bar{\mu}, \bar{t}) W \cdot w_2 = 0$ und $v_2 = W \cdot w_2$.*

Dann sind die beiden folgenden Zahlen β_1 und β_2 verschieden von Null:

$$\begin{aligned}
\beta_1 &= v_1^T (D_x^2 \mathbb{L}_{J_0} \cdot v_2) v_1 - 2 v_1^T D_x^T \mathbb{L}_{J_0} ((D_x^T \Pi)^\dagger)^T \cdot (v_1^T D_x^2 \Pi \cdot v_2) \\
&\quad - v_2^T D_x \mathbb{L}_{J_0} ((D_x^T \Pi)^\dagger)^T (v_1^T D_x^2 \Pi v_1), \\
\beta_2 &= D_t \mathbb{L}_{J_0} \cdot v_2 - D_t^T \Pi \cdot (D_x^T \Pi)^\dagger D_x^T \mathbb{L}_{J_0} \cdot v_2 .
\end{aligned}$$

Dabei sind folgende Notationen benutzt worden:

$$\begin{aligned}
v_1^T (D_x^2 \mathbb{L}_{J_0} \cdot v_2) v_1 &= v_1^T \left[D_x (D_x^T \mathbb{L}_{J_0} \cdot v_2) \right] v_1^T , \\
v_1^T D_x^2 \Pi v_2 &= \left(v_1^T D_x^2 h_i v_2, i \in I, v_1^T D_x^2 g_j v_2, j \in J_0 \right)^T .
\end{aligned}$$

Damit wird auch $v_1^T D_x^2 \Pi v_1$ erklärt.

Bemerkung 15
Auch in den Variationsungleichungen ist die etwas undurchsichtige Bedingung **Typ3: 4-VU** *unter der LICQ und der Bedingung* **Typ3: 3-VU** *äquivalent zur Nichtdegeneriertheit von $(\bar{x}, \bar{\lambda}, \bar{\mu}, \bar{t})$ als kritischer Punkt des Optimierungsproblems.*

$$\mathcal{P}_{VU} \quad \min_{(z,\lambda,\mu)\in\mathcal{X}_{VU}} \quad \Lambda(z,\lambda,\mu) = t,$$

$$\mathcal{X}_{VU} = \left\{(z,\lambda,\mu) \in I\!R^{n+m+|J_0(\bar{z})|+1} | \mathbb{H}_{J_0(z)}(z,\lambda,\mu) = 0\right\}$$

Bemerkung 16
Die Projektion von $\mathcal{X}_{VU}$ auf die z-Komponente bildet lokal die Menge Σ_{gc}. Das folgt aus der LICQ und der Komplementarität der Multiplikatoren $\bar{\mu}$. Die Menge Σ_{gc} kann lokal um $\bar{z}$ nach einer der x-Komponenten parametrisiert werden, und deswegen ist die Struktur Σ_{gc} wie im Bild 11 dargestellt.

Die Kurve Σ_{gc} bildet um $\bar{z}$ einen Rückkehrpunkt in t, weil $(\bar{x},\bar{\lambda},\bar{\mu},\bar{t})$ ein nichtdegenerierter kritischer Punkt des auf eine eindimensionale Mannigfaltigkeit eingeschränkten Problems $\mathcal{P}_{VU}$ ist. Die Richtung der Rückkehr, entweder für $t \geq \bar{t}$ oder entgegengesetzt, hängt entscheidend vom Vorzeichen der Bedingung zweiter Ordnung **ND2** der Nichtdegeneriertheit von $\mathcal{P}_{VU}$ ab. Es ist leicht zu sehen, daß die Multiplikatoren von $\mathcal{P}_{VU}$ im kritischen Punkt $(\bar{x},\bar{\lambda},\bar{\mu},\bar{t})$

$$\frac{-1}{\beta_2}(v_2, -(D_x^T\Pi)^\dagger D_x^T \mathbb{L}_{J_0} \cdot v_2) \tag{67}$$

sind und daß der Tangentialraum an die Menge $\mathcal{X}_{VU}$ $(T_{(\bar{z},\bar{\lambda},\bar{\mu})}\mathcal{X}_{VU})$ der lineare Teilraum ist, der vom Vektor $(v_1, (D_x^T\Pi)^\dagger D_x \mathbb{L}_{J_0} \cdot v_1)$ erzeugt wird. Bezeichnen wir mit $L(z,\lambda,\mu)$ die Lagrange-Funktion bezüglich $\mathcal{P}_{VU}$ im kritischen Punkt $(\bar{x},\bar{\lambda},\bar{\mu},\bar{t})$. Die Art der Optimalität von $(\bar{x},\bar{\lambda},\bar{\mu},\bar{t})$, und damit die Richtung der Rückkehr in t, hängt vom Vorzeichen der Zahl

$$D^2_{(z,\lambda,\mu)} L(\bar{z},\bar{\lambda},\bar{\mu})|_{T_{(\bar{z},\bar{\lambda},\bar{\mu})}\mathcal{X}_{VU}} = \frac{-1}{\beta_2} \cdot \beta_1$$

ab.

Daher folgt, daß $\bar{z}$ einen Maximalpunkt (bzw. Minimalpunkt) bezüglich t in Σ_{gc} bildet, falls $sign(\beta_1\beta_2) = 1$ (bzw. $sign(\beta_1\beta_2) = -1$). Damit wird das Verhalten im Bild 11 in Abhägigkeit von $sign(\beta_1\beta_2)$ klar.

Die Vektoren v_1 und v_2 in Bedingung **Typ3: 4-VU** sind eindeutig bestimmt, bis auf jeweils eine multiplikative Konstante. Die Konstante zu v_1 multipliziert im Quadrat auch β_1 und hat deswegen keinen Einfluß auf das Vorzeichen von β_1.

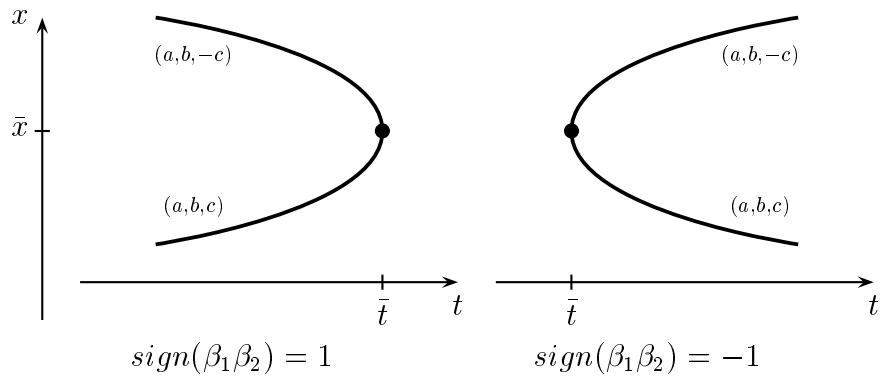


Bild 11: Typ 3 bei Variationsungleichungen

Die Konstante zu v_2 multipliziert $\beta_1$und auch β_2, daher sind die Multiplikatoren in (67) eindeutig. Das heißt, sie multipliziert $\beta_1\beta_2$ quadratisch, und damit ist $sign(\beta_1\beta_2)$ unabhängig von der Selektion von v_1 und v_2.

Wegen der LICQ und der Komplementarität ändern sich die linearen Indizes entlang der kritischen Kurve nicht. Der Beweis über die Vorzeichenänderung von $\det(D_x\mathbb{L}_{J_0}(\bar{z},\bar{\lambda},\bar{\mu})|_{T_{\bar{x}}M(\bar{t})})$ wird im weiter hinten zu findenden Satz 20 und in Bemerkung 31 erklärt.

Bemerkung 17
Auch in Punkten vom Typ 3 der Variationsungleichungen bleiben solche Punkte, die eine Rückkehr der Menge der stationären Punkte hervorrufen, selbst im Fall eines konvexen zulässigen Bereichs nicht nur ein stationärer Punkt, sondern auch eine Lösung des Problems. Das ist anders als für Optimierungsprobleme, wo diese Rückkehrpunkte stationär, aber wegen der Existenz einer Abstiegsrichtung keine Lösung sind.

Definition 31 *(vergleiche Definition 17)*
Ein verallgemeinerter kritischer Punkt $\bar{z} = (\bar{x}, \bar{t})$ einer Variationsungleichung $(F, H, G) \in C^2(\mathbb{R}^{n+1}, \mathbb{R}^{n+m+p})$ ist vom Typ 4, falls die Bedingungen **1.**, **2.**, **3.** *und* **4.** *der Definition 17 gelten und folgende Umformulierung der fünften Bedingung erfüllt ist:*

Typ4: 5-VU *: Es seien*

$$T = \{\eta \in \mathbb{R}^n | Dh_i(\bar{z})\eta = 0, Dg_j(\bar{z})\eta = 0, i \in I, j \in J_0(\bar{z})\}$$

und W eine Matrix, deren Spalten eine lineare Basis von T bilden. Weiter seien

$$\mathbf{L}(z) = \sum_{i \in I} \bar{\lambda}_i h_i(z) + \sum_{j \in J_0(\bar{z})} \bar{\mu}_j g_j(z) \tag{68}$$

und $A = D_t \mathbf{L}(\bar{z}) \cdot W^T D_x^2 \mathbf{L}(\bar{z}) W$. (Nach den Bedingungen **1.** *bis* **4.** *der Definition 17 ist $\mathbf{L}(z)$ definiert und A regulär). Dann ist die Zahl $F^T(\bar{z}) W A^{-1} W^T F(\bar{z})$ verschieden von Null.*

Bemerkung 18

Es sei bemerkt, daß in der Definition 17 vom Typ 4 für Optimierungsprobleme der Gradient der Zielfunktion nur im Ausdruck $D_x f(\bar{z}) W A^{-1} W^T D_x^T f(\bar{z})$ auftaucht. Insbesondere kommt die Hesse-Matrix der Zielfunktion in der Definition nicht vor. Daher spielt es überhaupt keine Rolle, ob der Vektor $D_x^T f(\bar{z})$ ein Gradient ist oder nicht, und damit ist die vorhergehende Definition 31 begründet.

Bei den Punkten vom Typ 4 der Variationsungleichungen sind auch die Regulärität der Matrix $A = D_t \mathbf{L}(\bar{z}) \cdot W^T D_x^2 \mathbf{L}(\bar{z}) W$ und die Bedingung

$$F^T(\bar{z}) W A^{-1} W^T F(\bar{z}) \neq 0$$

äquivalent unter den ersten drei Bedingungen der Definition zu der folgenden Nichtdegeneriertheit (der Beweis ist derselbe wegen der Unabhängigkeit der Definition von der Hesse-Matrix).

Da die Matrix $D_x \begin{pmatrix} H \\ G_{J_0(\bar{z})}(z) \end{pmatrix}$ den Rang $m + J_0(\bar{z}) - 1$ hat, gibt es eine bis auf eine multiplikative Konstante eindeutige, verschwindende Kombination der Zeilen mit Koeffizienten $(\bar{\lambda}_i, \bar{\mu}_j, i \in I, j \in J_0(\bar{z}))$. Wegen der dritten Bedingung der Definition vom Typ 4 sind alle $\bar{\mu}_j \neq 0$. Es sei $p \in J_0(\bar{z})$ ein beliebiger, aber fester Index und $\theta = (\theta_0, \theta_i, \theta_j, i \in I, j \in J_0(\bar{z}) \setminus \{p\})$. Dann ist der Punkt $(\bar{z}, \bar{\theta})$, mit $\bar{\theta} = (0, \bar{\lambda}_i, \bar{\mu}_j, i \in I, j \in J_0(\bar{z}) \setminus p)$, ein nicht degenerierter stationärer Punkt des Problems $\widetilde{\mathcal{P}}_{VU}$, in dem die LICQ erfüllt ist.

$$\begin{array}{lll} \widetilde{\mathcal{P}}_{VU} & \min & \tilde{\Lambda}(z, \theta) = t, \\ & (z, \theta) \in \widetilde{\mathcal{X}}_{VU} & \\ & & \widetilde{\mathcal{X}}_{VU} = \left\{ (z, \theta) \in I\!R^{n+m+|J_0(\bar{z})|+1} | \Gamma_{VU}(z, \theta) = 0 \right\} \end{array} \tag{69}$$

Dabei ist

$$\Gamma_{VU}(z,\theta) = \begin{pmatrix} \theta_0 F(z) - \sum_{i\in I} \theta_i D_x h_i(z) - \sum_{j\in J_0(\bar{z})\setminus\{p\}} \theta_j D_x g_j(z) - \bar{\mu}_p D_x g_p(z) \\ H(z) \\ G_{J_0(\bar{z})}(z) \end{pmatrix}.$$

Bemerkung 19
Aus $W^T F(\bar{z}) \neq 0$ folgt die lineare Unabhängigkeit der Vektoren

$$\{F(\bar{z}), D_x h_i(\bar{z}), D_x g_j(\bar{z}), i \in I, j \in J_0(\bar{z}) \setminus \{k\}\} \tag{70}$$

für beliebiges $k \in J_0(\bar{z})$. Das impliziert, daß zu jedem Punkt $z \in \Sigma_{gc}$ in einer Umgebung von $\bar{z}$ eindeutige Koeffizienten θ existieren, so daß $(z,\theta) \in \widetilde{\mathcal{X}}_{VU}$. Daher bildet die Projektion von $\widetilde{\mathcal{X}}_{VU}$ auf die Komponenten z lokal um $\bar{z}$ die Menge Σ_{gc}. Die Menge $\widetilde{\mathcal{X}}_{VU}$ kann mittels einer glatten Abbildung $(z(\tau),\theta(\tau))$, die auf einer Umgebung von Null definiert ist, parametrisiert werden. Dabei gilt $(z(0),\theta(0)) = (\bar{z},\bar{\theta})$. Es sei $(\dot{z}(0),\dot{\theta}(0))$ die Ableitung dieser Abbildung. Aus der LICQ in $(\bar{z},\bar{\theta})$ ist leicht zu sehen, daß $\dot{t}(0) = 0$ und $\dot{\theta}_0(0) \neq 0$. Aus dieser letzten Ungleichung folgt einerseits, daß in allen Punkten von Σ_{gc} in einer Umgebung von $\bar{z}$ die LICQ erfüllt ist. Andererseits wechseln die linearen Indizes gegenseitig ihre Werte für τ kleiner und größer als Null, denn die zugehörigen Muliplikatoren sind $\frac{1}{\theta_0(\tau)}(\theta_i(\tau), \theta_j(\tau), i \in I, j \in J_0(\bar{z}) \setminus \{p\}, \bar{\mu}_p)$.

Behauptung 7
Es sei im Punkt $\bar{z}$ vom Typ 4 eine Parametrisierung $(z(\tau),\theta(\tau))$ der Kurve $\widetilde{\mathcal{X}}_{VU}$ (und damit $z(\tau)$ der Kurve Σ_{gc}) wie vorher gewählt (mit $(z(0),\theta(0)) = (\bar{z},\bar{\theta})$)). Dann sind alle Punkte $z(\tau)$ mit $\tau \neq 0$ nichtdegeneriert. Das Vorzeichen von $\det(D_x \mathbb{L}_{J_0}(z(\tau))|_{T_{x(\tau)}M(t(\tau))})$ ist auf beiden Zweigen $\tau > 0$ und $\tau < 0$ dasselbe (bzw. verschieden), genau dann, wenn $n - m - J_0(\bar{z})$ gerade (bzw. ungerade) ist.

<u>Beweis zur Behauptung 7:</u>
Es sei $p \in J_0(\bar{z})$ fixiert. Wir benutzen folgende Bezeichnungen:
für die Matrizen

$$B_0(z) = D_x \begin{pmatrix} H \\ G_{J_0(\bar{z})\setminus\{p\}}(z) \\ F^T(z) \end{pmatrix}$$

und

$$B_1(z) = D_x \begin{pmatrix} H(z) \\ G_{J_0(\bar{z})\setminus\{p\}}(z) \end{pmatrix}$$

, für die linearen Teilräume

$$\begin{aligned} T_0(z) &= \{\xi \in I\!R^n | B_0(z)\xi = 0\} \\ T_1(z) &= \{\xi \in I\!R^n | B_1(z)\xi = 0\} \end{aligned}$$

und für die Abbildung

$$\mathcal{L}(z,\theta) = \theta_0 F(z) - \sum_{i\in I} \theta_i D_x h_i(z) - \sum_{j\in J_0(\bar{z})\setminus\{p\}} \theta_j D_x g_j(z) - \bar{\mu}_p D_x g_p(z).$$

Beide Matrizen haben, wie vorher erwähnt, linear unabhängige Zeilen. Für $i = 0, 1$ sei W_i eine $n\times(n-m-J_0(\bar{z})+i)$-Matrix, deren Spalten eine Basis von $T_i(\bar{z})$ bilden. Dabei sei $W_1 = \left[W_0 \vdots b_0\right]$ mit $b_0 = F(\bar{z}) - B_1^T(\bar{z}) \left[B_1^T(\bar{z})\right]^\dagger F(\bar{z})$.

Es existiert eine Umgebung $U_{\bar{z}}$ von $\bar{z}$, wo die Zeilen der Matrix $B_0(z)$ linear unabhängig bleiben und die Spalten der Matrix $\left[B_0^T(z) \vdots W_0\right]$ eine Basis des Raumes $I\!R^n$ bilden. Die $n \times (n - m - J_0(\bar{z}))$-Matrix $W_0(z) = W_0 - B_0^T(z) \left[B_0^T(z)\right]^\dagger W_0$ bildet für $z \in U_{\bar{z}}$ eine Basis von $T_0(z)$. Die Elemente der Matrix $W_0(z)$ hängen wegen des vollen Ranges von $B_0(\bar{z})$ ebenso glatt von z ab wie die Elemente von $B_0(z)$.

Betrachten wir solche τ in einer Umgebung von Null, mit $z(\tau) \in U_{\bar{z}}$. Wir wissen schon, daß die Punkte $z(\tau)$ im Σ_{gc} liegen und daß die LICQ erfüllt ist. Die Multiplikatoren haben die Gestalt $\frac{1}{\theta_0(\tau)}(\theta_i(\tau), \theta_j(\tau), i \in I, j \in J_0(\bar{z}) \setminus \{p\}, \bar{\mu}_p)$, und damit ist die Komplementarität erfüllt.

Daraus folgt, daß die Spalten der Matrix $W_0(z(\tau))$ eine Basis des Tangentialraumes $T_{x(\tau)}M(t(\tau))$ bilden und daß die Nichtdegeneriertheit von $z(\tau)$ von der Regularität der Matrix $W_0^T(z(\tau))D_x\mathbb{L}_{J_0}(z(\tau))W_0(z(\tau))$ abhängt, wobei

$$D_x\mathbb{L}_{J_0}(z(\tau)) = \frac{1}{\theta_0(\tau)} D_x\mathcal{L}(z(\tau), \theta(\tau)). \tag{71}$$

ist.

Falls die Matrix

$$W_0^T(z(0))D_x\mathcal{L}(z(0), \theta(0))W_0(z(0)) = W_0^T D_x^2 \mathbf{L}(\bar{z}) W_0 \tag{72}$$

regulär ist, wobei $\mathbf{L}(z)$ entsprechend Definition 31 verstanden wird, folgt aus der Beziehung (71) zunächst die Nichtdegeneriertheit von $z(\tau)$ für $\tau \neq 0$. Die Veränderung von $sign(\det(W_0^T(z(\tau))D_x\mathbb{L}_{J_0}(z(\tau))W_0(z(\tau))))$ ergibt sich aus der folgenden Gleichung.

$$\det\left(W_0^T(z(\tau))D_x\mathbb{L}_{J_0}(z(\tau))W_0(z(\tau))\right)=\left(\frac{1}{\theta_0(\tau)}\right)^{|n-m-J_0(\bar{z})|}\det\left(W_0^T(z(\tau))D_x\mathcal{L}(z(\tau),\theta(\tau))W_0(z(\tau))\right)$$

Es reicht dann, die Regularität der Matrix (72) zu zeigen. Laut Lemma 1 ($D_x^2\mathbf{L}(\bar{z})$ symmetrisch !) ist die Regularität dieser Matrix äquivalent zur Regularität der folgende Matrix.

$$\begin{pmatrix} D_x^2\mathbf{L}(\bar{z}) & B_1(\bar{z}) & F(\bar{z}) \\ B_1^T(\bar{z}) & 0 & 0 \\ F^T(\bar{z}) & 0 & 0 \end{pmatrix} \tag{73}$$

Die Regularität von $\begin{pmatrix} D_x^2\mathbf{L}(\bar{z}) & B_1(\bar{z}) \\ B_1^T(\bar{z}) & 0 \end{pmatrix}$ folgt unter Benutzung von Lemma 1 aus der Regularität von A, die in der Definition 31 vorausgesetzt wird. Laut Lemma 5 ist die Regularität von (73) äquivalent zur Tatsache, daß die Zahl

$$\begin{pmatrix} F^T(\bar{z}) & 0 \end{pmatrix} \begin{pmatrix} D_x^2\mathbf{L}(\bar{z}) & B_1(\bar{z}) \\ B_1^T(\bar{z}) & 0 \end{pmatrix}^{-1} \begin{pmatrix} F(\bar{z}) \\ 0 \end{pmatrix}$$

gleich zu

$$F^T(\bar{z})W_1\left[W_1^T D_x^2\mathbf{L}(\bar{z})W_1\right]^{-1} W_1^T F(\bar{z})$$

und verschieden von Null ist. Das ist aber genau das, was in der Bedingung **Typ4: 5-VU** vorausgesetzt wird. □

Bemerkung 20
Da die Vektoren (70) linear unabhängig sind und $\dot{t}(0) = 0$, ist der Vektor $\dot{x}(0)$ auch verschieden von Null. Daher ist die Menge Σ_{gc} um $\bar{z}$ nach einer der x-Komponenten parametrisierbar, und die Struktur von Σ_{gc} um $\bar{z}$ ergibt das Bild 12.

Um $\bar{z}$ bildet Σ_{gc} einen Rückkehrpunkt bezüglich t, weil $(\bar{z}, \bar{\theta})$ ein nichtdegenerierter kritischer Punkt des Problemes $\widetilde{\mathcal{P}}_{VU}$ ist. Die zu $(\bar{z}, \bar{\theta})$ gehörenden Multiplikatoren sind

$$\frac{-1}{D_t\mathbf{L}(\bar{z})}(0, (\bar{\lambda}, \bar{\mu})),$$

wobei $(\bar{\lambda}, \bar{\mu})$ wie in Bemerkung 18 gewählt sind und $\mathbf{L}(z)$ wie in (68). Da $(\bar{\lambda}, \bar{\mu})$ als Koeffizienten in $D_t\mathbf{L}(\bar{z})$ auftauchen, sind die Multiplikatoren unabhängig von der freien multiplikativen Konstante bei der Wahl von $(\bar{\lambda}, \bar{\mu})$.

Es sei mit L die Lagrangefunktion bezüglich $\widetilde{\mathcal{P}}_{VU}$ im kritischen Punkt $(\bar{z}, \bar{\theta})$ bezeichnet. Die für die Richtung der Rückkehr entscheidende Zahl $D^2_{(z,\theta)}L(\bar{z}, \bar{\theta})|_{T_{(\bar{z},\bar{\theta})}\widetilde{\mathcal{X}}_{VU}}$ hat in diesem Fall die Gestalt

$$D^2_{(z,\theta)}L(\bar{z}, \bar{\theta})|_{T_{(\bar{z},\bar{\theta})}\widetilde{\mathcal{X}}_{VU}} = \frac{-1}{D_t\mathbf{L}(\bar{z})} \cdot F^T(\bar{z})W\left[W^T D^2_x\mathbf{L}(\bar{z})W\right]^{-1} W^T F(\bar{z}),$$

dabei wird W wie in Definition 31 benutzt. Speziell folgt daraus die Beziehung

$$sign(D^2_{(z,\theta)}L(\bar{z}, \bar{\theta})|_{T_{(\bar{z},\bar{\theta})}\widetilde{\mathcal{X}}_{VU}}) = -sign(F^T(\bar{z})WA^{-1}W^T F(\bar{z})).$$

Es kann jetzt die Zahl

$$\alpha = sign(F^T(\bar{z})WA^{-1}W^T F(\bar{z})),$$

die auch unabhängig von der eventuellen Selektion von $(\bar{\lambda}, \bar{\mu})$ ist, für die geometrische Darstellung der Kurve Σ_{gc} genutzt werden. Davon wird, wie im Bild 12 zu sehen ist, für $\alpha = 1$ ($\alpha = -1$) ein lokaler Maximalpunkt (bzw. Minimalpunkt) in der t-Komponente gebildet.

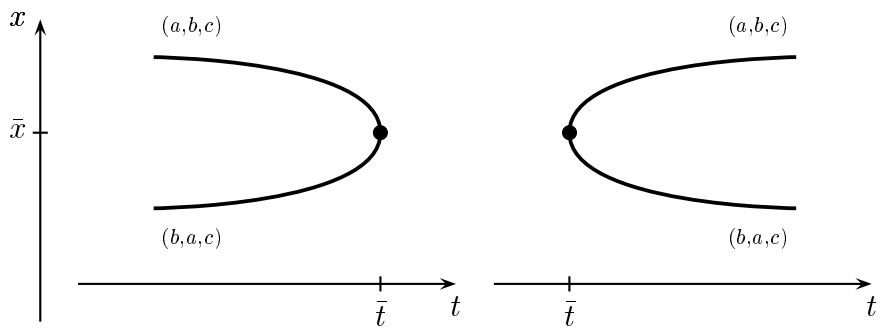
Laut Behauptung 7 entscheidet das Vorzeichen der Zahl

$$\beta = (-1)^{n-m-J_0(\bar{z})}$$

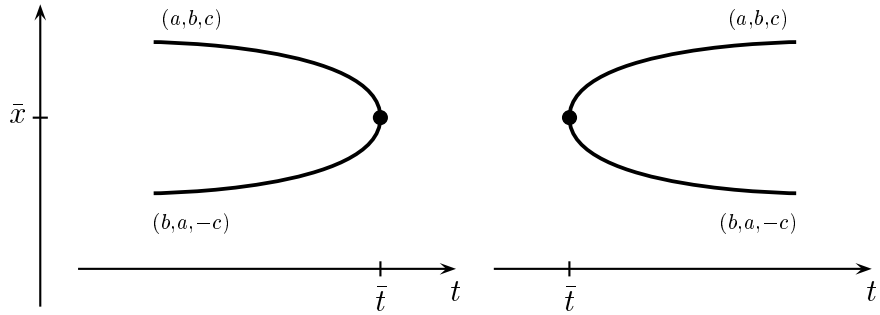
über das Verhalten von $sign(\det(D_x\mathbb{L}_{J_0}(z(\tau))|_{T_{x(\tau)}M(t(\tau))}))$. Damit ensteht das Bild 12 der lokalen Struktur.

Bemerkung 21
Im Fall von konvexen zulässigen Mengen (h_i affin und deswegen $J_0(\bar{z}) \neq \emptyset$) gibt es ein wichtiges Merkmal bei Punkten vom Typ 4. Es sei $(z(\tau), \theta(\tau))$ eine Parametrisierung wie in Behauptung 7, und wir nehmen an, daß für $\tau < 0$ die Punkte $z(\tau)$ stationär sind. Daraus folgt, daß alle $\bar{\mu}_j$, $j \in J_0(\bar{z})$, in Definition 31 das gleiche Vorzeichen haben (MFCQ ist verletzt).

Für $\tau < 0$ gilt die LICQ, und wegen der Konvexitätsvoraussetzung sind alle Matrizen $D^2_x g_j(z(\tau))$, $j \in J_0(\bar{z})$ in dem Tangentialraum negativ semidefinit. Daraus folgt, daß die Matrix $W^T D^2_x\mathbf{L}(\bar{z})W$ mindestens $n-m-J_0(\bar{z})$ negative



$sign(\alpha) = 1 \; sign(\beta) = 1$ $\qquad$ $sign(\alpha) = -1 \; sign(\beta) = 1$



$sign(\alpha) = 1 \; sign(\beta) = -1$ $\qquad$ $sign(\gamma) = -1 \; sign(\beta) = -1$

$$a + b = |J_0(\bar{z})| \qquad c = +(-1)$$

Bild 12: Typ 4 bei Variationsungleichungen

Eigenwerte besitzt, wobei W und $\mathbf{L}$, speziell mit $\bar{\mu}_j > 0$, $\forall j \in J_0(\bar{z})$, wie in Definition 31 erklärt sind. Daher ist ihre Anzahl positiver Eigenwerte 0 *oder* 1.

Die Konvexitätsvoraussetzung impliziert dann nach den Ergebnissen des Satzes 5.3 (Fall 2. und Fall 3.) in [33] über die homotopischen Veränderungen der zulässsigen Menge, daß die Matrix $W^T D_x^2 \mathbf{L}(\bar{z}) W$ negativ definit ist und daß die zulässige Menge im Parameterwert $\bar{t}$ nicht leer wird oder verschwindet.

Definition 32 *(vergleiche Definition 18)*
Ein verallgemeinerter kritischer Punkt $\bar{z} = (\bar{x}, \bar{t})$ einer Variationsungleichung $(F, H, G) \in C^2(\mathbb{R}^{n+1}, \mathbb{R}^{n+m+p})$ ist vom Typ 5, falls die Bedingungen **1.**, **2.** *und* **3.** *der Definition 18 gelten und folgende Umformulierung der vierten Bedingung erfüllt ist:*

Typ5: 4-VU *: Es seien $(\bar{\lambda}, \bar{\mu}) = (\bar{\lambda}_i, \bar{\mu}_j, i \in I, j \in J_0(\bar{z}))$ die bis auf eine multiplikative Konstante eindeutigen Koeffizienten des Gleichungssystems*

$$\sum_{i \in I} \bar{\lambda}_i D_x h_i(\bar{z}) + \sum_{j \in J_0(\bar{z})} D_x \bar{\mu}_j g_j(\bar{z}) = 0$$

und $(\bar{\alpha}, \bar{\beta}) = (\bar{\alpha}_i, \bar{\beta}_j, i \in I,\ j \in J_0(\bar{z}))$ die eindeutige Lösung von

$$\begin{pmatrix} F(\bar{z}) \\ 0 \end{pmatrix} - \sum_{i \in I} \bar{\alpha}_i D^T h_i(\bar{z}) - \sum_{j \in J_0(\bar{z})} \bar{\beta}_j D^T g_j(\bar{z}) = 0.$$

Dann sind alle Zahlen $\Delta_{jk} = \bar{\beta}_j - \bar{\beta}_k \cdot \frac{\bar{\mu}_j}{\bar{\mu}_k}$, $\forall j, k \in J_0(\bar{z})$, $j \neq k$, verschieden von Null.

Bemerkung 22
In der Definition 18 der Punkte vom Typ 5 für einparametrische Optimierungsprobleme taucht nicht die Hesse-Matrix der Zielfunktion auf. Damit spielt es keine Rolle, ob der Vektor $Df(\bar{z})$ tatsächlich ein Gradient ist. Es ist dann die vorhergehende Definition für Variationsungleichungen durch den Ersatz von $Df(\bar{z})$ mit $(F^T(\bar{z}), 0)$ begründet.

Sind die Bedingungen **1.**, **2.** *und* **3.** *der Definition 18 erfüllt, dann ist* **Typ5: 4-VU** *äquivalent zur Bedingung, daß jede Lösung $(\theta_i, \vartheta_j, i \in I, j \in$*

$J_0(\bar{z})$) des linearen Gleichungssystems

$$F(\bar{z}) - \sum_{i \in I} \theta_i D_x h_i(\bar{z}) - \sum_{j \in J_0(\bar{z})} \vartheta_j D_x g_j(\bar{z}) = 0 \tag{74}$$

höchstens einen Multiplikator $\vartheta_j, j \in J_0(\bar{z})$ hat, der gleich Null ist.

Bemerkung 23
Laut Bedingungen **1.** *und* **2.** *der Definition 18 sind die $n+1$ Vektoren*

$$\{Dh_i(\bar{z}), Dg_j(\bar{z}), i \in I, j \in J_0(\bar{z})\} \tag{75}$$

linear unabhängig. Damit ist $\bar{z}$ in einer Umgebung der einzige zulässige Punkt mit der aktiven Indexmenge $J_0(\bar{z})$. Aus Bedingung **3.** *in der genannten Definition (alle $\bar{\mu}_j, j \in J_0(\bar{z})$, von* **Typ5: 4-VU** *sind verschieden von Null) folgt die lineare Unabhängigkeit jeder Vektormenge der Form*

$$\{D_x h_i(\bar{z}), D_x g_j(\bar{z}), i \in I, j \in J_0(\bar{z}) \setminus \{k\}\}$$

mit $k \in J_0(\bar{z})$. Das heißt, daß jeder verallgemeinerte kritische Punkt in einer Umgebung von $\bar{z}$ zu einer Lösung eines der Gleichungssysteme

$$\mathbb{H}_{J_0(\bar{z})\setminus\{k\}}(z, \lambda, \mu) = \begin{pmatrix} \mathbb{L}_{J_0(\bar{z})\setminus\{k\}}(z, \lambda, \mu) \\ H(z) \\ G_{J_0(\bar{z})\setminus\{k\}}(z) \end{pmatrix} = 0 \tag{76}$$

mit $k \in J_0(\bar{z})$ gehört.

Nach Bedingung **Typ5: 4-VU** gibt es für jede der $n+1$ Indexmengen $I \cup J_0(\bar{z}) \setminus \{k\}$, $k \in J_0(\bar{z})$, eine eindeutige Lösung $(\bar{\lambda}^k, \bar{\mu}^k)$ des Systems

$$\mathbb{H}_{J_0(\bar{z})\setminus\{k\}}(\bar{z}, \lambda, \mu) = 0,$$

die auch die strikte Komplementarität erfüllt, d.h. $\bar{\mu}_j^k \neq 0$, $\forall j \in J_0(\bar{z}) \setminus \{k\}$ (es gilt $\bar{\mu}_j^k = \Delta_{jk}$).

Da $|I \cup J_0(\bar{z}) \setminus \{k\}| = n$, ist die Matrix $D_{(x,\lambda,\mu)}\mathbb{H}_{J_0(\bar{z})\setminus\{k\}}(\bar{z}, \bar{\lambda}^k, \bar{\mu}^k)$ regulär. Daraus folgt für jedes $k \in J_0(\bar{z})$ die Existenz einer Parametrisierung nach t um $(\bar{z}, \bar{\lambda}^k, \bar{\mu}^k)$ der Lösungsmenge von (76) $(x^k(t), \lambda^k(t), \mu^k(t))$, definiert in einer Umgebung von $\bar{t}$ mit $(x^k(\bar{t}), \lambda^k(\bar{t}), \mu^k(\bar{t})) = (\bar{z}, \bar{\lambda}^k, \bar{\mu}^k)$.

Bemerkung 24
Die Menge Σ_{gc} wird dann um $\bar{z}$ durch die zulässigen Stücke ($g_k \geq 0$) der $n+1$ Kurven $(x^k(t), t)$ gebildet. Aus der Unabhängigkeit der $n+1$ Vektoren (75) folgt für jedes $k \in J_0(\bar{z})$ die folgende Ungleichung.

$$\frac{d}{dt} g_k(x^k(\bar{t}), \bar{t}) = D_x^T g_k(\bar{z}) \dot{x}^k(\bar{t}) + D_t g_k(\bar{z}) \neq 0 \tag{77}$$

Aus dieser Beziehung folgt auch, daß $\dot{x}^{k_1}(\bar{t}) \neq \dot{x}^{k_2}(\bar{t})$ für $k_1 \neq k_2$ sind, denn

$$D_x^T g_{k_1}(\bar{z}) \dot{x}^{k_2}(\bar{t}) + D_t g_{k_1}(\bar{z}) = 0.$$

Das ist auch in der Darstellung der Menge Σ_{gc} in einer Umgebung der Punkte vom Typ 5 der Bilder 5 und 9 zu sehen.

Die Zahl (77) entscheidet, in welchem der beiden Teile $t \geq \bar{t}$ oder $t \leq \bar{t}$ die Kurve $(x^k(t), t)$ zulässig ist. Es seien $(\bar{\lambda}, \bar{\mu})$ die Koeffizienten aus der Definition 32 und $\mathbf{L}(z) = \sum_{i \in I} \bar{\lambda}_i h_i(z) + \sum_{j \in J_0(\bar{z})} \bar{\mu}_j g_j(z)$, dann ist folgender andere Ausdruck für (77) leicht zu finden:

$$\frac{d}{dt} g_k(x^k(\bar{t}), \bar{t}) = \frac{1}{\bar{\mu}_k} D_t \mathbf{L}(\bar{z}).$$

Daher ist für $\gamma_k = sign(\bar{\mu}_k \cdot D_t \mathbf{L}(\bar{z})) > 0$ (bzw. < 0) die Kurve $(x^k(t), t)$ für $t \geq \bar{t}$ (bzw. $t \leq \bar{t}$) enthalten in der Menge Σ_{gc}. Die linearen Indizes entlang der Kurve $(x^k(t), t)$ hängen von der Anzahl positiver oder negativer Komponenten des Vektors $(\Delta_{jk}, j \in J_0(\bar{z}) \setminus \{k\})$ ab.

Bemerkung 25
Falls die MFCQ in $\bar{z}$ verletzt ist, haben alle $\bar{\mu}_k$ das gleiche Vorzeichen. Damit laufen alle Kurven von Σ_{gc} in die gleiche Richtung. Das heißt, es entsteht ein Rückkehrpunkt in der Menge Σ_{gc}. In diesem Fall, der Beweis ist derselbe wie für Optimierungsprobleme, ist die Menge der stationären Punkte eine einzige Kurve $(x^k(t), t)$. Damit ist $\bar{z}$ ein Randpunkt.

Falls die MFCQ erfüllt ist und $\bar{z}$ ein stationärer Punkt ist, gibt es genau zwei Kurven $(x^{k_1}(t), t)$ und $(x^{k_2}(t), t)$ mit entgegengesetzten Richtungen, die aus stationären Punkten bestehen. Das folgt aus denselben Argumenten wie für den Fall von Optimierungsproblemen. Es heißt, daß unter der MFCQ eine Fortsetzung in der Menge der stationären Punkte existiert.

Die lokale Struktur der Mengen Σ_{gc} und Σ_{stat} in einer Umgebung der Punkte vom Typ 5 ist identisch mit der Struktur im Falle von Optimierungsproblemen. Deswegen verweisen wir auf die Bilder 5 und 9.

Bemerkung 26
Im Fall von konvexen zulässigen Mengen (h_i affin und $J_0(\bar{z}) \neq \emptyset$, denn $|I| <= n$) folgt aus den Sätzen 4.1 (Typ 3 und 4) und 5.3 (Fall 2. und Fall 3.) in [33], daß die zulässige Menge im Parameterwert $\bar{t}$ erst auftaucht oder verschwindet, wenn die MFCQ verletzt ist.

5.4 Generizität reguläre Variationsungleichungen

Wir sind jetzt in der Lage, die Klasse der regulären einparametrischen Variationsungleichungen einzuführen.

Definition 33
Gegeben sei eine einparametrische Variationsungleichung (F, H, G) und für $i = 1, \ldots, 5$ seien die folgenden Mengen definiert:

$$\Sigma^i_{gc}(F, H, G) = \{z \in \Sigma_{gc}(F, H, G) | z \text{ ist vom Typ } i\}\,.$$

Falls eine einparametrische Variationsungleichung klar definiert ist, schreiben wir kurz Σ_{gc} und Σ^i_{gc}.

Es sei

$$\mathcal{V} = \left\{(F, H, G) \in C^3(I\!R^{n+1}, I\!R^{n+m+p}) | \Sigma_{gc} = \cup_{i=1}^5 \Sigma^i_{gc}\right\}.$$

Definition 34
Eine einparametrische Variationsungleichung $(F, H, G) \in C^3(I\!R^{n+1}, I\!R^{n+m+p})$, die in der Menge $\mathcal{V}$ enthalten ist, nennen wir auch regulär im Sinne von Jongen, Jonker und Twilt, oder einfach regulär.

Als wesentliche Eigenschaften der Menge $\mathcal{V}$ regulärer Variationsungleichungen konzentrieren wir uns auf ihre Offenheit und Dichtheit. Als ersten Schritt beweisen wir den entscheidenden Fakt für die lokale Offenheit.

Lemma 6
Es seien $z_k \in I\!R^{n+1}$,$(F^k, H^k, G^k) \in C^3(I\!R^{n+1}, I\!R^{n+m+p})$ Folgen mit den Konvergenzbedingungen

1.- $z_k \to \bar{z}$ *und*

2.- $\forall |\alpha| \leq 3, D_\alpha(F^k, H^k, G^k)(z_k) \to D_\alpha(\bar{F}, \bar{H}, \bar{G})(\bar{z})$.

Es sei $\bar{z} \in \Sigma^i_{gc}(\bar{F}, \bar{H}, \bar{G})$ *mit* $i \in \{1, \ldots, 5\}$ *(bzw.* $\bar{z} \notin \Sigma_{gc}(\bar{F}, \bar{H}, \bar{G})$*). Dann gilt, für hinreichend große k,* $\{z_k\} \cap \Sigma_{gc}(F^k, H^k, G^k) \in \Sigma^i_{gc}(F^k, H^k, G^k) \cup \Sigma^1_{gc}(F^k, H^k, G^k)$ *(bzw* $z_k \notin \Sigma_{gc}(F^k, H^k, G^k)$*).*

Beweis:

Falls $\bar{z} \notin \Sigma_{gc}(\bar{F}, \bar{H}, \bar{G})$, hat laut Definition die Matrix

$$\left[\bar{F}(\bar{z}) \vdots D_x^T \bar{H}(\bar{z}) \vdots D_x^T \bar{G}_{J_0(\bar{z})}(\bar{z})\right]$$

linear unabhängige Spalten. Dann gibt es $(m + |J_0(\bar{z})| + 1)$ Zeilen, die eine quadratische Matrix mit einer von Null verschiedenen Determinante bilden. Da die Determinante (bestehend aus Summen und Produkten) stetig von den Matrixkomponenten abhängt, folgt aus Bedingung **2.-**, daß für hinreichend große k

$$\left[F^k(z_k) \vdots D_x^T H^k(z_k) \vdots D_x^T G^k_{J_0(\bar{z})}(z_k)\right]$$

auch regulär ist. Daraus folgt $z_k \notin \Sigma_{gc}(F^k, H^k, G^k)$, denn für hinreichend große k ist $J_0(z_k) \subset J_0(\bar{z})$.

Typ 1: $(\bar{z} \in \Sigma^1_{gc}(\bar{F}, \bar{H}, \bar{G}))$

Mit den selben auf die Matrix

$$\left[D_x^T \bar{H}(\bar{z}) \vdots D_x^T \bar{G}_{J_0(\bar{z})}(\bar{z})\right] \tag{78}$$

angewandten Argumenten wie vorherer kennt man das Erfülltsein der LICQ für hinreichend große k. Es bleibt **VU-ND1** und **VU-ND2** zu zeigen.

Da für hinreichend große k $J_0(z_k) \subset J_0(\bar{z})$ gilt, haben die zu $z_k \in \Sigma_{gc}(F^k, H^k, G^k)$ gehörenden Multiplikatoren $(\lambda^k, \mu^k) \in I\!R^{m+|J_0(\bar{z})|}$ (mit $\mu^k_j = 0$ für $j \in J_0(\bar{z}) \setminus J_0(z_k)$) die folgende Gestalt.

$$(\lambda^k, \mu^k) = \left[D_x^T H^k(z_k) \vdots D_x^T G^k_{J_0(\bar{z})}(z_k)\right]^\dagger F^k(z_k) \tag{79}$$

Damit konvergiert (λ^k, μ^k) gegen $(\bar{\lambda}, \bar{\mu})$, und aus **VU-ND1** für $\bar{z}$ folgt für hinreichend große k $J_0(z_k) = J_0(\bar{z})$ und **VU-ND1** für z_k. Daraus ergibt sich auch, siehe **2.-,** die Konvergenz von

$$D_x \mathbb{L}^k_{J_0(z_k)}(z_k) = D_x F^k(z_k) - \sum_{i \in I} \lambda_i^k D_x^2 h_i^k(z_k) - \sum_{j \in J_0(z_n)} \mu_j^k D_x^2 g_j^k(z_k) \tag{80}$$

gegen $D_x \overline{\mathbb{L}}_{J_0(\bar{z})}(\bar{z}) = D_x \bar{F}(\bar{z}) - \sum_{i \in I} \bar{\lambda}_i D_x^2 \bar{h}_i(\bar{z}) - \sum_{j \in J_0(\bar{z})} \bar{\mu}_j D_x^2 \bar{g}_j(\bar{z})$.

Es seien die Spalten von $\bar{W}$ eine Basis für $T_{\bar{x}} \bar{M}(\bar{t})$. Für hinreichend große k ist die Matrix $\left[D_x^T H^k(z_k) \vdots D_x^T G^k_{J_0(\bar{z})}(z_k) \vdots \bar{W} \right]$ regulär und die Spalten von

$$W_n = \bar{W} - \left[D_x^T H^k(z_k) \vdots D_x^T G^k_{J_0(\bar{z})}(z_k) \right] \left[D_x^T H^k(z_k) \vdots D_x^T G^k_{J_0(\bar{z})}(z_k) \right]^{\dagger} \bar{W} \tag{81}$$

bilden eine Basis für $T_{x_k} M(t_k)$. Es folgt dann aus **2.-** die Konvergenz von $W_k^T D_x \mathbb{L}^k_{J_0(z_k)}(z_k) W_k$ gegen $\bar{W}^T D_x \overline{\mathbb{L}}_{J_0(\bar{z})}(\bar{z}) \bar{W}$. Aus $\det(\bar{W}^T D_x \overline{\mathbb{L}}_{J_0(\bar{z})}(\bar{z}) \bar{W}) \neq 0$ ergibt sich **VU-ND2** für z_k. Damit gilt $z_k \in \Sigma^1_{gc}(F^k, H^k, G^k)$.

<u>Typ 2:</u> $(\bar{z} \in \Sigma^2_{gc}(\bar{F}, \bar{H}, \bar{G}))$

Wie wir für Punkte vom Typ 1 gesehen haben, ist die LICQ für hinreichend große k erfüllt. Es seien $(\bar{\lambda}, \bar{\mu})$ die zu $\bar{z}$ gehörenden Multiplikatoren. Ohne Beschränkung der Allgemeinheit (O.B.d.A.) sei $\bar{\mu}_p = 0, p \in J_0(\bar{z})$. Aus der Darstellung (79) der zu $z_k \in \Sigma_{gc}(F^k, H^k, G^k)$ gehörenden Multiplikatoren und mit den selben Argumenten wie für die Punkte vom Typ 1 lassen sich für hinreichend große k drei Fälle für z_k unterscheiden.

1. $J_0(z_k) = J_0(\bar{z}) \setminus \{p\}$. Dabei muß auch **VU-ND1** bzgl. z_k gelten. Die Bedingung **Typ2: 4-VU** ergibt mit den selben Argumenten wie für den Fall $\bar{z} \notin \Sigma^1_{gc}(\bar{F}, \bar{H}, \bar{G})$ die Bedingung **VU-ND2** bzgl. z_k. Dabei ist wichtig zu bemerken, daß die Spalten der gegen $\bar{W}_0$ konvergierenden Matrizen

$$W_0^k = \bar{W}_0 - B_1^k \left[B_1^k \right]^{\dagger} \bar{W}_0 \tag{82}$$

für hinreichend große k eine Basis vom $T_{x_k} M(t_k)$ bilden, wobei

$$B_1^k = \left[D_x^T H^k(z_k) \vdots D_x^T G^k_{J_0(\bar{z}) \setminus \{p\}}(z_k) \right] \tag{83}$$

und die Spalten von $\bar{W}_0$ eine Basis von $T_{\bar{x}}^{+} M(\bar{t})$ sind. Damit folgt: $z_k \in \Sigma^1_{gc}(F^k, H^k, G^k)$.

2. $J_0(z_k) = J_0(\bar{z})$ und **VU-ND1** bzgl. z_k ist erfüllt. Unter Benutzung der Matrizen (81) und derselben Argumente wie im Fall $\bar{z} \notin \Sigma^1_{gc}(\bar{F}, \bar{H}, \bar{G})$ ergibt die Bedingung **Typ2: 3-VU** das Erfülltsein von **VU-ND2** bzgl. z_k, und damit $z_k \in \Sigma^1_{gc}(F^k, H^k, G^k)$.

3. $J_0(z_k) = J_0(\bar{z})$, aber **VU-ND1** ist verletzt, denn $\mu^k_p = 0$ und $\mu^k_j \neq 0$, $\forall j \in J_0(z_k) \setminus \{p\}$. Die LICQ und die Bedingung **2.** der Punkte vom Typ 2 sind dann erfüllt. Bedingung **Typ2: 3-VU** folgt wie im vorigen Fall unter Benutzung der Matrizen (81). Die Bedingung **Typ2: 4-VU** folgt mit den selben Argumenten, aber unter Benutzung der Matrizen W^k_0, gegeben in (82), denn ihre Spalten sind Basen der Teilräume $T^+_{x_k} M(t_k)$. Aus der Konvergenz $(\lambda^k, \mu^k) \to (\bar{\lambda}, \bar{\mu})$ und **2.-** folgt, daß $D_z \mathbb{L}^k_{J_0(z_k)}(z_k) \to D_z \overline{\mathbb{L}}_{J_0(\bar{z})}(\bar{z})$. Da sich die Zahlen α_k, β_k und γ_k der Bedingung **Typ2: 5-VU** mit elementaren Operationen in Abhängigkeit von den Elementen der Matrizen

$$D_z \mathbb{L}^k_{J_0(z_k)}(z_k), \quad W^k_0, \left[W^{kT}_0 D_x \mathbb{L}^n_{J_0(z_k)}(z_k) W^k_0\right]^{-1}$$

und

$$\left[D^T H^k(z_k) \vdots D^T G^k_{J_0(z_k)}(z_k)\right]$$

schreiben lassen, konvergieren sie zu den zugehörigen Zahlen von $\bar{z}$. Damit ist für hinreichend große k die Bedingung **Typ2: 5-VU** bzgl. z_k erfüllt, und dann $z_k \in \Sigma^2_{gc}(F^k, H^k, G^k)$.

<u>Typ 3:</u> $(\bar{z} \in \Sigma^3_{gc}(\bar{F}, \bar{H}, \bar{G}))$

Analog dem Gedankengang des Falles $\bar{z} \notin \Sigma^1_{gc}(\bar{F}, \bar{H}, \bar{G})$ sind für hinreichend große k die LICQ, die **VU-ND1**-Bedingung und $J_0(\bar{z}) = J_0(z_k)$ erfüllt. Aus **2.-** und der entsprechenden Konvergenz der Multiplikatoren $(\lambda^k, \mu^k) \to (\bar{\lambda}, \bar{\mu})$ folgt, daß $D_z \mathbb{L}^k_{J_0(z_k)}(z_k) \to D_z \overline{\mathbb{L}}_{J_0(\bar{z})}(\bar{z})$. Andererseits sind die Matrizen wie in (81), welche Basen für $T_{x_k} M(t_k)$ bilden, konvergent gegen die Basis $\bar{W}$ des Tangentialraumes $T_{\bar{x}} M(\bar{t})$. Da die Matrizen $W^T_k D_x \mathbb{L}^k_{J_0(z_k)}(z_k) W_k$ gegen $\bar{W}^T D_x \overline{\mathbb{L}}_{J_0(\bar{z})}(\bar{z}) \bar{W}$ konvergieren, haben sie wegen **Typ3: 3-VU**, mit ähnlichen Argumenten wie für den Erhalt der LICQ, für hinreichend große k einen Rang nicht kleiner als $n - m - |J_0(\bar{z})| - 1$. Damit sind nur zwei Fälle möglich:

1. Die Matrix $W_k^T D_x \mathbb{L}^k_{J_0(z_k)}(z_k) W_k$ hat vollen Rang, und damit gilt $z_k \in \Sigma^1_{gc}(F^k, H^k, G^k)$.

2. Die Matrix $W_k^T D_x \mathbb{L}^k_{J_0(z_k)}(z_k) W_k$ hat den Rang $n - m - |J_0(\bar{z})| - 1$. Das bedeutet das Erfülltsein von **Typ3: 3-VU** bzgl. z_k. Um die Bedingung **Typ3: 4-VU** zu zeigen, müssen wir nachweisen, daß die zu z_k zugehörigen β_1^k und β_2^k verschieden von Null sind. Sie können aber so gewählt werden, daß $\beta_1^k \to \bar{\beta}_1$ und $\beta_2^k \to \bar{\beta}_2$. Das ist daraus ersichtlich, daß sich die beiden Zahlen mit elementaren Ausdrücken in Abhängigkeit der Komponenten von

$$D_z \mathbb{L}^k_{J_0(z_k)}(z_k), D_x^2 \mathbb{L}^k_{J_0(z_k)}(z_k),$$

$$\left[D^T H^k(z_k) \vdots D^T G^k_{J_0(\bar{z})}(z_k) \right],$$

v_1^k und v_2^k schreiben lassen. Es bleibt nur zu zeigen, daß $v_1^k \to \bar{v}_1$ und $v_2^k \to \bar{v}_1$. Beschränken wir uns auf v_1.

Da $v_1^k = W_k w_1^k$, $\bar{v}_1 = W_k \bar{w}_1$, reicht es, w_1^k und $\bar{w}_1$ so zu wählen, daß $w_1^k \to \bar{w}_1$. Ist $\bar{w}_1$ festgelegt und o.B.d.A. mit der ersten Komponente $[\bar{w}_1] = 1$, dann lassen sich die restlichen $n-m-|J_0(\bar{z})|-1$ Komponenten mit einer stetigen Funktion der Elemente der Matrix $\bar{W}^T D_x \overline{\mathbb{L}}_{J_0(\bar{z})}(\bar{z}) \bar{W}$ schreiben (nach **Typ3: 3-VU** sind die letzten $n - m - |J_0(\bar{z})| - 1$ Spalten von $\bar{W}^T D_x \overline{\mathbb{L}}_{J_0(\bar{z})}(\bar{z}) \bar{W}$ linear unabhängig). Die gleiche Funktion evaluiert in $W_k^T D_x \mathbb{L}^k_{J_0(z_k)}(z_k) W_k$, liefert dann die letzten $n - m - |J_0(\bar{z})| - 1$ Komponenten von w_1^k, und mit $\left[w_1^k\right] = 1$ ist die Konvergenz $w_1^k \to \bar{w}_1$ gezeigt. Damit folgt für hinreichend große k aus **Typ3: 4-VU** bzgl. $\bar{z}$ die gleiche Bedingung bzgl. z_k, was $z_k \in \Sigma^3_{gc}(F^k, H^k, G^k)$ heißt.

<u>Typ 4:</u> $(\bar{z} \in \Sigma^4_{gc}(\bar{F}, \bar{H}, \bar{G}))$
Wie in Bemerkung 19 sind alle Vektormengen

$$\{F(\bar{z}), D_x h_i(\bar{z}), D_x g_j(\bar{z}), i \in I, j \in J_0(\bar{z}) \setminus \{p\}\} \tag{84}$$

für beliebiges $p \in J_0(\bar{z})$ linear unabhängig. Daraus folgt, daß $J_0(z_k) = J_0(\bar{z})$ für hinreichend große k. Damit ist die Bedingung **1.** der Punkte vom Typ 4 erfüllt. Laut Bedingung **2.** der Definition vom Typ 4 hat die Matrix (78) den Rang $m + |J_0(\bar{z})| - 1$. Das erlaubt für hinreichend große k nur die zwei folgenden Fälle:

1. LICQ ist in z_k erfüllt. Es seien (λ_k, μ_k) die zu z_k gehörenden Multiplikatoren, womit

$$\mathbb{L}^k_{J_0(\bar{z})}(z_k, \lambda_k, \mu_k) = 0 \tag{85}$$

gilt.

Ist die Folge $\{(\lambda_k, \mu_k)\}$ beschränkt, dann konvergiert sie o.B.d.A. gegen einen Vektor $(\bar{\lambda}, \bar{\mu})$. Nach einem Grenzübergang in (85) ergibt sich $\overline{\mathbb{L}}_{J_0(\bar{z})}(\bar{z})(\bar{z}, \bar{\lambda}, \bar{\mu}) = 0$, was der linearen Unabhängigkeit von (84) widerspricht. Das heißt o.B.d.A. $\|(\lambda_k, \mu_k)\| \to +\infty$.

Betrachten wir die folgenden Gleichungen.

$$\frac{1}{\|(\lambda_k, \mu_k)\|}\mathbb{L}^k_{J_0(\bar{z})}(z_k, \lambda_k, \mu_k) = 0 \tag{86}$$

Falls die **VU-ND1** in z_k verletzt ist, können wir o.B.d.A. annehmen, daß die Komponente $[\mu_k]_p = 0$ ist mit $p \in J_0(\bar{z})$. $\{\frac{1}{\|(\lambda_k, \mu_k)\|}(\lambda_k, \mu_k)\}$ konvergiere O.B.d.A. gegen $(\bar{\lambda}, \bar{\mu}) \neq 0$. Nach einem Grenzübergang in (86) kommt man zur folgende Beziehung.

$$\bar{\mathbf{L}}(\bar{z}) = \sum_{i \in I} \bar{\lambda}_i \bar{h}_i(\bar{z}) + \sum_{j \in J_0(\bar{z})} \bar{\mu}_j \bar{g}_j(\bar{z}) = 0 \tag{87}$$

Dobei gilt auch $\bar{\mu}_p = 0$. Das widerspricht aber der Bedingung **4.** bei der Definition des Typs 4 für $\bar{z}$. **VU-ND1** ist in z_k also erfüllt.

Ist die Bedingung **VU-ND2** für z_k verletzt, benutzen wir wieder die Folge $\{\frac{1}{\|(\lambda_k, \mu_k)\|}(\lambda_k, \mu_k)\}$, die o.B.d.A. gegen $(\bar{\lambda}, \bar{\mu}) \neq 0$ konvergiert. Dabei erfüllt $(\bar{\lambda}, \bar{\mu})$ die Beziehung (87) und o.B.d.A. $\bar{\mu}_p = 1$. Die Matrix

$$\bar{B}^0 = \left[\bar{F}(\bar{z}) \vdots D_x^T \bar{H}(\bar{z}) \vdots D_x^T \bar{G}_{J_0(\bar{z}) \setminus \{p\}}(\bar{z})\right] \tag{88}$$

hat vollen Rang und wegen **2.-** auch die Matrizen

$$B_k^0 = \left[F^k(z_k) \vdots D_x^T H^k(z_k) \vdots D_x^T G^k_{J_0(\bar{z}) \setminus \{p\}}(z_k)\right].$$

Es seien $\bar{W}^0$ eine vollen Rang besitzende Matrix der Größe $n \times (n - m - |J_0(\bar{z})|)$ mit $\left[\bar{W}^0\right]^T \bar{B}^0 = 0$. Mit $W_k^0 = \bar{W}^0 - B_k^0 \left[B_k^0\right]^\dagger \bar{W}^0$ konvergieren

$W_k^0 \to \bar{W}^0$, und nach **VU-ND1** bzgl. z_k bilden ihre Spalten eine Basis von $T_{x_k}M(t_k)$. Nach Bedingung **4.** und **Typ4: 5-VU** der Definition vom Typ 4 ist die Matrix $\left[\bar{W}^0\right]^T D_x\bar{\mathbf{L}}(\bar{z})\bar{W}^0$ regulär (siehe auch Beweis der Behauptung 7). Die Bedingung **VU-ND2** bzgl. z_k folgt jetzt aus der Konvergenz von

$$\det(\left[W_k^0\right]^T \left[\frac{1}{\|(\lambda_k,\mu_k)\|} D_x\mathbb{L}^k_{J_0(\bar{z})}(z_k,\lambda_k,\mu_k)\right] W_k^0)$$

gegen $\det(\left[\bar{W}^0\right]^T D_x\bar{\mathbf{L}}(\bar{z})\bar{W}^0)$. Damit gilt $z_k \in \Sigma^1_{gc}(F^k,H^k,G^k)$.

2. Der Rang von

$$\left[D_x^T H^k(z_k) \vdots D_x^T G^k_{J_0(z_k)}(z_k)\right] \tag{89}$$

ist $m + |J_0(\bar{z})| - 1$, was der Bedingung **2.** der Definition vom Typ 4 entspricht. Nach Bedingung **2.** und **3.** der Definition vom Typ 4 bzgl. $\bar{z}$ hat die Matrix

$$\bar{B}_1 = \left[D_x^T \bar{H}(\bar{z}) \vdots D_x^T \bar{G}_{J_0(\bar{z})\setminus\{p\}}(\bar{z})\right] \tag{90}$$

für beliebiges, aber jetzt festes $p \in J_0(\bar{z})$ vollen Rang. Es seien $(\bar{\lambda},\bar{\mu})$ die Beziehung (87) erfüllende Koeffizienten mit o.B.d.A. $\bar{\mu}_p = 1$. Sie haben dann die Gestalt $(\bar{\lambda},\bar{\mu}) = (-\left[\bar{B}_1\right]^\dagger D_x^T g_p(\bar{z}),1)$. Die Matrizen B_1^k (wie in (83) erklärt) konvergieren gegen $\bar{B}_1$ und haben deswegen für hinreichend große k vollen Rang. Für die Koeffizienten (λ^k,μ^k) einer verschwindenden Linearkombination der Spalten von (89) können wir dann $\mu_1^k = 1$ annehmen und daher die Gestalt

$$(\lambda^k,\mu^k) = (-\left[B_1^k\right]^\dagger D_x^T g_p(z_k),1)$$

sichern. Bedingung **3.** der Definition vom Typ 4 bzgl. z_k folgt aus der gleichen Bedingung bzgl. $\bar{z}$, denn $(\lambda^k,\mu^k) \to (\bar{\lambda},\bar{\mu})$.

Aus dieser Konvergenz folgt also die Konvergenz von $D_z\mathbf{L}_k(z_k,\lambda^k,\mu^k) = D_z\left[\sum_{i\in I}\lambda_i^k h_i^k(z^k) + \sum_{j\in J_0(\bar{z})}\mu_j^k g_j^k(z^k)\right]$ gegen $D_z\bar{\mathbf{L}}(\bar{z},\bar{\lambda},\bar{\mu})$. Es seien die Spalten von $\bar{W}_0$ eine Basis des Raumes $\bar{T} = \{\eta \in I\!R^n | \eta^T\bar{B}_1 = 0\}$.

Die Matrizen W_0^k (wie im (82) erklärt) bilden für hinreichend große k auch eine Basis für $T_k = \{\eta \in I\!R^n | \eta^T B_1^k = 0\}$. Damit konvergieren die Matrizen

$$A_k = D_t \mathbf{L}_k(z_k, \lambda^k, \mu^k) \left[W_0^k\right]^T \left[D_x \mathbf{L}_k(z_k, \lambda^k, \mu^k)\right] W_0^k$$

gegen $\bar{A} = D_t \bar{\mathbf{L}}(\bar{z}, \bar{\lambda}, \bar{\mu}) \left[\bar{W}_0\right]^T \left[D_x \bar{\mathbf{L}}(\bar{z}, \bar{\lambda}, \bar{\mu})\right] \bar{W}_0$. Die Bedingung **4.** der Definition vom Typ 4 bzgl. z_k folgt dann aus der gleichen Bedingung bzgl. $\bar{z}$.

Da $F^k(z_k) \to \bar{F}(\bar{z})$, folgt aus der vorher gesehenen Konvergenz und aus **Typ4: 5-VU** bzgl. $\bar{z}$ die letzte Bedingung **Typ4: 5-VU** bzgl z_k. Damit gilt $z_k \in \Sigma_{gc}^4(F^k, H^k, G^k)$.

<u>Typ 5:</u> $(\bar{z} \in \Sigma_{gc}^5(\bar{F}, \bar{H}, \bar{G}))$

Nach den Bedingungen **1.** bis **3.** der Definition vom Typ 5 ist für einen beliebigen Index $p \in J_0(\bar{z})$ die quadratische Matrix $\bar{B}^1$ (wie in (90) erklärt) regulär. Daraus folgt für hinreichend große k auch die Regularität der Matrizen B_1^k (wie in (83) erklärt). Da $J_0(z_k) \subset J_0(\bar{z})$, sind alle $z_k \in \Sigma_{gc}(F^k, H^k, G^k)$ auch kritische Punkte.

Bedingung **Typ5: 4-VU** impliziert, daß im Vektor

$$(\bar{\theta}, \bar{\vartheta}) = (\bar{\theta}_i, i \in I, \bar{\vartheta}_j, j \in J_0(\bar{z}) \setminus \{p\}) = \left[\bar{B}_1\right]^{-1} F(\bar{z})$$

alle Komponenten $\bar{\vartheta}_j, j \in J_0(\bar{z}) \setminus \{p\}$, verschieden von Null sind. Damit sind auch alle solchen Komponenten bei $(\theta^k, \vartheta^k) = \left[B_1^k\right]^{-1} F(z_k)$ für hinreichend große k verschieden von Null. Das läßt nur zwei Möglichkeiten zu:

1. $J_0(z_k) = n - m$, d.h. ein solcher Index p wurde aus $J_0(z_k)$ weggelassen, und dabei ist auch **VU-ND1** erfüllt. Die Bedingung **VU-ND2** folgt automatisch, denn $T_{x_k} M(t_k) = \{0\}$. Das heißt: $z_k \in \Sigma_{gc}^1(F^k, H^k, G^k)$.

2. $J_0(z_k) = n - m + 1$. Hier ist die Bedingung **1.** der Definition vom Typ 5 schon erfüllt. Aus **2.-** konvergiert det $\left[D^T H^k(z_k) \vdots D^T G_{J_0(\bar{z})}^k(z_k)\right]$ gegen det $\left[D^T \bar{H}(\bar{z}) \vdots D^T \bar{G}_{J_0(\bar{z})}(\bar{z})\right]$,und damit folgt Bedingung **2.** der Definition vom Typ 5 bzgl. z_k unmittelbar aus derselben Bedingung bzgl. $\bar{z}$.

Mit den gleichen Gedankengängen wie für die Bedingung **3.** bei Punkten vom Typ 4 folgt die Bedingung **3.** vom Typ 5 in z_k.

Da alle Komponenten ϑ_j^k, $j \in J_0(\bar{z}) \setminus \{p\}$, von $(\theta^k, \vartheta^k) = \left[B_1^k\right]^{-1} F(z_k)$ verschieden von Null sind, gilt dann letztlich **Typ5: 4-VU**, und damit gilt $z_k \in \Sigma_{gc}^1(F^k, H^k, G^k)$. *Box*

Satz 19
Es seien $(\bar{F}, \bar{H}, \bar{G}) \in C^3(I\!R^{n+1}, I\!R^{n+m+p})$ *und* $\bar{z} = (\bar{x}, \bar{t})$ *fixiert. Es sei* $\bar{z} \in \Sigma_{gc}^i(\bar{F}, \bar{H}, \bar{G})$ *(bzw.* $\bar{z} \notin \Sigma_{gc}(\bar{F}, \bar{H}, \bar{G})$*) mit* $i \in \{1, \ldots, 5\}$*. Dann existieren eine offene Umgebung* $U_{\bar{z}}$ *von* $\bar{z}$ *und eine positive Zahl* $r_{\bar{z}}$*, so daß für alle Abbildungen* $(\tilde{F}, \tilde{H}, \tilde{G}) \in V^3((\bar{F}, \bar{H}, \bar{G}), U_{\bar{z}}, r_{\bar{z}})$ *folgendes gilt:* $\Sigma_{gc}(\tilde{F}, \tilde{H}, \tilde{G}) \cap U_{\bar{z}} \subset \Sigma_{gc}^i(\tilde{F}, \tilde{H}, \tilde{G}) \cup \Sigma_{gc}^1(\tilde{F}, \tilde{H}, \tilde{G})$ *(bzw.* $\Sigma_{gc}(\tilde{F}, \tilde{H}, \tilde{G}) \cap U_{\bar{z}} = \emptyset$*).*

<u>Beweis:</u>
Es sei der Satz für $\bar{z}$ und $(\bar{F}, \bar{H}, \bar{G})$ nicht erfüllt, und wir betrachten z.B. den Fall $\bar{z} \in \Sigma_{gc}^3(\bar{F}, \bar{H}, \bar{G})$. Wählen wir die Umgebungen $U_k = \{z \in I\!R^{n+1} | \, \|z - \bar{z}\| \leq \frac{1}{k}\}$ und die positiven Zahlen $r_k = \frac{1}{k}$, ann existieren die Folgen

$$\begin{aligned} (F^k, H^k, G^k) &\in V^3((\bar{F}, \bar{H}, \bar{G}), U_k, r_k), \\ z_k &\in U_k, \end{aligned} \tag{91}$$

so daß

$$z_k \in \Sigma_{gc}(F^k, H^k, G^k) \setminus \{\Sigma_{gc}^3(F^k, H^k, G^k) \cup \Sigma_{gc}^1(F^k, H^k, G^k)\} \tag{92}$$

gilt.

Da z_k in U_k enthalten ist, folgt $\|z_k - \bar{z}\| \leq \frac{1}{k}$. Das heißt, $z_k \to \bar{z}$. Es sei $\alpha \in \mathbb{N}^{n+1}$ mit $|\alpha| \leq 3$. Laut Bedingung (91), auf (F^k, H^k, G^k) angewandt, gilt dann

$$|D_\alpha(F^k, H^k, G^k)(z_k) - D_\alpha(\bar{F}, \bar{H}, \bar{G})(z_k)| < \frac{1}{k}. \tag{93}$$

Alle Abbildungen $D_\alpha(\bar{F}, \bar{H}, \bar{G})(z)$ sind stetig, und deswegen konvergieren $D_\alpha(\bar{F}, \bar{H}, \bar{G})(z_k)$ gegen $D_\alpha(\bar{F}, \bar{H}, \bar{G})(\bar{z})$. Damit ist auch wegen (93) die Konvergenz $D_\alpha(F^k, H^k, G^k)(z_k) \to D_\alpha(\bar{F}, \bar{H}, \bar{G})(\bar{z})$ gesichert. Somit sind die Voraussetzungen des Lemma 6 erfüllt, und es ist ein Wiederspruch zu (92) erzeugt. □

Bevor wir zum Störungssatz kommen, konzentrieren wir uns auf eine Beziehung, die Satz 10.2.2 in [32] für Gradientenabbildungen entspricht. Dieses Ergebnis wird für die Untersuchung der Punkte vom Typ 3 und beim Beweis des Störungssatzes benutzt.

Es sei $q \leq n$ eine feste ganze Zahl, und wir betrachten Abbildungen der Form

$$T(x,u,t) = \begin{pmatrix} T^1(x,u,t) \\ T^2(x,t) \end{pmatrix}, \tag{94}$$

wobei $x \in I\!R^n$, $u \in I\!R^q$, $T : I\!R^{n+q+1} \to I\!R^{n+q}$, $T^1 : I\!R^{n+q+1} \to I\!R^n$ und damit $T^2 : I\!R^{n+1} \to I\!R^q$. Wir betrachten dann zwei Matrizen.

$$B_L = \begin{pmatrix} C_1 & C_2 \\ 0 & I_q \end{pmatrix} \tag{95}$$

und

$$B_R = \begin{pmatrix} C_3 & 0 \\ C_4 & I_q \end{pmatrix} \tag{96}$$

Dabei sind C_1 und C_3 quadratische reguläre Matrizen der Größe n, und I_q enstpricht der Einheitsmatrix der Größe q. Damit sind B_L und B_R auch quadratische reguläre Matrizen der Größe $n+q$.

Es sei mit

$$T_{(B_L,B_R)}(y,v,t) = B_L \cdot T(B_R \cdot (y,v),t) = \begin{pmatrix} T^1_{(B_L,B_R)}(y,v,t) \\ T^2_{(B_L,B_R)}(y,t) \end{pmatrix} \tag{97}$$

eine neue Abbildung bezeichnet. Wegen der Struktur vom B_L und B_R hat diese Abbildung die gleiche Abhängigkeit von den Variablen (y,v,t) wie T, was in der vorhergehenden Beziehung mittels $T^1_{(B_L,B_R)}$ und $T^2_{(B_L,B_R)}$ angedeutet wird.

Aus den vorherigen Vereinbarungen folgt, daß (x,u,t) genau dann eine Nullstelle von $T(x,u,t)$ ist, wenn auch $(B_R^{-1} \cdot (x,u),t)$ eine Nullstelle von $T_{(B_L,B_R)}(y,v,t)$ ist. Definiert man die Menge

$$\Sigma(T) = \{(x,u,t) | T(x,u,t) = 0\},$$

ergibt sich folgende Beziehung.

$$\Sigma(T) = \begin{pmatrix} B_R & 0 \\ 0 & 1 \end{pmatrix} \Sigma(T_{(B_L,B_R)}) \tag{98}$$

Betrachten wir folgende Annahme für eine Abbildung $T(x,u,t)$.

$$P_1(T): \qquad 0 \text{ ist ein Regulärwert von } T$$

Aus der Beziehung

$$DT_{(B_L,B_R))}(y,v,t) = B_L \cdot DT(B_R \cdot (y,v),t) \cdot \begin{pmatrix} B_R & 0 \\ 0 & 1 \end{pmatrix} \tag{99}$$

folgt dann die Bemerkung 27.

Bemerkung 27
Es seien B_L und B_R reguläre Matrizen wie in (95) bzw. (96) gewählt. Weiter seien die Abbildungen $T(x,u,t)$ und $T_{(B_L,B_R)}(y,v,t)$ wie in (94) bzw. (97). $P_1(T)$ gilt genau dann, wenn $P_1(T_{(B_L,B_R)})$ gilt.

Wir stellen jetzt einige Eigenschaften der Funktionen (94) auf, die äquivalent auch für die Funktionen (97) gelten. Dazu führen wir folgende Menge ein:

$$\Sigma^s(T) = \left\{(x,u,t) \in \Sigma(T) | \det(D_{(x,u)}T(x,u,t)) = 0\right\}.$$

Bemerkung 28
Aus (99) folgt für $(x,u) = B_R \cdot (y,v)$ die Beziehung

$$D_{(y,v)}T(y,v,t) = B_L \cdot D_{(x,u)}T(x,u,t) \cdot B_R,$$

und, da B_L und B_R regulär sind, gilt:

$$\Sigma^s(T) = \begin{pmatrix} B_R & 0 \\ 0 & 1 \end{pmatrix} \Sigma^s(T_{(B_L,B_R)}).$$

Behauptung 8
Es seien $(x,u,t) \in \Sigma^s(T)$ und $(y,v,t) \in \Sigma^s(T_{(B_L,B_R)})$ mit $(x,u) = B_R \cdot (y,v)$. Die Bedingung $P_1(T)$ (und damit auch $P_1(T_{(B_L,B_R)})$) sei mindestens in einer Umgebung von (x,u,t) erfüllt. Dann ist der Punkt (x,u,t) ein nichtdegenerierter kritischer Punkt von $t|_{\Sigma(T)}$ (Optimierungsproblem mit der Zielfunktion $F(x,u,t) = t$ und der zulässigen Menge $\Sigma(T)$) genau dann, wenn (y,v,t) ein nichtdegenerierter kritischer Punkt von $t|_{\Sigma(T_{(B_L,B_R)})}$ ist.

Beweis:

Aus $(x,u,t) \in \Sigma^s(T)$, $(y,v,t) \in \Sigma^s(T_{(B_L,B_R)})$ und $P_1(T)$, $P_1(T_{(B_L,B_R)})$ folgen die LICQ und daß beide Punkte kritisch sind. Es bleibt damit nur die Äquivalenz der Bedingungen zweiter Ordnung **ND2** zu überprüfen.

Es sei (y,v,t) nichtdegeneriert bzgl. $t|_{\Sigma(T_{(B_L,B_R)})}$. Mit $m_{(y,v)}$ bezeichnen wir die zugehörigen Multiplikatoren und mit $(\tau_{(y,v)},1)$ einen Vektor, der den eindimensionalen Tangentialraum $T_{(y,v,t)}\Sigma(T_{(B_L,B_R)})$ linear erzeugt. Nach (99) gilt, daß die zu (x,u,t) gehörenden Multiplikatoren $m_{(x,u)}$ die Gestalt $m_{(x,u)} = B_L^T \cdot m_{(y,v)}$ haben und daß der Vektor $(\tau_{(x,u)},1) = (B_R \cdot \tau_{(y,v)},1)$ den Tangentialraum $T_{(x,u,t)}\Sigma(T)$ linear erzeugt.

Bezeichnen wir mit $\mathbb{L}\,(x,u,t) = t - m_{(x,u)}^T \cdot T(x,u,t)$ und $\mathbb{L}_{(B_L,B_R)}(y,v,t) = t - m_{(y,v)}^T \cdot T_{(B_L,B_R)}(y,v,t)$ die Lagrange-Funktionen bezüglich der Probleme $t|_{\Sigma(T)}$ (bzw. $t|_{\Sigma(T_{(B_L,B_R)})}$). Die Beschränkung von $D^2\mathbb{L}_{(B_L,B_R)}(y,v,t)$ auf den Tangentialraum $T_{(y,v,t)}\Sigma(T_{(B_L,B_R)})$ hat dann die folgende Gestalt.

$$D^2\mathbb{L}_{(B_L,B_R)(y,v,t)}|_{T_{(y,v,t)}\Sigma(T_{(B_L,B_R)})} = -(\tau_{(y,v)}^T,1)\left[D^2 m_{(y,v)}^T \cdot T_{(B_R(y,v),t)}\right](\tau_{(y,v)},1)$$

$$\begin{aligned} &= -(\tau_{(y,v)}^T \cdot B_R^T,1)\left[D^2 m_{(y,v)}^T \cdot B_L \cdot T_{(B_R(y,v),t)}\right](B_R \cdot \tau_{(y,v)},1) \\ &= -(\tau_{(x,u)}^T,1)\left[D^2 m_{(x,u)}^T \cdot T_{(x,u,t)}\right](\tau_{(x,u)},1) \end{aligned}$$

Damit ist die Äquivalenz beider Bedingungen ND2 geklärt und der Beweis fertig. □

Bemerkung 29

Es seien die Bedingung $P_1(T)$ erfüllt und $(y(\theta),v(\theta),t(\theta))$ eine Parametrisierung der Kurve $\Sigma(T_{(B_L,B_R)})$ in einer Umgebung eines Punktes $(y,v,t) \in \Sigma^s(T_{(B_L,B_R)})$, wobei die Parametrisierung in einer Umgebung vom Null definiert ist und $(y(0),v(0),t(0)) = (y,v,t)$. Nach (98) ist die Abbildung

$$\begin{pmatrix} x(\theta) \\ u(\theta) \\ t(\theta) \end{pmatrix} = \begin{pmatrix} B_R & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y(\theta) \\ v(\theta) \\ t(\theta) \end{pmatrix}$$

eine lokale Parametrisierung der Kurve $\Sigma(T)$ in einer Umgebung des Punktes (x,u,t).

Da $(y, v, t) \in \Sigma^s(T_{(B_L,B_R)})$ *und* $(x, u, t) \in \Sigma^s(T)$ *sind, haben die Funktionen*

$$\Psi_{(B_L,B_R)}(\theta) = \det\left[D_{(y,v)}T_{(B_L,B_R)}(y(\theta), v(\theta), t(\theta))\right]$$

und

$$\Psi(\theta) = \det\left[D_{(x,u)}T(x(\theta), u(\theta), t(\theta))\right] \tag{100}$$

eine Nullstelle in Null. Es gilt also

$$\Psi_{(B_L,B_R)}(\theta) = \det[B_L] \cdot \det[B_L] \cdot \Psi(\theta).$$

Daraus folgt, daß die Ableitung $\Psi'(0)$ *genau dann verschwindet, wenn auch* $\Psi'_{(B_L,B_R)}(0)$ *gleich Null ist.*

Wegen der speziellen Abhängigkeit (94) der Abbildungen T und $T_{(B_L,B_R)}$ von den jeweiligen Variablen haben beide Ableitungen $D_{(y,v)}T_{(B_L,B_R)}$ und $D_{(x,u)}T$ folgende Struktur:

$$D_{(x,u)}T = \begin{pmatrix} D_xT_1 & D_uT_1 \\ D_xT_2 & 0 \end{pmatrix}.$$

Definition 35
Es sei mit $\mathbf{M}_1(n+q)$ *die Menge aller quadratischen Matrizen der Größe* $n+q$ *bezeichnet, in denen die aus den letzten* q *Zeilen und Spalten gebildete Teilmatrix die Nullmatrix ist.* $\mathbf{M}_1(n+q)$ *ist direkt mit dem Raum* $\mathbb{R}^{n(n+2q)}$ *identifizierbar, und seien die offene Menge*

$$\mathbf{M}_2(n+q) = \left\{ \begin{pmatrix} A & B \\ C & 0 \end{pmatrix} \in \mathbf{M}_1(n+q) \mid \text{rank } (B) = \text{ rank } (C) = q \right\}$$

und die Mannigfaltigkeit

$$\mathbf{M}_3(n+q, k) = \{P \in \mathbf{M}_2(n+q) \mid \text{rank } (P) = k\}$$

definiert, wobei k *eine feste natürliche Zahl mit* $2q \leq k \leq n+q$ *ist.*

Laut der vorigen Definition gilt $D_{(x,u)}T \in \mathbf{M}_1(n+q)$.

Bemerkung 30
Bezeichnen wir die Elemente $P \in \mathbf{M}_3(n+q,k)$ mit

$$P = \begin{pmatrix} A & B \\ C & 0 \end{pmatrix}.$$

Es sei $\bar{P} \in \mathbf{M}_3(n+q,k)$ fest. Laut Definition haben die Matrizen $\bar{B}$ und $\bar{C}$ den Rang q und $\bar{P}$ den Rang k. Aus der Struktur von $\bar{P}$ folgt dann die Existenz zweier Indexmengen $\{n+1,\ldots,n+q\} \subset I_1, I_2 \subset \{1,\ldots,n+q\}$, so daß $|I_1| = |I_2| = k$ und die Teilmatrix, die aus den Zeilen I_1 und den Spalten I_2 gebildet wird (bezeichnet mit $\bar{P}_{(I_1,I_2)}$), regulär ist.

Bezeichnen wir mit $I_1^c = \{1,\ldots,n+q\} \setminus I_1$ (bzw. I_2^c) die restlichen Indexmengen und mit $\bar{P}_{(I_1^c,I_2)}$, $\bar{P}_{(I_1,I_2^c)}$,$\bar{P}_{(I_1^c,I_2^c)}$ die anderen Teilmatrizen von $\bar{P}$. Wählen wir eine Umgebung $U(\bar{P}) \subset \mathbf{M}_1(n+q)$ von $\bar{P}$ so, daß $U(\bar{P}) \subset \mathbf{M}_2(n+q)$ gilt ($\mathbf{M}_2(n+q)$ ist selber offen) und daß $\forall P \in U(\bar{P})$ die Teilmatrix $P_{(I_1,I_2)}$ regulär ist.

Die Mannigfaltigkeit $\mathbf{M}_3(n+q,k)$ läßt sich jetzt in der Umgebung $U(\bar{P})$ durch die Gleichungen

$$P_{(I_1^c,I_2^c)} - P_{(I_1^c,I_2)} \left[P_{(I_1,I_2)}\right]^{-1} P_{(I_1,I_2^c)} = 0$$

beschreiben. Die Gradienten dieser Funktionen sind linear unabhängig (die Ableitung nach den Elementen der Teilmatrix $P_{(I_1^c,I_2^c)}$ bilden eine Einheitsmatrix der Größe $(n+q-k)^2$). Damit ist $\mathbf{M}_3(n+q,k)$ eine Mannigfaltigkeit der Kodimension $(n+q-k)^2$.

Definition 36
Es sei $F : I\!R^{m_1} \to I\!R^{m2}$ eine differenzierbare Abbildung. Die m_1-dimensionale Mannigfaltigkeit $Graf(F)$ wird wie folgt definiert:

$$Graf(F) = \left\{(x,y) \in I\!R^{m_1+m_2} | y - F(x) = 0\right\}.$$

Es sei $M_2 \subset I\!R^{m_2}$ eine Mannigfaltigkeit. F und M_2 schneiden sich transversal ($F \overset{T}{\cap} M_2$), falls $Graf(F) \overset{T}{\cap} I\!R^{m_1} \times M_2$.

Es sei $T(x,u,t)$ eine Abbildung wie in (94), und wir betrachten folgende Bedingung:

$$P_2(T): \qquad (T, D_{(x,u)}T) \overset{T}{\cap} \{0_{(n+q)}\} \times \mathbf{M}_3(n+q, n+q-1).$$

Wegen der Struktur der Matrizen B_L und B_R gilt, daß

$$(T(\bar{x},\bar{u},\bar{t}), D_{(x,u)}T(\bar{x},\bar{u},\bar{t})) \in \{0_{(n+q)}\} \times \mathbf{M}_3(n+q, n+q-1)$$

genau dann, wenn auch

$$(T_{(B_L,B_R)}(\bar{y},\bar{v},\bar{t}), D_{(y,v)}T_{(B_L,B_R)}(\bar{y},\bar{v},\bar{t})) \in \{0_{(n+q)}\} \times \mathbf{M}_3(n+q, n+q-1)$$

erfüllt wird, wobei $(\bar{y},\bar{v},\bar{t}) = (B_R^{-1} \cdot (\bar{x},\bar{u}), \bar{t})$ ist.

Behauptung 9
Es seien B_L, B_R wie in (95) bzw. (96) und $T(x,u,t)$, $T_{(B_L,B_R)}(y,v,t)$ wie in (94) bzw. (97) gewählt. Weiter seien $(T(\bar{x},\bar{u},\bar{t}), D_{(x,u)}T(\bar{x},\bar{u},\bar{t}))$ und $(T_{(B_L,B_R)}(\bar{y},\bar{v},\bar{t}), D_{(y,v)}T_{(B_L,B_R)}(\bar{y},\bar{v},\bar{t}))$ in $\{0_{(n+q)}\} \times \mathbf{M}_3(n+q, n+q-1)$ mit $(\bar{y},\bar{v},\bar{t}) = (B_R^{-1} \cdot (\bar{x},\bar{u}), \bar{t})$. Die Bedingung $P_2(T)$ gilt genau dann in einer Umgebung von $((\bar{x},\bar{u},\bar{t}), 0_{(n+q)}, D_{(x,u)}T(\bar{x},\bar{u},\bar{t}))$, wenn sie auch in einer Umgebung von $((\bar{y},\bar{v},\bar{t}), 0_{(n+q)}, D_{(y,v)}T_{(B_L,B_R)}(\bar{y},\bar{v},\bar{t}))$ gilt.

<u>Beweis:</u>

Es seien $z_1 \in I\!R^{n+q}$ und $z_2 \in \mathbf{M}_1(n+q)$. Mit $\Phi(z_2)$ bezeichnen wir eine differenzierbare Funktion, deren Nullstellen in einer Umgebung von $D_{(x,u)}T(\bar{x},\bar{u},\bar{t})$ die Mannigfaltigkeit $\mathbf{M}_3(n+q, n+q-1)$ darstellen. Die Mannigfaltigkeiten $Graf(T(x,u,t), D_{(x,u)}T(x,u,t))$ und $I\!R^{n+q} \times \mathbf{M}_1(n+q) \times \{0_{(n+q)}\} \times \mathbf{M}_3(n+q, n+q-1)$ seien in einer Umgebung ihres gemeinsamen Punktes $((\bar{x},\bar{u},\bar{t}), 0_{(n+q)}, D_{(x,u)}T(\bar{x},\bar{u},\bar{t}))$ transversal geschnitten. Das ist äquivalent dazu, daß die Abbildung

$$\Omega_1(x,u,t,z_1,z_2) = \begin{pmatrix} z_1 - T(x,u,t) \\ z_2 - D_{(x,u)}T(x,u,t) \\ z_1 \\ \Phi(z_2) \end{pmatrix}, \tag{101}$$

deren Nullstellen in einer Umgebung von $(\bar{x},\bar{u},\bar{t}, 0_{(n+q)}, D_{(x,u)}T(\bar{x},\bar{u},\bar{t}))$ die gemeinsamen Punkte beider Mannigfaltigkeiten beschreiben, eine Jacobi-Matrix in $(\bar{x},\bar{u},\bar{t}, 0_{(n+q)}, D_{(x,u)}T(\bar{x},\bar{u},\bar{t}))$ mit linear unabhängigen Zeilen hat.

Betrachten wir jetzt die folgende reguläre lineare Koordinatentransformation.

$$\begin{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix} \\ t \\ z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} B_R \begin{pmatrix} y \\ v \end{pmatrix} \\ t \\ B_L^{-1}\gamma_1 \\ B_L^{-1}\gamma_2 B_R^{-1} \end{pmatrix} \tag{102}$$

Dabei sind $\gamma_1 \in I\!R^{n+q}$ und $\gamma_2 \in \mathbf{M}_1(n+q)$. Es sei bemerkt, daß die Abbildung $\Lambda(\gamma_2) = B_L^{-1}\gamma_2 B_R^{-1}$ linear und regulär ist. Die Matrizen B_L^{-1} und B_R^{-1} behalten dieselbe Struktur von B_L bzw. B_R. Daher folgt, daß $\Lambda(\gamma_2)$ zur Mannigfaltigkeit $\mathbf{M}_3(n+q, n+q-1)$ genau dann gehört, wenn γ_2 auch dazu gehört.

Ist $\bar{\gamma}_2 = B_L D_{(x,u)} T(\bar{x}, \bar{u}, \bar{t}) B_R \in \mathbf{M}_3(n+q, n+q-1)$, gibt es eine Umgebung von $\bar{\gamma}_2$ in $\mathbf{M}_1(n+q)$ so, daß die Nullstellen der differenzierbaren und im $\bar{\gamma}_2$ regulären Abbildung $\Psi(\gamma_2) = \Phi(B_L^{-1}\gamma_2 B_R^{-1})$ die Mannigfaltigkeit $\mathbf{M}_3(n+q, n+q-1)$ in dieser Umgebung darstellen.

In den neuen Koordinaten $(y, v, t, \gamma_1, \gamma_2)$ schreibt sich die Abbildung Ω_1 wie folgt:

$$\Omega_2(y, v, t, \gamma_1, \gamma_2) = \begin{pmatrix} B_L^{-1}\gamma_1 - T(B_R(y,v), t) \\ B_L^{-1}\gamma_2 B_R^{-1} - D_{(x,u)} T(B_R(y,v), t) \\ B_L^{-1}\gamma_1 \\ \Phi(B_L^{-1}\gamma_2 B_R^{-1}) \end{pmatrix}.$$

Diese Abbildung hat natürlich immer noch eine Jakobi-Matrix im Punkt $(\bar{y}, \bar{v}, \bar{t}, 0_{(n+q)}, \bar{\gamma}_2) = (B_R^{-1}(\bar{x}, \bar{u}), \bar{t}, 0_{(n+q)}, B_L D_{(x,u)} T(\bar{x}, \bar{u}, \bar{t}) B_R)$ mit unabhängigen Zeilen. Die Menge der Nullstellen von Ω_2 in einer Umgebung dieses Punktes ist die gleiche wie die der neuen Abbildung (aus Ω_2 durch eine lineare Transformation im Bildraum konstruiert) Ω_3.

$$\begin{aligned} \Omega_3(y, v, t, \gamma_1, \gamma_2) &= \begin{pmatrix} \gamma_1 - B_L T(B_R(y,v), t) \\ \gamma_2 - B_L D_{(x,u)} T(B_R(y,v), t) B_R \\ \gamma_1 \\ \Psi(\gamma_2) \end{pmatrix} \\ &= \begin{pmatrix} \gamma_1 - T_{(B_L, B_R)}(y, v, t) \\ \gamma_2 - D_{(y,v)} T_{(B_L, B_R)}(y, v, t) \\ \gamma_1 \\ \Psi(\gamma_2) \end{pmatrix} \end{aligned}$$

Die Abbildung Ω_3 hat in

$$(\bar{y}, \bar{v}, \bar{t}, 0_{(n+q)}, \bar{\gamma}_2) = (\bar{y}, \bar{v}, \bar{t}, T_{(B_L, B_R)}(\bar{y}, \bar{v}, \bar{t}), D_{(y,v)} T_{(B_L, B_R)}(\bar{y}, \bar{v}, \bar{t}))$$

eine Nullstelle und eine Jakobi-Matrix mit unabhängigen Zeilen. Da die Nullstellen der ersten zwei Teile von Ω_3 in einer Umgebung dieses Punktes

$Graf(T_{(B_L,B_R)}(y,v,t), D_{(y,v)}T_{(B_L,B_R)}(y,v,t))$ darstellen und die Nullstellen der anderen zwei Teile die Mannigfaltigkeit $I\!R^{n+q} \times \mathbf{M}_1(n+q) \times \{0_{(n+q)}\} \times \mathbf{M}_3(n+q, n+q-1)$ beschreiben, gilt dann die Bedingung $P_2(T_{(B_L,B_R)})$ in einer Umgebung von $(\bar{y}, \bar{v}, \bar{t}, 0_{(n+q)}, \bar{\gamma}_2)$. $\square$

Satz 20 *(siehe Satz 10.2.2 in [32] für Gradientenabbildungen)*
Es seien T eine Abbildung wie in (94) und $(\bar{x}, \bar{u}, \bar{t})$ eine ihrer Nullstellen mit folgenden Eigenschaften:

- *$D_{(x,u,t)}T(\bar{x}, \bar{u}, \bar{t})$ hat den Rang $n+q$ (d.h. $P_1(T)$ ist mindestens in einer Umgebung von $(\bar{x}, \bar{u}, \bar{t})$ erfüllt).*
- $(T(\bar{x}, \bar{u}, \bar{t}), D_{(x,u)}T(\bar{x}, \bar{u}, \bar{t})) \in \{0_{(n+q)}\} \times \mathbf{M}_3(n+q, n+q-1)$

Dann sind die folgenden drei Bedingungen äquivalent:

1. *$(\bar{x}, \bar{u}, \bar{t})$ ist ein nichtdegenerierter kritischer Punkt von $t|_{\Sigma(T)}$ (siehe Behauptung 8).*
2. *Es seien $(x(\theta), u(\theta), t(\theta))$ eine lokale Parametrisierung von $\Sigma(T)$ mit $(x(0), u(0), t(0)) = (\bar{x}, \bar{u}, \bar{t})$ und $\Psi(\theta)$ wie in (100). Dann gilt $\Psi'(0) \neq 0$.*
3. *$P_2(T)$ ist in einer Umgebung von $((\bar{x}, \bar{u}, \bar{t}), 0_{(n+q)}, D_{(x,u)}T(\bar{x}, \bar{u}, \bar{t}))$ erfüllt.*

<u>Beweis:</u>
Aus den Voraussetzungen an $(\bar{x}, \bar{u}, \bar{t})$ ergeben sich folgende Bedingungen:

- $(\bar{x}, \bar{u}, \bar{t}) \in \Sigma^s(T)$ und damit $\Psi(0) = 0$.
- $(\bar{x}, \bar{u}, \bar{t})$ ist ein kritischer Punkt bezüglich $t|_{\Sigma(T)}$, in dem die LICQ erfüllt ist.

Da $D_{(x,u)}T(\bar{x}, \bar{u}, \bar{t}) \in \mathbf{M}_3(n+q, n+q-1)$, kann man zwei Matrizen B_L und B_R wie in (95) bzw. (96) so wählen, daß die folgende Struktur für $B_L \cdot D_{(x,u)}T(\bar{x}, \bar{u}, \bar{t}) \cdot B_R$ entsteht.

$$\begin{pmatrix} 0 & 0_{n-1}^T & 0_q^T \\ 0_{n-1} & A_1 & B_1 \\ 0_q & B_2^T & 0 \end{pmatrix} \tag{103}$$

Dabei sind 0_{n-1} und 0_q die Nullvektoren in $I\!R^n$ bzw. $I\!R^q$, A_1 eine quadratische $(n-1)\times(n-1)$ Matrix und B_1,B_2 $n\times(q-1)$ Matrizen. Die Teilmatrix

$$\begin{pmatrix} A_1 & B_1 \\ B_2^T & 0 \end{pmatrix} \tag{104}$$

hat vollen Rang.

Bei der Betrachtung der Bemerkungen 27 und 29 und der Behauptungen 8 und 9 kann man o.B.d.A. annehmen, daß $D_{(x,u)}T(\bar{x},\bar{u},\bar{t})$ die Struktur in (103) besitzt.

Wir bezeichnen mit $T_j^1(x,u,t)$ die Komponente j der Teilabbildung T^1 von T. Wegen der Voraussetzungen an T (speziell aus dem vollen Rang von $D_{(x,u,t)}T(\bar{x},\bar{u},\bar{t})$) gilt, daß $D_tT_1^1(\bar{x},\bar{u},\bar{t})\neq 0$ ist.

Es sei $(\bar{x},\bar{u},\bar{t})$ ein nichtdegenerierter kritischer Punkt von $t|_{\Sigma(T)}$. Die Multiplikatoren in diesem Punkt sind die Komponenten des Vektors

$$\frac{1}{D_tT_1^1(\bar{x},\bar{u},\bar{t})}(1,0_{n+q-1}).$$

Da $D_tT_1^1(\bar{x},\bar{u},\bar{t})$ verschieden von Null und (104) regulär sind, folgt aus der Struktur (103) von $D_{(x,u,t)}T(\bar{x},\bar{u},\bar{t})$, daß der Tangentialraum an $\Sigma(T)$ in $(\bar{x},\bar{u},\bar{t})$ der vom Vektor $(1,0_{n+q})$ erzeugte eindimensionale lineare Teilraum ist. Das, zusammen mit der Gestalt der Multiplikatoren, ergibt, daß die Bedingung zweiter Ordnung **ND2** bzgl. $t|_{\Sigma(T)}$ im Punkt $(\bar{x},\bar{u},\bar{t})$ äquivalent zur Ungleichung

$$D_{x_1}^2T_1^1(\bar{x},\bar{u},\bar{t})\neq 0 \tag{105}$$

ist.

Der Rest des Beweises besteht darin, zu zeigen, daß die anderen beiden Bedingungen auch äquivalent zur Ungleichung (105) sind.

Es seien $(x(\theta),u(\theta),t(\theta))$ und $\Psi(\theta)$ wie in der zweiten Bedingung. Für zwei Indizes $k,l\in\{1,\ldots,n+q\}$ bezeichnen wir mit $\Psi_{k,l}(\theta)$ die Determinante des (k,l)-Minor (die durch Streichung der Zeile k und der Spalte l enstandene $(n+q-1)\times(n+q-1)$-Teilmatrix) der Matrix $D_{(x,u)}T(x(\theta),u(\theta),t(\theta))$. Dann gilt:

$$\Psi(\theta)=\sum_{j=1}^{n}D_{x_j}T_1^1\big(x(\theta),u(\theta),t(\theta)\big)\cdot\Psi_{1,j}(\theta)+\sum_{k=1}^{q}D_{u_k}T_1^1\big(x(\theta),u(\theta),t(\theta)\big)\cdot\Psi_{1,n+k}(\theta)$$

Wegen der Struktur (103) von $D_{(x,u,t)}T(\bar{x},\bar{u},\bar{t})$ gilt:

$$\begin{aligned}
\Psi_{1,1}(0) &\neq 0 \\
\Psi_{1,j}(0) &= 0,\ \forall j = 2,\ldots,n+q \\
D_{x_j}T_1^1(\bar{x},\bar{u},\bar{t}) &= 0,\ \forall j = 1,\ldots,n \\
D_{u_k}T_1^1(\bar{x},\bar{u},\bar{t}) &= 0,\ \forall k = 1,\ldots,q
\end{aligned}$$

Deshalb reduziert sich die Ableitung $\Psi'(0)$ wie folgt:

$$\begin{aligned}
\Psi'(0) &= \frac{d}{d\theta}\left[D_{x_1}T_1^1(x(0),u(0),t(0))\right]\cdot\Psi_{1,1}(0) \\
&= D_{(x,u,t)}D_{x_1}T_1^1(\bar{x},\bar{u},\bar{t})\cdot\begin{pmatrix}\dot{x}(0)\\ \dot{u}(0)\\ \dot{t}(0)\end{pmatrix}\cdot\Psi_{1,1}(0)
\end{aligned}$$

Da die Ableitung der Parametrisierung in Null $(\dot{x}(0),\dot{u}(0),\dot{t}(0))$ zum Tangentialraum an $\Sigma(T)$ in $(\bar{x},\bar{u},\bar{t})$ gehört, der durch den Vektor $(1,0_{n+q})$ erzeugt wird, erhält man

$$\Psi'(0) = D_{x_1}^2T_1^1(\bar{x},\bar{u},\bar{t})\cdot\dot{x}_1(0)\cdot\Psi_{1,1}(0),$$

wobei $\dot{x}_1(0)$ und $\Psi_{1,1}(0)$ verschieden von Null sind. Damit ist die zweite Bedingung äquivalent zur Ungleichung (105).

Es seien jetzt $z \in I\!R^{n+q}$ und $\gamma \in M_2(n+q)$. Wegen der Struktur (103) von $D_{(x,u,t)}T(\bar{x},\bar{u},\bar{t})$ ist die Mannigfaltigkeit $M_3(n+q,n+q-1)$ in einer Umgebung dieses Punktes durch die Nullstellen einer Abbildung $\Lambda_4(\gamma)$ mit folgender Struktur (für Details siehe Bemerkung 30)

$$\Lambda_4(\gamma) = \gamma_{1,1} - \tilde{\Lambda}_4(\gamma)$$

gegeben. Speziell gilt $D_\gamma\tilde{\Lambda}_4(D_{(x,u)}T(\bar{x},\bar{u},\bar{t})) = 0$.

Die dritte Bedingung ist somit äquivalent dazu, daß die Jakobi-Matrix der folgenden Abbildung in der Nullstelle

$$(\bar{x},\bar{u},\bar{t},\bar{z},\bar{\gamma}) = (\bar{x},\bar{u},\bar{t},0_{n+q},D_{(x,u)}T(\bar{x},\bar{u},\bar{t}))$$

den Rang $2(n+q)+n(n+2q)+1$ hat.

$$\Lambda = \begin{pmatrix}\Lambda_1\\ \Lambda_2\\ \Lambda_3\\ \Lambda_4\end{pmatrix} = \begin{pmatrix} z - T(x,u,t)\\ \gamma - D_{(x,u)}T(x,u,t)\\ z\\ \gamma_{1,1} - \tilde{\Lambda}_4(\gamma)\end{pmatrix}$$

In der folgenden Darstellung dieser Jakobi-Matrix bezeichnen wir mit 0_k den $k-$dimensionalen Nullvektor (als Spaltenvektor) und mit $\bigotimes$ eine Teilmatrix, die für die Analyse nicht von Bedeutung ist.

$$\begin{array}{c} \\ \Lambda_1 \\ \Lambda_2 \\ \Lambda_2 \\ \Lambda_3 \\ \Lambda_4 \end{array} \begin{array}{c} \begin{array}{ccccc} D_z & D_{x_1} & D_{(x,u,t)\setminus\{x_1\}} & D_{\gamma_{1,1}} & D_{\gamma\setminus\{\gamma_{1,1}\}} \end{array} \\ \begin{bmatrix} I_{n+q} & 0_{n+q} & D_{(x,u,t)\setminus\{x_1\}}T(\bar{x},\bar{u},\bar{t}) & 0 & 0 \\ 0 & -D^2_{x_1}T^1_1(\bar{x},\bar{u},\bar{t}) & \bigotimes & 1 & 0^T_{n(n+2q)-1} \\ 0 & \bigotimes & \bigotimes & 0_{n(n+2q)-1} & I_{n(n+2q)-1} \\ I_{n+q} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0^T_{n(n+2q)-1} \end{bmatrix} \end{array}$$

Nach Voraussetzung ist die quadratische Matrix $D_{(x,u,t)\setminus\{x_1\}}T(\bar{x},\bar{u},\bar{t})$ regulär und eine einfache Betrachtung der Matrix liefert den Beweis, daß ihre Regularität äquivalent zur Ungleichung (105) ist. $\square$

Bemerkung 31
Laut Bemerkung 15 ist in den Pukten vom Typ 3 der Variationsungleichungen die erste der drei äquivalenten Aussagen des vorhergehenden Satzes 20 erfüllt. Dabei ist $T = \mathbb{H}_{J_0(x,t)}(x,\lambda,\mu,t)$ mit $u = (\lambda,\mu)$ und speziell

$$T_2(x,t) = \begin{bmatrix} H(x,t) \\ G_{J_0(\bar{x},\bar{t})}(x,t) \end{bmatrix}.$$

Es sei $(x(\theta),u(\theta),t(\theta))$ eine lokale Parametrisierung von $\{\mathbb{H}_{J_0(x,t)} = 0\}$ in einer Umgebung von $(\bar{x},\bar{\lambda},\bar{\mu},\bar{t}) = (x(0),u(0),t(0))$, und wir bezeichnen mit $D_x\mathbb{L}_{J_0}(\theta) = D_x\mathbb{L}_{J_0}(x(\theta),u(\theta),t(\theta))$ und $D_xT_2(\theta) = D_xT_2(x(\theta),t(\theta))$. Dann gilt, daß für $\theta < 0$ und $\theta > 0$ die Vorzeichen von

$$\Psi(\theta) = \det \begin{bmatrix} D_x\mathbb{L}_{J_0}(\theta) & D_x^TT_2(\theta) \\ D_xT_2(\theta) & 0 \end{bmatrix}$$

verschieden sind. Es sei $\bar{W}$ eine $n \times (m + |J_0(\bar{x},\bar{t})|)$ Matrix, deren Spalten eine Basis für den Tangentialraum $T_{\bar{x}}M(\bar{t})$ bilden. Die Matrizen

$$W(\theta) = \bar{W} - D_x^TT_2(\theta)\left[D_x^TT_2(\theta)\cdot D_xT_2(\theta)\right]^{-1}D_xT_2(\theta)\bar{W}$$

bilden dann für jedes θ in einer Umgebung von Null eine Basis für den Tangentialraum $T_{\bar{x}(\theta)}M(t(\theta))$. Aus der Gleichung

$$\begin{bmatrix} W^T{}_{(\theta)} & 0 \\ D_xT_{2(\theta)} & 0 \\ 0 & I_{m+|J_0(\bar{x},\bar{t})|} \end{bmatrix} \begin{bmatrix} D_x\mathbb{L}_{J_0}{}_{(\theta)} & D_x^TT_{2(\theta)} \\ D_xT_{2(\theta)} & 0 \end{bmatrix} \begin{bmatrix} W(\theta) & D_x^TT_{2(\theta)} & 0 \\ 0 & 0 & I_{m+|J_0(\bar{x},\bar{t})|} \end{bmatrix}$$

$$= \begin{bmatrix} W^T{}_{(\theta)}D_x\mathbb{L}_{J_0}{}_{(\theta)}W_{(\theta)} & W^T{}_{(\theta)}D_x\mathbb{L}_{J_0}{}_{(\theta)}D_x^TT_{2(\theta)} & 0 \\ D_xT_{2(\theta)}D_x\mathbb{L}_{J_0}{}_{(\theta)}W_{(\theta)} & D_xT_{2(\theta)}D_x\mathbb{L}_{J_0}{}_{(\theta)}D_x^TT_{2(\theta)} & D_xT_{2(\theta)}D_x^TT_{2(\theta)} \\ 0 & D_xT_{2(\theta)}D_x^TT_{2(\theta)} & 0 \end{bmatrix}$$

ergibt sich die Beziehung

$$\Psi_{(\theta)} \cdot \left(\det \begin{bmatrix} W_{(\theta)} & D_x^TT_{2(\theta)} \end{bmatrix}\right)^2 =$$

$$= \det\left[W^T{}_{(\theta)}D_x\mathbb{L}_{J_0}{}_{(\theta)}D_x^TT_{2(\theta)}\right] \cdot \left(\det\left[D_xT_{2(\theta)}D_x^TT_{2(\theta)}\right]\right)^2 .$$

Damit haben $\Psi(\theta)$ und $\det(D_x\mathbb{L}_{J_0}(x(\theta),u(\theta),t(\theta))|_{T_{\bar{x}(\theta)}M(t(\theta))})$ das gleiche Vorzeichen, was den Vorzeichenwechsel der letzten Zahl entlang der kritischen Kurve Σ_{gc} beweist.

Für den folgenden Störungssatz vereinbaren wir folgende Bezeichnungen. Es sei eine einparametrische Variationsungleichung $(\bar{F},\bar{H},\bar{G})(x,t)$ festgelegt. Wir identifizieren den Raum der $k \times l$ Matrizen mit $I\!R^{kl}$ und betrachten folgende Parameter: $A \in I\!R^{n^2}$ eine quadratische Matrix, $b \in I\!R^n$ ein Spaltenvektor, $C \in I\!R^{(m+p)n}$ eine Matrix mit $m+q$ Zeilen und n Spalten und $d \in I\!R^{m+q}$ ein Spaltenvektor. Da nur A und b für die Störung der Abbildung F, und C und d für die Störung der Restriktionen (H,G) benutzt werden, betrachten wir zwei zusammengefaßte Parameter $\mathcal{A} = (A,b)$ und $\mathcal{C} = (C,d)$. Als gemeinsamen Parameter benutzen wir $\mathcal{P} = (\mathcal{A},\mathcal{C}) \in I\!R^{n^2+n+n(m+p)+m+p}$.

Für einen Parameter $\mathcal{P}$ bezeichnen wir mit

$$(\bar{F},\bar{H},\bar{G},\mathcal{P}) = (\bar{F}(x,t) + Ax + b, (\bar{H}(x,t), G(x,t)) + Cx + d)$$

die mittels $\mathcal{P}$ gestörte Variationsungleichung. Mit $(\bar{H},\bar{G},\mathcal{C})$, $(\bar{H},\mathcal{C})_i$, $i \in I$ und $(\bar{G},\mathcal{C})_j$, $j \in J$ seien die gestörten Restriktionen und ihre jeweiligen Komponenten bezeichnet. Analog bedeutet $(\bar{F},\mathcal{A})$ die Störung der Abbildung $\bar{F}$.

Satz 21
Es sei $(\bar{F}, \bar{H}, \bar{G})(x,t)$ eine feste einparametrische Variationsungleichung. Jede meßbare Teilmenge von

$$\left\{\mathcal{P} \in I\!R^{n^2+n+n(m+p)+m+p} | (\bar{F}, \bar{H}, \bar{G}, \mathcal{P}) \notin \mathcal{V}\right\}$$

hat das Lebesgue-Maß Null.

Der Beweis dieses Ergebnisses ist praktisch identisch mit dem Beweis des Störungssatzes für die Klasse $\mathcal{F}$ (siehe [61]). Deshalb geben wir keinen vollständigen Beweis, sondern nur die Hauptideen des Beweises in [61] mit kleinen Hinweisen darauf, wo für unseren Satz andere Argumente benötigt werden.

Beweis:

Da die Struktur der zulässigen Mengen und auch die ihrer Störung dieselbe ist, ist das Lemma 1 in [61] auch hier anwendbar. Die Überlegungen bei den Punkten vom Typ 4 und 5 (die Hesse-Matrix der Zielfunktion ist ohne Bedeutung) des zweiten Schrittes sind auch direkt zu übernehmen. Bemerkt sei nur, daß für fast alle $\mathcal{A} = (A, b) \in I\!R^{n^2+n}$ (für nicht symmetrische Matrizen A) folgendes gilt:

$$\begin{aligned} F(\bar{x}) + A\bar{x} + b &\notin T \\ (F(\bar{x}) + A\bar{x} + b)^T M (F(\bar{x}) + A\bar{x} + b) &\neq 0, \end{aligned}$$

wobei M eine feste, von Null verschiedene $n \times n-$Matrix, $T \subset I\!R^n$ ein fester linearer echter Teilraum und $\bar{x}$,$F(\bar{x}) \in I\!R^n$ zwei feste Vektoren sind.

Im ersten Schritt betrachtet man zwei feste Indizes $J_1 \subset J_0 \subset \{1, \ldots, p\}$ mit $0 \leq q = m + |J_0| \leq n$ und zwei Zahlen η und s mit $\eta \in \{0, 1\}$ bzw $0 \leq 2(q + \eta - 1) \leq s \leq n + q + \eta - 1$. In den Variablen $(x, \lambda, \mu, t, \gamma_1, \gamma_2)$, wobei $\mu \in I\!R^{J_0}$, $\gamma_1 \in I\!R^n$ und $\gamma_2 \in M_2(n + q + \eta - 1)$, betrachten wir die Mannigfaltigkeit

$$M_4^\eta = I\!R^n \times I\!R^m \times \{\mu \in I\!R^{J_0} | \mu_j = 0, j \in J_1\} \times \{0_{n+q}\} \times M_3(n + q + \eta - 1, s),$$

die die Kodimension $|J_1| + n + q + (n + q + \eta - 1 - s)^2$ besitzt.

Wir legen einen Index in J_0 (z.B. 1) fest und definieren die Abbildungen

$$\mathbb{L}^\eta_{J_0}(\mathcal{P}, x, \lambda, \mu, t) \quad = \quad (\bar{F}, \mathcal{A}) - \sum_{i=1}^{m} \lambda_i D_x^T (\bar{H}, \mathcal{C})_i - \eta \mu_1 D_x^T (\bar{G}, \mathcal{C})_1$$

$$
\begin{aligned}
& & & - \sum_{j \in J_0 \setminus \{1\}} \mu_j D_x^T (\bar{G}, \mathcal{C})_j \\
\mathbb{H}^{\eta}_{J_0}(\mathcal{P},x,\lambda,\mu,t) & = & & \begin{bmatrix} \mathbb{L}^{\eta}_{J_0}(\mathcal{P},x,\lambda,\mu,t) \\ (\bar{H},\mathcal{C}) \\ (1-\eta)\mu_1 + (\bar{G},\mathcal{C})_1 \\ (\bar{G},\mathcal{C})_{J_0\setminus\{1\}} \end{bmatrix} \\
\mathbb{M}^{1}(\mathcal{P},x,\lambda,\mu,t) & = & & D_{(x,\lambda,\mu)}\mathbb{H}^{1}_{J_0}(\mathcal{P},x,\lambda,\mu,t) \in M_1(n+q) \\
\mathbb{M}^{0}(\mathcal{P},x,\lambda,\mu,t) & = & & \begin{bmatrix} D_x\mathbb{L}^{1}_{J_0} & D_\lambda \mathbb{L}^{0}_{J_0} & D_{\mu\setminus\{\mu_1\}}\mathbb{L}^{0}_{J_0} \\ D_x(\bar{H},\mathcal{C}) & 0 & 0 \\ D_x(\bar{G},\mathcal{C})_{J_0\setminus\{1\}} & 0 & 0 \end{bmatrix} \in M_1(n+q-1)
\end{aligned}
$$

Man kann zeigen, daß

$$
Graf(\mathbb{H}^{\eta}_{J_0}, \mathbb{M}^{\eta}) \overset{\top}{\pitchfork} I\!R^{n^2+n+n(m+p)+m+p} \times M_4^{\eta}
$$

für alle möglichen Wahlen von J_1, J_0, η, s und des festgelegten Index. Nach dem parametrisierten Satz von Sard folgt dann für fast alle feste $\mathcal{P}$, daß für jede solche Selektion die Bedingung

$$
Graf(\mathbb{H}^{\eta}_{J_0}, \mathbb{M}^{\eta}) \overset{\top}{\pitchfork} M_4^{\eta} \tag{106}
$$

auch gilt, wobei jetzt $Graf(\mathbb{H}^{\eta}_{J_0}, \mathbb{M}^{\eta}) \subset I\!R^{n+q+1+n+q+n(n+2q)}$.

Die Kodimension von $Graf(\mathbb{H}^{\eta}_{J_0}, \mathbb{M}^{\eta})$ ist $n+q+n(n+2q)$. Falls (106) sich nicht auf die leere Menge reduziert, gilt dann $|J_1| + (n+q+\eta-1-s)^2 \leq 1$.

Für jeden verallgemeinerten kritschen Punkt $(\bar{x}, \bar{t})$, von $(\bar{F}, \bar{H}, \bar{G}, \mathcal{P})$ ($\mathcal{P}$ ist schon fest und (106) gilt), in dem die LICQ erfüllt ist, existieren eine aktive Indexmenge $J_0 = J_0(\bar{x}, \bar{t})$, Multiplikatoren $(\bar{\lambda}, \bar{\mu})$ und $J_1 = \{j \in J_0 | \bar{\mu}_j = 0\}$, so daß die Bedingung (106) für $\eta = 1$ erfüllt ist.

Falls $|J_1| + (n+q+-s)^2 = 0$, sind $D_{(x,\lambda,\mu)}\mathbb{H}^{1}_{J_0}(\mathcal{P}, \bar{x}, \bar{\lambda}, \bar{\mu}, \bar{t})$ regulär (**VU-ND2**) und $J_1 = \emptyset$ (**VU-ND1**). Damit ist der Punkt vom Typ 1.

Falls $|J_1| = 0$ und $s = n+q-1$ sind, gelten auch $J_1 = \emptyset$ (**VU-ND1**) und nach Definition von $M_3(n+q, n+q-1)$ auch **Typ3: 3-VU**. Aus der Bedingung (106) folgt zunächst, daß die Matrix $D_{(x,\lambda,\mu,t)}\mathbb{H}^{1}_{J_0}(\mathcal{P}, \bar{x}, \bar{\lambda}, \bar{\mu}, \bar{t})$ vollen Rang (gleich $n+q$) hat. Nach Satz 20 und der Bemerkung 15 ist die Bedingung **Typ3: 4-VU** erfüllt, und damit ist der Punkt vom Typ 3.

Falls $|J_1| = 1$ und $(n+q+-s)^2 = 0$ sind, wird die zweite Bedingung der Punkte vom Typ 2 erfüllt (ein einziger Multiplikator z.B. $\bar{\mu}_1$ ist Null). Aus

(106) folgt also **Typ2: 2-VU**. Andererseits gilt diese Bedingung in $(\bar{x}, \bar{\lambda}, \bar{\mu}, \bar{t})$ für $\eta = 1$ und $s = n+q-1$, wobei J_1 und J_0 dieselben Indexmengen bleiben. Daraus folgen **Typ2: 3-VU** und **Typ2: 4-VU**.

Als Folgerung der beiden vohergehenden Sätze 19 und 21 kann man jetzt folgende wichtige Aussage über die regulären Varationsungleichungen herleiten:

Satz 22
Die Menge $\mathcal{V}$ ist offen und dicht in $C^3(\mathbb{R}^{n+1}, \mathbb{R}^{n+m+p})$, versehen mit der C^3-starken Topologie.

Beweis:
Der Beweis wird mit Hilfe der Zerlegung der Einheit (siehe z.B. [2], [31]) aus den Sätzen 19 und 21 hergeleitet. Die dabei benutzte Konstruktion ist die sogenannte lokal→global Prozedur, üblich in der Differential Topologie (siehe z.B. [2], [31]). Diese Konstruktion ist für den Beweis der Generizität der Klasse $\mathcal{F}$ auch in [13] zu finden. □

Da die geometrische Struktur und die wichtigsten Eigenschaften der Singularitäten für Variationsungleichungen erhalten geblieben sind, gelten solche Sätze über die globale Struktur der kritischen Kurven, die auf diesen Eigenschaften basieren, auch in diesem Fall. Damit ergeben sich die folgenden Bemerkungen:

Bemerkung 32 *(siehe auch Satz 6)*
Es sei $(F, H, G) \in \mathcal{V}$ eine reguläre einparametrische Variationsungleichung. Aus den Bemerkungen 11 und 12 folgt die Offenheit der Menge Σ^1_{gc} in Σ_{gc}. Andererseits folgt aus der Bemerkung 13 (für Typ 2), der Bemerkung 16 (für Typ 3), den Bemerkungen 19 und 20, der Behauptung 7 (für Typ 4) und den Bemerkungen 23 und 24 (für Typ 5), daß alle Punkte in $\Sigma_{gc} \setminus \Sigma^1_{gc}$ isoliert und Randpunkte von Σ^1_{gc} sind. Damit gilt:

- *Σ^1_{gc} ist offen und dicht in Σ_{gc},*
- *$\Sigma_{gc} \setminus \Sigma^1_{gc} = \cup_{i=2}^5 \Sigma^i_{gc}$ ist diskret.*

Bemerkung 33 *(siehe auch Satz 7)*
Es sei $(F, H, G) \in \mathcal{V}$ eine reguläre einparametrische Variationsungleichung.

Aus der Bemerkung 12 folgt die Offenheit der Menge Σ^1_{stat} in Σ_{stat}. Aus der Änderung der linearen Indizes in einer Umgebung der Punkte vom Typ 2 und 3, geschildert in den Bildern 10 und 11, folgt, daß solche Punkte keine Randpunkte von Σ_{stat} sind. Nach Definition sind die Punkte vom Typ 4 nicht in Σ_{stat}. Ist aber ein solcher Punkt $(x,t) \in \Sigma^4_{gc}$ in $\overline{\Sigma_{stat}}$, dann ist die MFCQ verletzt und $\overline{\Sigma_{stat}}$ ist in einer Umgebung von (x,t) eine eindimensionale Mannigfaltigkeit. Ferner gilt in diesem Fall, daß (x,t) genau dann ein Randpunkt von $\overline{\Sigma_{stat}}$ ist, wenn $J_0(x,t) \neq \emptyset$. Andererseits folgt aus der Bemerkung 25, daß die Punkte vom Typ 5 in Σ_{stat} genau dann keine Randpunkte von Σ_{stat} sind, wenn die MFCQ erfüllt ist. Aus der Struktur vom Σ_{stat} in einer Umgebung der Singularitäten folgt damit:

$\overline{\Sigma_{stat}}$ ist eine stückweise, zweimal stetig differenzierbare Mannigfaltigkeit mit Rand, und ein Punkt ist genau dann ein Randpunkt, wenn die MFCQ verletzt ist und $J_0 \neq \emptyset$ ist.

Bemerkung 34
Der Satz 8 gilt auch mit $(F,H,G) \in \mathcal{V}$. Der Beweis ist derselbe, denn die Definition der Punkte vom Typ 4 und 5 ist praktisch für Optimierungsprobleme und Variationsungleichungen dieselbe, und die Gültigkeit der MFCQ in einem Punkt hängt nicht von der Zielfunktion oder der Abbidung F ab.

5.5 Anwendung bei der Konvergenz von Kurvenverfolgungsprozeduren

Als Abschluß betrachten wir einige Aussagen, die für die Konvergenz von Kurvenverfolgungsprozeduren für reguläre Variationsungleichungen nützlich sind.

Bemerkung 35
Es seien $(\bar{x},\bar{t})$ ein Punkt vom Typ 2 oder 5 und zwei Kurven $x_1(t)$, $x_2(t)$ in einer Umgebung von $\bar{t}$ definiert, so daß

$$\{(x_i(t),t) | t < \bar{t}\} \subset \Sigma^1_{gc}, \ i = 1,2$$
$$(x_i(t),t) \to (\bar{x},\bar{t}), \ \text{für } t \nearrow \bar{t} \ .$$

Laut Bemerkungen 13 und 24 gilt $(-\dot{x}_1,-1) \neq (-\dot{x}_2,-1)$. Damit gibt es zwei offene Kugeln von Radius r_1 (bzw r_2) $B(-\dot{x}_1,r_1)$ (bzw. $B(-\dot{x}_2,r_2)$)

mit $\overline{B(-\dot{x}_1, r_1)} \cap \overline{B(-\dot{x}_2, r_2)} = \emptyset$. *Daraus folgt, daß der von den Mengen* $B(-\dot{x}_1, r_1) \times \{-1\}$ *(bzw.* $B(-\dot{x}_2, r_2) \times \{-1\}$*) erzeugte Kegel* K_1 *(bzw.* K_2*), ohne den Nullvektor genommen, auch offen ist, und* $\overline{K_1} \cap \overline{K_2} = \{0\}$. *Andererseits existiert ein* $\tilde{t} < \bar{t}$, *so daß* $x_i(t) \in (\bar{x}, \bar{t}) - K_i$, $\forall t \in (\tilde{t}, \bar{t})$, $i = 1, 2$. *Für* $\tilde{t}$ *nahe genug bei* $\bar{t}$ *gelten für die offenen Mengen*

$$U_i = ((\bar{x}, \bar{t}) - K_i) \cap I\!R^n \times (\tilde{t}, \bar{t}), \ i = 1, 2$$

folgende Eigenschaften (jeweils $i = 1, 2$*).*

$$\begin{aligned} \{(x_i(t), t) | t \in (\tilde{t}, \bar{t})\} &\subset U_i \\ (\Sigma_{gc} \setminus \{(x_i(t), t) | t \in [\tilde{t}, \bar{t})\}) \cap \overline{U_i} &= \{(\bar{x}, \bar{t})\} \notin \Sigma^1_{gc} \end{aligned} \tag{107}$$

Dieselbe Überlegung kann bei mehreren Kurven im Fall der Punkte vom Typ 5 durchgeführt werden, und man findet für jede Kurve $x(t)$ *mit den Eigenschaften von* $x_1(t)$ *eine offene Menge* U *und eine Zahl* $\tilde{t}$ *so daß die Bedingungen 107 bzgl.* $x(t)$, U *und* $\tilde{t}$ *erfüllt sind.*

Bemerkung 36
Es sei $(\bar{x}, \bar{t})$ *ein Punkt vom Typ 3 oder 4. Laut Bemerkungen 15, 16, 19 und 20 kann man die Kurve* Σ_{gc} *in einer Umgebung von* $(\bar{x}, \bar{t})$ *nach einer der Komponenten von* x *(z.B.* x_1*) parametrisieren. Bezeichnen wir mit* y *den Vektor* $y = (x_2, \ldots, x_n)$, *dann existieren positive Zahlen* r_{x_1}, r_y *und* r_t *und eine differenzierbare Abbildung* $(y(x_1), t(x_1))$, *so daß* Σ_{gc} *in der offenen Umgebung* $(\bar{x}_1 - r_{x_1}, \bar{x}_1 + r_{x_1}) \times B(\bar{y}, r_y) \times B(\bar{t}, r_t)$ *die Menge*

$$\{(x_1, y(x_1), t(x_1) | x_1 \in (\bar{x}_1 - r_{x_1}, \bar{x}_1 + r_{x_1})\}$$

darstellt. Dabei seien $(\bar{x}_1, y(\bar{x}_1), t(\bar{x}_1)) = (\bar{x}, \bar{t})$ *und* $t(x_1) < \bar{t}$, $\forall x_1 \neq \bar{x}_1$. *Wir bezeichnen* $(x(x_1), t(x_1)) = (x_1, y(x_1), t(x_1))$.

Man kann die Zahlen r_{x_1}, r_y *und* r_t *auch so klein wählen, daß der Durchschnitt des Randes von* $(\bar{x}_1 - r_{x_1}, \bar{x}_1 + r_{x_1}) \times B(\bar{y}, r_y) \times B(\bar{t}, r_t)$ *mit* Σ_{gc} *aus den Punkten* $\{(x(\bar{x}_1 - r_{x_1}), t(\bar{x}_1 - r_{x_1})), (x(\bar{x}_1 + r_{x_1}), t(\bar{x}_1 + r_{x_1}))\}$ *besteht. Damit gelten für die offenen Mengen* $U^- = (\bar{x}_1 - r_{x_1}, \bar{x}_1) \times B(\bar{y}, r_y) \times B(\bar{t}, r_t)$, $U^+ = (\bar{x}_1, \bar{x}_1 + r_{x_1}) \times B(\bar{y}, r_y) \times B(\bar{t}, r_t)$ *folgende Beziehungen.*

$$\begin{aligned} \{(x(x_1), t(x_1)) | x_1 \in (\bar{x}_1 - r_{x_1}, \bar{x}_1)\} &\subset U^- \\ (\Sigma_{gc} \setminus \{(x(x_1), t(x_1)) | x_1 \in [\bar{x}_1 - r_{x_1}, \bar{x}_1)\}) \cap \overline{U^-} &= \{(\bar{x}, \bar{t})\} \notin \Sigma^1_{gc} \\ \{(x(x_1), t(x_1)) | x_1 \in (\bar{x}_1, \bar{x}_1 + r_{x_1}]\} &\subset U^+ \\ (\Sigma_{gc} \setminus \{(x(x_1), t(x_1)) | x_1 \in (\bar{x}_1, \bar{x}_1 + r_{x_1}]) \cap \overline{U^+} &= \{(\bar{x}, \bar{t})\} \notin \Sigma^1_{gc} \end{aligned}$$

Nach den Bemerkungen 15 und 18 ist $\bar{x}_1$ die einzige Nullstelle von $\dot{t}(x_1)$ in $(\bar{x}_1 - r_{x_1}, \bar{x}_1 + r_{x_1})$ (die nichtdegenierten kritischen Punkte sind isoliert !). Daraus folgt, daß $\dot{t}(x_1)$ auf jedem der Intervalle $(\bar{x}_1 - r_{x_1}, \bar{x}_1)$ und $(\bar{x}_1, \bar{x}_1 + r_{x_1})$ dasselbe Vorzeichen besitzt. Da $(t(x_1) \leq \bar{t} = t(\bar{x}_1))$ gilt, ist $\dot{t}(x_1)$ für $(\bar{x}_1 - r_{x_1}, \bar{x}_1)$ positiv und für $(\bar{x}_1, \bar{x}_1 + r_{x_1})$ negativ. Damit kann man zwei Intervalle $[\tilde{t}^-, \bar{t}]$,$[\tilde{t}^+, \bar{t}]$ und zwei stetige, in Innern ihres Definitionbereiches differenzierbare Funktionen

$$\tilde{x}_1^- : [\tilde{t}^-, \bar{t}] \to [\bar{x}_1 - r_{x_1}, \bar{x}_1]$$
$$\tilde{x}_1^+ : [\tilde{t}^+, \bar{t}] \to [\bar{x}_1, \bar{x}_1 + r_{x_1}]$$

definieren, die inverse Funktionen von $t(x_1)$ sind. Es sind $\tilde{x}_1^-$ (bzw. $\tilde{x}_1^+$) strikt monoton wachsend (fallend) und $\tilde{x}_1^-(\tilde{t}^-) = \bar{x}_1 - r_{x_1}$, $\tilde{x}_1^-(\tilde{t}^+) = \bar{x}_1$ (bzw. $\tilde{x}_1^+(\tilde{t}^+) = \bar{x}_1$, $\tilde{x}_1^-(\tilde{t}^+) = \bar{x}_1$). Wir bezeichnen die stetigen Abbildungen

$$\tilde{x}^-(t) = x(\tilde{x}_1^-(t)) : (\tilde{t}^-, \bar{t}) \to (\bar{x}_1 - r_{x_1}, \bar{x}_1) \times B(\bar{y}, r_y)$$
$$\tilde{x}^+(t) = x(\tilde{x}^+(t)) : (\tilde{t}^+, \bar{t}) \to (\bar{x}_1, \bar{x}_1 + r_{x_1}) \times B(\bar{y}, r_y).$$

Die Menge Σ_{gc}^1 zerfällt dann in einer Umgebung von $(\bar{x}, \bar{t})$ in die folgenden zwei Zusammenhangskomponenten.

$$\{(\tilde{x}^-(t), t) | t \in (\tilde{t}^-, \bar{t})\} \subset U^-$$
$$\{(\tilde{x}^+(t), t) | t \in (\tilde{t}^+, \bar{t})\} \subset U^+$$

Dabei gilt auch

$$(\Sigma_{gc} \setminus \{(\tilde{x}^-(t), t) | t \in (\tilde{t}^-, \bar{t})\}) \cap \overline{U^-} = \{(\bar{x}, \bar{t})\} \notin \Sigma_{gc}^1$$
$$(\Sigma_{gc} \setminus \{(\tilde{x}^+(t), t) | t \in (\tilde{t}^+, \bar{t})\}) \cap \overline{U^+} = \{(\bar{x}, \bar{t})\} \notin \Sigma_{gc}^1.$$

Lemma 7

Es seien $(F, H, G) \in \mathcal{V}$ eine reguläre einparametrische Variationsungleichung und S eine Zusammenhangskomponente von Σ_{gc}^1. Dann existieren ein offenes Intervall $I_S = (a, b)$ (mit $a \in I\!R \cup \{-\infty\}$ und $b \in I\!R \cup \{+\infty\}$) und eine zweimal stetig differenzierbare Funktion $x_S : I \to I\!R^n$ mit folgenden Eigenschaften:

1. $S = \{(x, t) \in I\!R^n \times I_S | x_S(t) = x\}$.

2. *Ist a (bzw. b) endlich, dann tritt nur einer der folgenden beiden Fälle auf:*

   - $\lim_{t\searrow a}(x_S(t),t)$ *(bzw.* $\lim_{t\nearrow b}(x_S(t),t)$*) konvergiert gegen einen Punkt in* $\Sigma_{gc}\setminus\Sigma^1_{gc}$.
   - $\lim_{t\searrow a}\|x_S(t)\|$ *(bzw.* $\lim_{t\nearrow b}\|x_S(t)\|$*) konvergiert gegen* $+\infty$.

Beweis:

Es sei $(x_0,t_0)\in S$ fest. Nach Bemerkung 12 existieren ein offenes Intervall $I_0=(\bar t-\varepsilon_0,\bar t+\varepsilon_0)$, eine offene Kugel $B(x_0,r_0)=\{x|\,\|x-x_0\|<r_0\}\subset I\!R^n$ (wobei $0<r_0,\varepsilon_0<1$) und eine zweimal stetig differenzierbare Abbildung $x_{I_0}:I_0\to I\!R^n$ mit folgenden Eigenschaften.

$$\begin{aligned} G(x_{I_0}) &\subset B(x_0,r_0)\times I_0 \\ \Sigma_{gc}\cap B(x_0,r_0)\times I_0 &= S\cap B(x_0,r_0)\times I_0=G(x_{I_0}) \end{aligned} \tag{108}$$

Dabei bedeutet $G(x_I)=\{(x,t)\in I\!R^n\times I_0|x_I(t)=x\}$ für Abbildungen $x_I: I\subset I\!R\to I\!R^n$. Man kann die Konstanten r_0 und ε_0 der offenen Mengen I_0 und $B(x_0,r_0)$ so klein wählen, daß die Abschließung $\overline{B(x_0,r_0)\times I_0}=\overline{B(x_0,r_0)}\times\overline{I_0}$ keinen Punkt in $\Sigma_{gc}\setminus G(x_{\overline{I_0}})$ besitzt.

Weiter seien I ein offenes Intervall und x^1_I, x^2_I zwei stetige Abbildungen mit

$$\begin{aligned} x^1_I,x^2_I &: I\to I\!R^n \\ G(x_I) &\subset \Sigma^1_{gc} \quad . \end{aligned}$$

Die Menge

$$\{t\in I|x^1_I(t)=x^2_I(t)\} \tag{109}$$

ist wegen der eben für (x_0,t_0) genannten lokalen Eigenschaften der Punkte vom Typ 1 und der Stetigkeit beider Abbildungen offen. Andererseits ist die Menge

$$\{t\in I|x^1_I(t)\neq x^2_I(t)\} \tag{110}$$

wegen der Stetigkeit beider Abbildungen ebenfalls offen. Da I zusammenhängend ist, muß eine der Mengen (110) oder (109) leer sein, und die andere mit I übereinstimmen.

Wir wollen jetzt die Menge $\mathcal{I}$ derjenigen offenen Intervalle I mit:

$$\begin{aligned} I_0 &\subset I \\ \exists x_I &: I \to I\!R^n \text{ stetig, mit } x_I(t) = x_{I_0}(t)\ \forall t \in I_0 \\ G(x_I) &\subset \Sigma^1_{gc} \end{aligned} \tag{111}$$

betrachten.

Die Abbildung x_I ist für jedes Intervall $I \in \mathcal{I}$ eindeutig bestimmt und zweimal stetig differenzierbar. Für zwei Intervalle $I_1, I_2 \in \mathcal{I}$ ergibt sich, daß $x_{I_1}(t) = x_{I_2}(t)$ für alle t im offenen Intervall $I_1 \cap I_2 \supset I_0$.

Es sei I_S die Vereinigung aller Intervalle in $\mathcal{I}$. Leicht zu sehen ist, daß I_S ebenfalls ein offenes Intervall mit eventuell unendlichen Grenzen darstellt. Für $t \in I_S$ definieren wir $x_S(t) = x_{I_t}(t)$ für ein $I_t \in \mathcal{I}$ mit $t \in I_t$. Wie vorher gesehen, ist x_S wohldefiniert auf I_S. Diese Abbildung x_S ist auch zweimal stetig differenzierbar, da sie mit x_{I_t} in einer Umgebung von t übereinstimmt. Daraus folgt $I_S \in \mathcal{I}$.

Somit gilt, daß $G(x_{I_S})$ eine zusammenhängende Teilmenge von Σ^1_{gc} ist. Da $G(x_{I_S}) \cap S \neq \emptyset$, ist $G(x_{I_S}) \subset S$. Für jeden Punkt (x,t) in $G(x_{I_S})$ kann man, wie für (x_0, t_0) gemacht, zwei offene Mengen $B(x, r_{(x,t)})$ und $I_{(x,t)}$ wählen (die Abbildung $x_{I_{(x,t)}}$ stimmt mit x_{I_S} in $I_{(x,t)}$ überein), so daß

$$\begin{aligned} G(x_{I_S}) \cap I\!R^n \times I_{(x,t)} &= G(x_{I_S}) \cap B(x, r_{(x,t)}) \times I_{(x,t)} \\ \overline{B(x, r_{(x,t)})} \times \overline{I_{(x,t)}} \cap (\Sigma_{gc} \setminus G(x_{I_S})) &= \emptyset \quad . \end{aligned}$$

Das liefert eine offene Überdeckung

$$\mathcal{U} = \{B(x, r_{(x,t)}) \times I_{(x,t)} | (x,t) \in G(x_{I_S})\}$$

von $G(x_{I_S})$.

Teilen wir $G(x_{I_S})$ in $G(x_{I_S})^+ = \{(x,t) \in G(x_{I_S}) | t \geq t_0\}$ (bzw. $G(x_{I_S})^-$ mit $t \leq t_0$). Wir zeigen, daß eine abgeschlossene Menge existiert, die $G(x_{I_S})^+$ enthält und $S \setminus G(x_{I_S})$ nicht schneidet. Mit derselben Überlegung gilt dasselbe für $G(x_{I_S})^-$ und damit gilt $G(x_{I_S}) = S$. Damit wäre die erste Behauptung bewiesen. Die zweite ergibt sich aus den folgenden Überlegungen bezüglich $G(x_{I_S})^+$ (und analog bzgl. $G(x_{I_S})^-$).

Bezeichnen wir mit $I_S^+ = \{t \in I_S | t \geq t_0\}$ und $\mathcal{U}^+ = \{B(x, r_{(x,t)}) \times I_{(x,t)} \in \mathcal{U} | I_{(x,t)} \cap I_S^+ \neq \emptyset\}$. Falls $I_S^+ = [t_0, +\infty)$ ist, kann man o.B.d.A.

(eine solche Überdeckung kann aus $\mathcal{U}^+$ konstruiert werden) annehmen, daß für alle $t \in I_S^+$ eine Umgebung existiert, die nur endlich viele Mengen aus $\{I_{(x,t)} | B(x, r_{(x,t)}) \times I_{(x,t)} \in \mathcal{U}^+\}$ schneidet. Damit ist die Menge

$$\bigcup \{\overline{B(x, r_{(x,t)})} \times \overline{I_{(x,t)}} | B(x, r_{(x,t)}) \times I_{(x,t)} \in \mathcal{U}^+\} \tag{112}$$

abgeschlossen. Da (112) keinen Punkt in $S \setminus G(x_{I_S})$ enthält, sind wir in diesem Fall fertig.

Es sei dann $I_S^+ = [t_0, t_0^+)$ mit $t_0 < t_0^+ < +\infty$. Betrachten wir eine Folge $\{\zeta_l\}$ mit

$$\begin{aligned} \zeta_l &\nearrow t_0^+, \text{ für } l \to +\infty \\ (x_{I_S}(\zeta_l), \zeta_l) &\to (y, t_0^+). \end{aligned} \tag{113}$$

Da Σ_{gc} abgeschlossen ist, gilt $(y, t_0^+) \in \Sigma_{gc} = \bigcup_{i=1}^5 \Sigma_{gc}^i$. Laut Bemerkungen 35 und 36 und wegen der Struktur in einer Umgebung der Punkte vom Typ 1 gibt es ein offenes Intervall (t_0^-, t_0^+), eine offene Umgebung V_y von y, und endlich viele Abbildungen $x^k_{(t_0^-, t_0^+)}$, $k = 1, ..., \kappa$ $(\leq 1\kappa \leq n+1)$ mit

$$\begin{aligned} x^k_{(t_0^-, t_0^+)} &: (t_0^-, t_0^+) \to V_y, \ k = 1, ..., \kappa \\ x^{k_1}_{(t_0^-, t_0^+)}(t) &\neq x^{k_2}_{(t_0^-, t_0^+)}(t), \ \forall t \in (t_0^-, t_0^+), k_1 \neq k_2 \\ \Sigma_{gc}^1 \cap V_y \times (t_0^-, t_0^+) &= \bigcup_{k=1}^{\kappa} G(x^k_{(t_0^-, t_0^+)}) \quad . \end{aligned}$$

Aus (113) und $G(x_{I_S}) \subset \Sigma_{gc}^1$ folgt die Existenz einer Abbildung (z.B. $x^1_{(t_0^-, t_0^+)}$) mit

$$G(x_{I_S}) \cap G(x^1_{(t_0^-, t_0^+)}) \neq \emptyset.$$

Daraus ergibt sich dann

$$G(x_{I_S}) \cap I\!R^n \times (t_0^-, t_0^+) = G(x^1_{(t_0^-, t_0^+)}).$$

Nach Definition von I_S folgt aus dieser Beziehung, daß $y \notin \Sigma_{gc}^1$. Ebenso folgt, daß höchstens ein Häufungspunkt wie (y, t_0^+) existieren kann. Daher existieren nur die zwei folgenden Fälle:

1. $\lim_{t \nearrow t_0^+} x_{I_S}(t) = (y, t_0^+) \in \bigcup_{i=2}^5 \Sigma_{gc}^i$. Entsprechend den Bemerkungen 35 und 36 finden wir ein offenes Intervall (t_0^-, t_0^+) und eine offene Menge $U \subset I\!R^{n+1}$, so daß

$$G(x_{I_S}) \cap I\!R^n \times (t_0^-, t_0^+) \subset U$$
$$(\Sigma_{gc} \setminus G(x_{I_S})) \cap \overline{U} = \{(y, t_0^+)\} \notin \Sigma_{gc}^1 \quad .$$

Andererseits können wir wegen der Kompaktheit von $[t_0, t_0^-]$ eine endliche Familie von offenen Mengen $\mathcal{W}$ aus $\mathcal{U}^+$ auswählen, die $G(x_{I_S}) \cap I\!R^n \times [t_0, t_0^-]$ überdeckt. Damit ist $G(x_{I_S})^+$ in $\mathcal{W} \cup U$ enthalten, und

$$\overline{\mathcal{W} \cup U} = \bigcup \{\overline{B(x, r_{(x,t)})} \times \overline{I_{(x,t)}} | B(x, r_{(x,t)}) \times I_{(x,t)} \in \mathcal{W}\} \cup \overline{U}$$

abgeschlossen. Ebenso gilt

$$\overline{\mathcal{W} \cup U} \cap (S \setminus G(x_{I_S})) = \emptyset,$$

und wir sind fertig.

2. $\lim_{t \nearrow t_0^+} \|x_{I_S}(t)\| = +\infty$. Da $G(x_{I_S})^+$ keinen Häufungspunkt in $I\!R^n \times \{t_0^+\}$ besitzt und für alle $\varepsilon > 0$ die Menge $G(x_{I_S})^+ \cap I\!R^n \times [t_0, t_0^+ - \varepsilon]$ kompakt ist, enthält $G(x_{I_S})^+$ alle ihre Häufungspunkte (d.h. $G(x_{I_S})^+$ ist abgeschlossen). Betrachten wir die kompakten Mengen

$$Z_j = \{x \in I\!R^n | j - 1 \leq \|x\| \leq j\} \times [t_0, t_0^+], \quad j = 1, 2, ... \quad .$$

Für jedes j wählen wir aus $\mathcal{U}^+$ eine endliche Überdeckung $\mathcal{W}_j$ der kompakten Menge $Z_j \cap G(x_{I_S})^+$ und bezeichnen mit

$$\mathcal{W} = \bigcup_{j=1}^{\infty} \mathcal{W}_j$$

die aus ihrer Vereinigung entstandene Überdeckung von $G(x_{I_S})^+$. Da $r_{(x_k,t_k)} \leq 1$ für alle $B(x, r_{(x,t)}) \times I_{(x,t)} \in \mathcal{W}$, gibt es für alle $x \in I\!R^{n+1}$ eine Umgebung, die nur endlich viele Elemente aus $\mathcal{W}$ schneidet. Damit ist die Menge

$$\overline{\mathcal{W}} = \bigcup \{\overline{B(x, r_{(x,t)})} \times \overline{I_{(x,t)}} | B(x, r_{(x,t)}) \times I_{(x,t)} \in \mathcal{U}_1^+\}$$

abgeschlossen und

$$\overline{\mathcal{W}} \cap (S \setminus G(x_{I_S})) = \emptyset.$$

Mit dieser Konstruktion ist der Beweis beendet. □

Bemerkung 37
Es sei $(\bar{x}, \bar{t}) \in \bigcup_{j=2}^{5} \Sigma_{gc}$. *Laut Bemerkungen 13, 24 und 36 gibt es endlich viele zweimal stetig differenzierbare Abbildungen* $x_k(t)$, $k \in \{1, 2, 3, ..., \bar{k}\}$ *mit* $(2 \leq \bar{k} \leq n+1)$ *und zwei Intervalle* $(\bar{t} - \varepsilon, \bar{t})$, $(\bar{t}, \bar{t} + \varepsilon)$, *wobei jede der Abbildungen* $x_k(t)$ *für eines der beiden Intervalle definiert ist, so daß in einer Umgebung von* $(\bar{x}, \bar{t})$ *die Beziehung*

$$\Sigma_{gc}^1 = \bigcup_{k=1}^{\bar{k}} G(x_k)$$

gilt. Also gilt für alle $k_1 \neq k_2$ $G(x_{k_1}) \cap G(x_{k_2}) = \emptyset$. *Entsprechend dem vorhergehenden Lemma 7 gehört jede der Mengen* $G(x_k)$ *zu verschiedenen Zusammenhangskomponenten von* Σ_{gc}^1.

Jetzt kann man Konvergenzaussagen ähnlich Satz 15 aufbauen.

Satz 23
Es sei $(F, H, G) \in \mathcal{V}$ *eine reguläre einparametrische Variationsungleichung mit folgenden zusätzlichen Eigenschaften:*

C1 *Es existiert eine kompakte Menge* $C \subset \mathbb{R}^n$, *so daß*

$$M(t) = \{x \in \mathbb{R}^n | H(x,t) = 0, G(x,t) \geq 0\} \subset C$$

für alle $t \in [0,1]$ *gilt.*

C2 *Die Variationsungleichung* $(F(.,0), H(.,0), G(.,0))$ *besitzt einen einzigen stationären Punkt* $(x_0, 0)$, *der nichtdegeneriert ist (d.h.* $(x_0, 0) \in \Sigma_{gc}^1$*).*

C3 *Für alle* $t \in [0,1]$ *und alle zulässigen Punkte in* $M(t)$ *gilt die MFCQ.*

Dann existiert eine stückweise zweimal stetig differenzierbare Kurve in $\mathbb{R}^n \times [0,1]$, *die den Punkt* $(x_0, 0)$ *mit einem Punkt* $(x_1, 1)$ *verbindet, so daß für alle Paare* (x,t) *auf der Kurve der Punkt* x *ein stationärer Punkt der Variationsungleichung* $P(t)$ *ist.*

Beweis:

Da die linearen Indizes um die Punkte vom Typ 1 unverändert bleiben, sind alle Zusammenhangskomponenten von Σ^1_{stat} auch Zusammenhangskomponenten von Σ^1_{gc}. Bezeichnen wir mit S_0 die Zusammenhangskomponente von Σ^1_{stat}, die den Punkt $(x_0, 0)$ enthält. Laut Lemma 7 existieren ein Intervall (t_0, t_1) (mit $t_0 < 0 < t_1$) und eine zweimal stetig differenzierbare Abbildung $y_0(t) : (t_0, t_1) \to I\!R^n$ mit

$$S_0 = G(y_0(t)) = \{(x, t) \in I\!R^n \times (t_0, t_1) | x = y_0(t)\}.$$

Falls $t_1 > 1$ gilt, sind wir fertig. Da $S_0 \cap I\!R^n \times [0, 1] \subset C$, folgt aus dem gleichen Lemma 7 die Existenz eines Punktes $(x_1, t_1) \in \Sigma_{gc} \setminus \Sigma^1_{gc}$ mit $\lim_{t \to t_1} y_0(t) = x_1$. Es gilt $(x_1, t_1) \in \overline{\Sigma_{stat}}$, und wegen Bedingung **C3** gilt auch $(x_1, t_1) \in \Sigma_{stat}$. Damit sind wir fertig, falls $t_1 = 1$. Da $t_0 < 0$ gilt, folgt auch

$$(\overline{S_0} \setminus \Sigma^1_{gc}) \cap I\!R^n \times [0, 1] = \{(x_1, t_1)\}. \tag{114}$$

Da die Punkte vom Typ 4 keine stationären Punkte sind, gibt es nur drei Möglichkeiten für (x_1, t_1):

1. $(x_1, t_1) \in \Sigma^2_{stat}$. Wegen der Änderung der linearen Indizes, dargestellt im Bild 10 und der Bemerkung 37, gibt es eine und nur eine von S_0 verschiedene Zusammenhangskomponente S_1 von Σ^1_{stat}, die eine Umgebung von (x_1, t_1) schneidet.

2. $(x_1, t_1) \in \Sigma^3_{stat}$. Nach Bemerkungen 35 und 37 existiert eine und nur eine von S_0 verschiedene Zusammenhangskomponente S_1 von Σ^1_{gc}, die eine Umgebung von (x_1, t_1) schneidet. Da in einer Umgebung von (x_1, t_1) die linearen Indizes unverändert bleiben (siehe Bild 11), gilt auch $S_1 \subset \Sigma^1_{stat}$.

3. $(x_1, t_1) \in \Sigma^5_{stat}$. Entsprechend Bemerkung 25 gibt es eine und nur eine von S_0 verschiedene Zusammenhangskomponente S_1 von Σ^1_{stat}, die eine Umgebung von (x_1, t_1) schneidet.

Nach Lemma 7 existieren ein $t_2 \neq t_1$ und eine zweimal stetig differenzierbare Abbildung $y_1(t)$, definiert für das Intervall $(\min(t_1, t_2), \max(t_1, t_2))$ mit $\lim_{t \to t_1} y_1(t) = x_1$. Aus der Bedingung **C2** folgt $t_2 \geq 0$. Falls $t_2 < 1$, können wir die gleichen Argumente wie für S_0 benutzen und finden dabei einen Punkt

$(x_2, t_2) \in \Sigma^2_{stat} \cup \Sigma^3_{stat} \cup \Sigma^5_{stat}$, eine von S_1 verschiedene Zusammenhangskomponente S_2 von Σ^1_{stat} und eine Umgebung $U_{(x_2,t_2)}$, so daß folgende Aussagen gelten:

- $\lim_{t \to t_2} y_1(t) = x_2$,
- $\Sigma^1_{stat} \cap U_{(x_2,t_2)} = U_{(x_2,t_2)} \cap (S_1 \cup S_2)$,
- $S_1 \cap S_2 = S_0 \cap S_2 = S_1 \cap S_0 = \emptyset$,
- $(\overline{S_1} \setminus \Sigma^1_{gc}) \cap I\!R^n \times [0,1] = \{(x_1, t_1), (x_2, t_2)\}$.

Da $(x_2, t_2) \in \Sigma_{stat}$, folgt aus der Bedingung **C2** , daß $t_2 > 0$. Diese Analyse kann solange wiederholt werden, wie die neue Zusammenhangskomponente von Σ^1_{stat} keinen Punkt in $\{x, t | t \geq 1\}$ besitzt. Dadurch finden wir Punkte $\{(x_k, t_k)\} \subset C \times [0,1]$, $k = 1, 2, ..$ aus $\Sigma^2_{stat} \cup \Sigma^3_{stat} \cup \Sigma^5_{stat}$, und zweimal stetig differenzierbare Abbildungen $y_k(t)$, definiert für die offenen Intervalle $(\min(t_k, t_{k+1}), \max(t_k, t_{k+1}))$, mit $\lim_{t \to t_k} y_k(t) = x_k$ und $\lim_{t \to t_{k+1}} y_k(t) = x_{k+1}$. Dabei sind $S_k = G(y_k)$ zueinander verschiedene Zusammenhangskomponenten von Σ^1_{stat}.

Da alle Punkte von $\Sigma_{gc} \setminus \Sigma^1_{gc}$ isoliert sind, ist dieses Verhalten nur endlich oft möglich. Daraus folgt die Existenz eines $\bar{k}$ mit $1 \leq \bar{k} < \infty$, und $t_{\bar{k}} \geq 1$. Ist $t_{\bar{k}} = 1$, dann ist $(x_{\bar{k}}, 1)$ in Σ_{stat} und $\lim_{t \nearrow 1} y_{\bar{k}-1}(t) = x_{\bar{k}}$. Mit Hilfe der Abbildungen $y_0(t), y_1(t), ..., y_{\bar{k}-1}(t)$ bildet man jetzt trivialerweise die Kurve, deren Existenz im Satz behauptet wird. □

Behauptung 10
Es seien alle Voraussetzungen des Satzes 23 erfüllt, jedoch die Bedingung **C2** *ersetzt durch*

D2 *Die Variationsungleichung* $(F(.,0), H(.,0), G(.,0))$ *besitzt nur nichtdegenerierte stationäre Punkte und die Anzahl der stationären Punkte ist ungerade.*

Dann existiert eine stückweise zweimal stetig differenzierbare Kurve von stationären Punkten in $I\!R^n \times [0,1]$, *die einen Punkt* $(x_0, 0)$ *mit einem Punkt* $(x_1, 1)$ *verbindet.*

Beweis:

Betrachtet man die Konstruktion im vorhergehenden Beweis des Satzes 23 ausgehend von einem Punkt $(x_0^0, 0) \in \Sigma_{stat}^1$, erhält man endlich viele Zahlen $\{t_k\}_{k=1}^{\bar{k}}$ (eventuell $t_{\bar{k}} \in \{-\infty, +\infty\}$) mit

$$\begin{aligned} 0 &< t_k < 1, \ k = 1, ..., \bar{k} - 1 \\ t_k &\neq t_{k+1}, \ k = 1, ..., \bar{k} - 1 \\ t_{\bar{k}} &\notin (0, 1) \quad . \end{aligned}$$

Zu jedem t_k, $k = 1, ..., \bar{k} - 1$ (und auch für $k = \bar{k}$, falls $t_{\bar{k}}$ endlich ist) existiert ein x_k^0, so daß

$$\begin{aligned} (x_k^0, t_k) &\in \Sigma_{stat} \setminus \Sigma_{stat}^1 \\ (x_{k_1}^0, t_{k_1}) &\neq (x_{k_2}^0, t_{k_2}), \ \forall k_1 \neq k_2 \quad . \end{aligned}$$

Falls $t_{\bar{k}} \geq 1$, haben wir die gesuchte Kurve gefunden.

Andernfalls folgt aus **D2** $t_{\bar{k}} < 0$. Nach Konstruktion existiert eine Zusammenhangskomponente $S_{\bar{k}-1}^0$ von Σ_{stat}^1, die die Ebene $\{(x,t)|t = 0\}$ genau in einem Punkt $(x_0^1, 0) \in \Sigma_{stat}^1$ schneidet. Aus **D2** muß $(x_0^3, 0)$ existieren mit

$$\begin{aligned} (x_0^3, 0) &\in \Sigma_{stat}^1 \\ (x_0^3, 0) &\notin \{(x_0^0, 0), (x_0^1, 0)\} \quad . \end{aligned}$$

Dieser Prozeß kann wieder auf $(x_0^3, 0)$ angewendet werden, und solange nicht die gesuchte Kurve geliefert wird, findet man paarweise neue Punkte in $\Sigma_{stat}^1 \cap \{(x,t)|t = 0\}$. Wegen **D2** bleibt schließlich ein einziger noch nicht gewählter stationärer Punkt $(x_l^0, 0) \in \Sigma_{stat}^1$ übrig. Dieselben Argumente wie für $(x_0^0, 0)$ liefern dann (da $t_{\bar{k}} \geq 1$ seien muß bzgl. $(x_l^0, 0)$) die gesuchte Kurve. □

Behauptung 11

Es seien alle Voraussetzungen des Satzes 23 erfüllt, jedoch die Bedingung **C2** *ersetzt durch* **E2**.

E2 *Die Anzahl der Elemente der Menge*

$$(\Sigma_{stat}^1 \cup \Sigma_{stat}^5 \cup \{(x,t) \in \Sigma_{stat}^2 | \delta = 1\}) \cap \{(x,t)|t = 0\}, \tag{115}$$

wobei δ wie in (62) verstanden wird, ist ungerade.

Dann existiert eine stückweise zweimal stetig differenzierbare Kurve von stationären Punkten in $\mathbb{R}^n \times [0,1]$, die einen Punkt $(x_0, 0)$ mit einem Punkt $(x_1, 1)$ verbindet.

Beweis:

Laut Bemerkung 6 und Bild 10 sind die Punkte von Σ^2_{stat} mit $\delta = -1$ Rückkehrpunkte von Σ_{stat}. Damit gibt es für jeden Punkt in $\Sigma_{stat} \cap \{(x,t)|t = 0\}$ und außerhalb (115) entweder zwei oder keine Zusammnehangskomponente von Σ^1_{stat}, die die Menge

$$\{(x,t)|t > 0\} \tag{116}$$

schneidet.

Andererseits gibt es in jedem Punkt in (115) eine und nur eine Zusammenhangskomponente von Σ^1_{stat}, die (116) schneidet. Damit existiert ein $\varepsilon > 0$, so daß die Voraussetzungen der vorhergehenden Behauptung 10 (vertauscht man 0 durch ε) erfüllt sind. □

Behauptung 12

Es seien $(F, H, G) \in \mathcal{V}$ eine reguläre einparametrische Variationsungleichung und die Bedingungen **C1** *und* **C3** *aus Satz 23 erfüllt. Weiter seien*

$$\begin{aligned} \Sigma_{stat} \cap \{(x,t)|t = 0\} &\subset \Sigma^1_{stat} \\ \Sigma_{stat} \cap \{(x,t)|t = 1\} &\subset \Sigma^1_{stat} \quad . \end{aligned}$$

Es gilt dann die folgende Beziehung.

$$|\Sigma_{stat} \cap \{(x,t)|t = 0\}| = |\Sigma_{stat} \cap \{(x,t)|t = 1\}| \mod 2 \tag{117}$$

Beweis:

Nach der Konstruktion des Beweises von Satz 23 kann man die endlich vielen Punkte aus $\Sigma_{stat} \cap \{(x,t)|t = 0\}$ in zwei disjunkte Gruppen einteilen. Eine Gruppe hat eine gerade Zahl von Elementen und die andere Gruppe ist 1-1-zugeordnet mit einem Teil von $\Sigma_{stat} \cap \{(x,t)|t = 1\}$. Dasselbe kann für $\Sigma_{stat} \cap \{(x,t)|t = 1\}$ bezüglich $\Sigma_{stat} \cap \{(x,t)|t = 0\}$ gemacht werden, wobei dieselbe Zuordnung entsteht. Das impliziert dann die Beziehung (117). □

Behauptung 13
Es seien alle Voraussetzungen der vorhergehenden Behauptung 12 erfüllt. Es seien für $s = 0, 1$ die Mengen

$$\begin{aligned} K_s^{+1} &= \left\{(x,s) \in \Sigma_{stat} | sign(\det(D_x \mathbb{L}_{J_0(x,s)}|_{T_x M(s)})) = +1\right\} \\ K_s^{-1} &= \left\{(x,s) \in \Sigma_{stat} | sign(\det(D_x \mathbb{L}_{J_0(x,s)}|_{T_x M(s)})) = -1\right\} \end{aligned}$$

bezeichnet. Es gilt dann die folgende Beziehung.

$$|K_0^{+1}| - |K_0^{-1}| = |K_1^{+1}| - |K_1^{-1}| \tag{118}$$

<u>Beweis:</u>
Betrachtet man die Konstruktion im vorhergehenden Beweis des Satzes 23 ausgehend von einem Punkt $(x_0^0, 0) \in \Sigma_{stat}^1$, erhält man endlich viele Zahlen $\{t_k\}_{k=1}^{\bar{k}}$ (eventuell $t_{\bar{k}} \in \{-\infty, +\infty\}$) mit

$$\begin{aligned} 0 &< t_k < 1, \ k = 1, ..., \bar{k} - 1 \\ t_k &\neq t_{k+1}, \ k = 1, ..., \bar{k} - 1 \\ t_{\bar{k}} &\notin [0, 1] \quad . \end{aligned}$$

Zu jedem t_k, $k = 1, ..., \bar{k} - 1$ (und auch für $k = \bar{k}$, falls $t_{\bar{k}}$ endlich ist) existiert ein x_k^0, so daß

$$\begin{aligned} (x_k^0, t_k) &\in \Sigma_{stat}^2 \cup \Sigma_{stat}^3 \cup \Sigma_{stat}^5 \\ (x_{k_1}^0, t_{k_1}) &\neq (x_{k_2}^0, t_{k_2}), \ \forall k_1 \neq k_2 \end{aligned}$$

gilt. Man findet ebenfalls Zusammenhangskomponenten von Σ_{stat}^1

$$S_k, k = 0, ..., \bar{k} - 1$$

und zweimal stetig differenzierbare Abbildungen

$$\begin{aligned} y_0 &: [0, t_1) \to I\!R^n \\ y_k &: (\min(t_k, t_{k+1}), \max(t_k, t_{k+1})) \to I\!R^n, \quad k = 1, ..., \bar{k} - 1. \end{aligned}$$

mit

$$\begin{aligned} G(y_0) &\subset S_0 \\ G(y_k) &= S_k, \quad k = 1, ..., \bar{k} - 1 \end{aligned}$$

Andererseits existieren Umgebungen U_k von (x_k^0, t_k) für $k = 1, ..., \bar{k} - 1$ mit

$$U_k \cap \Sigma_{stat}^1 = U_k \cap (S_{k-1} \cup S_k).$$

Da das Vorzeichen von $\det(D_x \mathbb{L}_{J_0}|_{T_x M(t)})$ in einer Umgebung der Punkte vom Typ 1 konstant bleibt, ist dieses Vorzeichen auch konstant auf jeder der Zusammenhangskomponenten S_k.

Betrachten wir ein festes $(x_{\tilde{k}}^0, t_{\tilde{k}})$ mit $1 \leq \tilde{k} \leq \bar{k} - 1$. Es gibt 3 Möglichkeiten.

- $(x_{\tilde{k}}^0, t_{\tilde{k}}) \in \Sigma_{stat}^2$. Unter Beachtung der Behauptung 6 und des daraus resultierenden Bildes 10 kommt man zu folgender Fallunterscheidung: Ist $(x_{\tilde{k}}^0, t_{\tilde{k}})$ kein Rückkehrpunkt in Σ_{stat} (d.h. $\min(t_{\tilde{k}-1}, t_{\tilde{k}+1}) < t_{\tilde{k}} < \max(t_{\tilde{k}-1}, t_{\tilde{k}+1})$), dann ist das Vorzeichen von $\det(D_x \mathbb{L}_{J_0}|_{T_x M(t)})$ auf $S_{\tilde{k}-1}$ und auf $S_{\tilde{k}}$ gleich. Andererseits, falls $(x_{\tilde{k}}^0, t_{\tilde{k}})$ ein Rückkehrpunkt in Σ_{stat} ist (d.h. $t_{\tilde{k}-1}, t_{\tilde{k}+1} < t_{\tilde{k}}$ oder $t_{\tilde{k}-1}, t_{\tilde{k}+1} > t_{\tilde{k}}$), dann ist das Vorzeichen von $\det(D_x \mathbb{L}_{J_0}|_{T_x M(t)})$ auf $S_{\tilde{k}-1}$ und auf $S_{\tilde{k}}$ verschieden.

- $(x_{\tilde{k}}^0, t_{\tilde{k}}) \in \Sigma_{stat}^3$. Hier entsteht stets ein Rückkehrpunkt in Σ_{stat}, und laut Bemerkung 31 ist das Vorzeichen von $\det(D_x \mathbb{L}_{J_0}|_{T_x M(t)})$ auf $S_{\tilde{k}-1}$ und auf $S_{\tilde{k}}$ verschieden.

- $(x_{\tilde{k}}^0, t_{\tilde{k}}) \in \Sigma_{stat}^5$. Da die MFCQ erfüllt ist, gilt $\min(t_{\tilde{k}-1}, t_{\tilde{k}+1}) < t_{\tilde{k}} < \max(t_{\tilde{k}-1}, t_{\tilde{k}+1})$. Andererseits ist das Vorzeichen von $\det(D_x \mathbb{L}_{J_0}|_{T_x M(t)})$ auf $S_{\tilde{k}-1}$ und auf $S_{\tilde{k}}$ gleich, und nach Vereinbarung gleich $+1$.

Aus dieser Überlegung folgt, daß das Vorzeichen von $\det(D_x \mathbb{L}_{J_0}|_{T_x M(t)})$ auf S_k gleich ist (und gleich dem Vorzeichen auf S_0) für alle $k \in \{0, 1, ..., \bar{k}-1\}$ mit $t_k < t_{k+1}$. Andererseits ist dieses Vorzeichen verschieden von dem auf S_0 für alle S_k mit $t_k > t_{k+1}$.

Da $t_{\bar{k}} \notin [0, 1]$, existiert ein eindeutig bestimmter Punkt x^1, so daß genau eine der folgenden Bedingungen erfüllt ist.

- $(x^1, 0) \in \Sigma_{stat}$. Außerdem gilt $(x^1, 0) \in S_{\bar{k}-1}$ Damit ist $t_{\bar{k}} < t_{\bar{k}-1}$ und $sign(\det(D_x \mathbb{L}(x_0^0, 0)_{J_0}|_{T_x M(0)})) \neq sign(\det(D_x \mathbb{L}(x^1, 0)_{J_0}|_{T_x M(0)}))$.

- $(x^1, 1) \in \Sigma_{stat}$. Außerdem gilt $(x^1, 1) \in S_{\bar{k}-1}$ Damit ist $t_{\bar{k}} > t_{\bar{k}-1}$ und $sign(\det(D_x \mathbb{L}(x_0^0, 0)_{J_0}|_{T_x M(0)})) = sign(\det(D_x \mathbb{L}(x^1, 1)_{J_0}|_{T_x M(1)}))$.

Dieser Prozeß kann für alle stationären Punkte mit Komponente $t = 0$ wiederholt werden. Am Ende wird jedem stationären Punkt $(x, 0) \in \Sigma_{stat}$ eine andere $(q(x), s)$ $(s \in \{0, 1\})$ eindeutig zugeordnet, so daß

$$sign(\det(D_x\mathbb{L}(x,0)_{J_0}|_{T_xM(0)})) = sign(\det(D_x\mathbb{L}(q(x),s)_{J_0}|_{T_xM(s)})) \Leftrightarrow s = 1$$

gilt.

Das wird ausgehend von $t = 1$ nach $t = 0$ ausgeführt. Dabei wird die Zuordnung auf die Menge

$$\{(y, 1) \in \Sigma_{stat} | \exists (x, 0) \in \Sigma_{stat} \text{ mit } q(x) = y\}$$

erhalten.

Aus dieser Konstruktion folgt die Existenz zweier Zahlen $p_0, p_1 \geq 0$ mit

$$\begin{aligned} |K_0^{+1}| - p_0 &= |K_1^{+1}| - p_1 \\ |K_0^{-1}| - p_0 &= |K_1^{-1}| - p_1 \end{aligned}$$

Daraus folgt die angekündigte Beziehung. □

5.6 Weitere Entwicklung

In diesem kleinen Abschnitt wollen wir Fragestellungen auflisten, die ausgehend von dem in diesem Kapitel entwickelten Material betrachtet werden können. Damit wollen wir die mögliche weitere Entwicklung in dieser Richtung schildern und die gemachten Untersuchungen rechtfertigen:

- Ausgehend von der definierten Regularität einparametrischen Variationsungleichungen ist man jetzt in der Lage verschiedene Verfahren zur Lösung von Variationsungleichungen genauer zu untersuchen, indem man sie als Einbettung darstellt und die Werkzeuge der Regularität ausnutzt. Dabei kommt man zu einer kritischen Analyse von Konvergenzbedingungen. Man kann auch, wie im Fall von Optimierungsproblemen verbesserte Verfahren entwickeln. Als Beispiel in dieser Richtung können die Arbeiten [3], [4], [11] genannt werden, die dieser Ansatz für die Klasse $\mathcal{F}$ wiederspiegeln.

- Die entwickelte Regularität dient als Basis für die Untersuchung von Problemen, die Variationsungleichungen enthalten. Ein Beispiel dafür sind die sogenannten Optimierungsprobleme mit Equilibrium constraints (siehe z.B. [48]).

- Die Ergebnisse, gegeben in dem vorhergehendem Abschnitt, stellen eine Grundlage für die Anwendung von Kurvenverfolgungsmethoden zur Lösung von Variationsungleichungen dar. Insbesondere kann das Programpaket PAFO in diesem Sinne erweitert werden.

Literaturverzeichnis

[1] Allgower E.L. und K. Georg: Introduction to numerical continuation methods, Springer Verlag, Berlin, 1990.

[2] Bröcker Th.: Differentiable Germs and Catasthrophes, translated by L. Lander, London Math. Soc. Lectures Note Series 17, Cambridge Univ. Press 1975.

[3] Dentcheva, D., R. Gollmer, J. Guddat und J.-J. Rückmann: Pathfollowing methods in nonlinear optimization II: exact penalty methods, in M. Florenzano et. al. (eds.), Approximation and Optimization II (Proceedings of the 2nd. International Conference Approximation and Optimization in the Caribbean, Havanna, Cuba, 1993). Ser. Approximation and Optimization, Peter Lang Verlag, Frankfurt a.M., Berlin, Bern, New York, Paris, Wien 1995, 200-30.

[4] Dentcheva, D., J. Guddat, J.-J. Rückmann und K. Wendler: Pathfollowing methods in nonlinear optimization III: multiplier embedding, in Mathematical Methods of Operation Research, ZOR, 1994.

[5] Facchinei F. und J. Soares: A new merit function for nonlinear complementarity problems and a related algorithm, SIAM J. Optim, Vol. 7, No. 1, (1997), 225-247.

[6] Fiacco A.V.: Introduction to Sensitivity and Stability Analysis in Nonlinear Programming, Academic Press, New York, 1983.

[7] Fiacco A.V. und G.P. McCormick: Nonlinear Programming: Sequential Unconstrained Minimization Techniques, Wiley, New York, 1968.

[8] Fukushima M und K. Taji: A new merit function and a succesive quadratic programming algorithm for variational inequality problems, SIAM J. Optimization, Vol 6, No. 3, (1996) 704-713.

[9] Fukushima M. und N. Yamashita: Equivalent unconstrained minimization and global error bounds for variational inequality problems, SIAM J. Control Optim., Vol. 35, No. 1, (1997) 273-284.

[10] Gfrerer H., J. Guddat, Hj. Wacker und W. Zulehner: Path-following methods for Kuhn-Tucker curves by an active index set strategy. In A Bagchi and H.Th. Jongen (eds.) Systems and Optimization, Lectures Notes in Control and Information Sciences 66, Springer Verlag, Berlin, Heidelberg, New York, Tokyo, (1985), 111-131.

[11] Gollmer R., J. Guddat, F. Guerra, D. Nowack und J.-J. Rückmann: Pathfollowing methods in nonlinear optimization: penalty embedding, In: J. Guddat et al. (eds.), Parametric Optimization and Related Topics III, Peter Lang Verlag, Frankfurt a.M., Berlin, Bern, New York, Paris, Wien, (1993), 163-214.

[12] Gollmer R., Kausmann U., Nowack D. und Wendler K. (1997): Computerprogramm PAFO. Humboldt Universität Berlin, Institut für Mathematik.

[13] Gomez W., J. Guddat, H.Th. Jongen, J.-J. Rückmann und C. Solano: Curvas críticas y saltos en optimización no lineal. Lehrbuch auf Spanisch, (Manuskript, Berlin 1998).

[14] Guddat J., F. Guerra und H.Th. Jongen: Parametric Optimization: singularities, pathfollowing and jumps. John Wiley, Chichester and Teubner, Stuttgart 1990.

[15] Guddat J., F. Guerra und D. Nowack: On the role of the Mangasarian-Fromowitz contraints qualification for penalty, exact penalty and Lagrange multiplier methods, in Mathematical programming with data perturbations, edited by A.V. Fiacco, Lectures notes in pure and applied mathematics vol.195, Marcel Dekker, Inc New York, 1997.

[16] Guddat J., H.Th. Jongen und J.-J. Rückmann: On stability and stationary points in nonlinear optimization. J. Austral. Math. Soc. Ser. B 28 (1986), 36-56.

[17] Guddat J., H.Th. Jongen und D. Nowack: Parametric optimization: pathfollowing with jumps, in A. Gomez, F. Guerra, M.A. Jimenez and G. Lopez (Hgs) Approximation and Optimization, Lect. Notes Math. 1354, Springer-Verlag, Berlin, 1988, pp. 43-55.

[18] Guddat J. und J.-J. Rückmann: One-parametric optimization: jumps in the set of generalized critical points, Control and Cybernetics 23, 1/2, 139-151, 1994.

[19] Gou-chen, Feng und Yu Bo: Combined Homotopy Interior Point Method for Nonlinear Programming Problems, in Advances in Numerical Mathematics, Proceedings of the Second Japan-China Seminar on Numerical Mathematics, Lecture Notes in Num. Appl. Anal., 14, 9-16 (1995).

[20] Harker P.T. und Jong-Shi Pang: Finite-dimensional variational inequality and nonlinear complementarity problems: A survey of theory, algorithms and applications. Mathematical Programming 48 (1990) 161-220.

[21] Harker P.T. und B. Xiao: A nonsmooth Newton method for variational inequalities, I: theory, Mathematical Programming 65 (1994) 151-194.

[22] He B.: A Class of Projection and Contraction Methods for Monotone Variational Inequalities. Appl. Math. Optim. 35, 69-76 (1997).

[23] Henn M., P. Jonker und F. Twilt: On the critical sets of one-parameter quadratic optimization problems, Durier, Roland et al. (eds.) Recent developments in optimization, 7th French-German conference on optimization, Dijon, France June 1994, Springer Verlag Berlin, Lect. Notes Econ. Math.Syst. 429, 183-197 (1995).

[24] Hettich R. und P. Zencke: Numerische Methoden zur Approximation und semi-infiniten Optimierung, Teubner Studienbücher, Stuttgart, 1982.

[25] Hettich, R., H.Th. Jongen und O. Stein: On continous deformations of semi-infinite optimization problems, in: M. Florenzano, J. Guddat, M. Jimenez, H.Th. Jongen, G. Lopez Lagomasino, F. Marcellan (eds.): Approximation and Optimization in the Caribbean II, Peter Lang Verlag, Frankfurt a.M., 1995, 406-424.

[26] Hirsch M.W.: Differential Topology, Grad. Texts Math., Vol. 33, Springer Verlag, Berlin, 1976.

[27] Fletcher R.: Penalty Functions, in Mathematical Programming: The State of the Art, A. Bachem, M. Grötschel and B. Korte, eds., Springer Verlag, New York, 1983.

[28] Jiang H. und L. Qi: A new nonsmooth equations approach to nonlinear complementarity problems: SIAM J. Control Optim., Vol. 35, No. 1, (1997) 178-193.

[29] John F.: Extremum problems with inequalities as subsidiary conditions, in Studies and Essays, R. Courant anniversary volume, Interscience, New York, 1948, pp. 187-204.

[30] Jongen H.Th.:Parametric Optimization: Critical points and Local Minima, In: Lectures in Applied Mathematics, Vol. 26, American Mathematical Society (Eds.: E.L. Allgower and K. Georg): "Computational Solution of Nonlinear Systems of Equations, pp. 317-335 (1990).

[31] Jongen H.Th., P. Jonker und F. Twilt: Nonlinear Optimization in $\mathbb{R}^n$ I: Morse Theory, Chebyshev Approximation, Peter Lang Verlag, Frankfurt a. M., 1983.

[32] Jongen H.Th., P. Jonker und F. Twilt: Nonlinear Optimization in $\mathbb{R}^n$ II: Transversality, Flows, Parametric Aspects, Peter Lang Verlag, Frankfurt a. M., 1986.

[33] Jongen H.Th., P. Jonker und F. Twilt: On one-parameter-families of sets defined by (in)equality constraints, Nieuw Archief voor Wiskunde (3) 30 (1982) 307-322.

[34] Jongen H.Th., P. Jonker und F. Twilt: One parametric families of optimization problems: equality constraints, JOTA 48 (1986), 141-161.

[35] Jongen H.Th., P. Jonker und F. Twilt: Critical sets in parametric optimization, Math. Program. 34 (1986), 333-353.

[36] Jongen, H.Th., D. Klatte und K. Tammer: Implizit functions and sensitivity of stationary points, math. Programming 49 (1990), 123-138.

[37] Jongen, H.Th., T. Möbert, J-J. Rückmann und K. Tammer: On inertia and Schur complement in optimization. Linear Algebra Appl. 95 (1987), 97-109.

[38] Jongen, H.Th. und O. Stein: On generic one-parametric semi-infinite optimization, SIAM J. Optim. 7, No. 4, 1103-1137, 1997.

[39] Jongen, H.Th. und O. Stein: Parametric semi-infinite programming: jumps in the set of local minimizers, In J. Guddat, H.Th. Jongen, F. Nožička, G. Still, F. Twilt (eds): Parametric Optimization and Related Topics IV, Peter lang Verlag, Farnkfurt a.M., 1997, 161-175.

[40] Klatte D.: On strongly stable local minimizers in nonlinear programs. In: J. Guddat et al. (Herausgeber), Advances in Mathematical Optimization, Akademie-Verlag, Berlin, 1988, 104-113.

[41] Klatte D.: Strong stability of stationary solutions and iterated local minimization.In: J. Guddat et al. (Herausgeber), Parametric Optimization and Related Topics II, Akademie-Verlag, Berlin, 1991, 119-136.

[42] Klatte D., B. Kummer und R. Walzebok: Conditions for optimality and strong stability in nonlinear programs without assuming twice differentiability of data. IIASA Working Paper WP-89-089, Laxenburg, Austria, 1989.

[43] Klatte D. und K. Tammer: Strong stability of stationary solutions and Karush-Kuhn-Tucker points in nonlinear optimization. Ann. Oper. Res. 27 (1990), 285-308.

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

[45] Kojima M. und R. Hirabayashi, Continuos deformation of nonlinear programs, Math, Program. Stud. 21 (1984), 150-198.

[46] Kummer B.: The inverse of a Lipschitz function in $I\!R^n$: Complete characterization by directional derivatives, Preprint Nr. 195, Humboldt-Universität zu Berlin, Sektion Mathematik, 1988

[47] Kummer B.:Lipschitzian inverse functions, directional derivatives and appplications in $C^{1,1}$-optimization, Journal Optimization Theory Appl., 70, (1991) 561-582.

[48] Luo Zhi-Quan, Jong-Shi Pang und Daniel Ralph: Mathematical Programs with Equilibrium Constraints, Cambridge University Press, 1996.

[49] Marcotte P. und D.L. Zhu: Co-coercivity and its role in the convergence of iterative schemes for solving variational inequalities, SIAM J. Optimiization, Vol. 6, No. 3 (1996), 714-726.

[50] Milnor J.:Topologie from the differentiable viewpoint, The University Press of Virginia (1965).

[51] Mangasarian O.L. und S. Fromovitz,: The Fritz John necessary optimality conditions in the presence of equality and inequality constraints, J. Math. Anal. Applic. 17 (1967), 37-47.

[52] Ouellete, D.V.: Schur complements and statistics, Lin. Alg. Appl. 36 (1981), 187-295.

[53] Pateva D.: On the singularities in linear one-parametric optimization problems, Optimization 22, (1991), 193-220.

[54] Poore A.B. und Q. Al-Hassan: The expanded Lagrangian system for constrained optimization problems, in SIAM J. Control and Opt, Vol 26, No. 2, (1988), pp 417-27.

[55] Poore A.B. und Q. Al-Hassan: Bifurcation analysis for singularities on a tangent space for quadratic penalty-barrier and multiplier methods for solving constrained optimization problems, part I, in J. Math. Anal. Appl. 197, (1996), pp. 658-78.

[56] Poore A.B. und Q. Al-Hassan: Analysis of bifurcation due to loss of linear independence and strict complementarity for penalty methods for solving constrained optimization problems, part II, in J. Math. Anal. Appl. 197, (1996), pp. 756-85.

[57] Poore A.B. und C.A. Thiart: Bifurcation problems in nonlinear parametric programming, in Mathematical Programming 39 (1987), pp. 189-205.

[58] Poore A.B. und C.A. Thiart: A bifurcation analysis of the nonlinear parametric programming problem, in Mathematical Programming 47 (1990), pp. 117-141.

[59] Robinson S.M.: Strongly regular generalized equations. Math. Oper. Res. 5 (1980), 43-62.

[60] Rückmann J.-J.: Einparametrische nichtkonvexe Optimierung: Strukturuntersuchungen und eine Verallgemeinerung des Einbettungsprinzips. Dissertation, Technische Hochschule Leipzig, Sektion Mathematik und Informatik, 1988.

[61] Rückmann J.-J. und K. Tammer: On linear-quadratic perturbations in one-parametric nonlinear optimization. Systems Science 18 (1992) 1, 37-48

[62] Sternberg S.: Lectures on Differential Geometry, Prentice Hall, Inc. 1964.

[63] Stein O.: On Parametric Semi-infinite Optimization, Dissertation, Shaker-Verlag, Aachen, 1997.

[64] Tammer K.: Strong stability of stationary points for $C^{1,1}$-optimization problems. 21. Jahrestagung "Mathematical Optimization", Benneckenstein, 1989.

[65] Tammer K.: Parametric linear complementarity problems, in Mathematical programming with data perturbations, edited by A.V. Fiacco, Lectures notes in pure and applied mathematics vol.195, Marcel Dekker, Inc New York, 1997.

[66] Wendler K.:Implementation of a Pathfollowing Procedure for Solving Nonlinear One-Parametric Optimization Problems. In: Brosowski, B et al. (eds.) Multicriteria Desicion. In: Ser. Approximation and Optimization. Verlag Peter Lang, Frankfurt a.M., Berlin, Bern, New York, Paris, Wien, 139-163.

LEBENSLAUF

Walter Gómez Bofill

26. Oktober 1969 : Geboren in Habana, Kuba.

1975 - 1981 : Besuch der Grundschule in Habana.

1981 - 1984 : Besuch der Mittelschule in Habana.

1984 - (Juli)1987 : Besuch der Oberschule in Habana, mit Erlangung der Hochschulereife.

Sept.1987 - Juli 1988 : Beginn des mathematischen Studiums an der Universität Habana. (2 Semester)

Sept.1988 - Februar 1989 : 6 monatiger Sprachkurs in Deutsch.

Februar 1989 - Juli 1990 : Fortführung des Mathematik-Studiums für 3 Semester an der Technischen Universität Dresden.

Sept. 1990 - Januar 1994 : Beendigung des Mathematik-Studiums an der Universität Habana.

April 1994 : Fertigstellung der Diplomarbeit.

- Betreuer: Prof. Dr. Guddat.
- Titel: "Vektoroptimierung: Singularitäten, Regularisierungen"

Sept. 1994 - Nov 1995 : Mitarbeiter an der Mathematik-Fakultät der Universität Habana.

Dez. 1995 - Nov. 1998 : Doktorand an der Humboldt Universität zu Berlin mit einem DAAD Stipendium.

Zusätzliche Tätigkeiten :

- Feb. 1991 - Juli 1992: Mitarbeit am Projekt "Multilineare Suche" am Institut für Mathematik und Physik der Akademie der Wissenschaften von Kuba.
- Feb. 1991 - Feb 1994: Studentische Hilfskraft an der Math. Fakultät der Uni. Habana.

Publikationen

- "On multi-directional search in Optimization", in Zusammenarbeit mit Dr. Alfredo Gómez von der ICIMAF, Kuba. In Proceedings 3rd. International Conference on Approximatión and Optimization, Puebla, 1995. Veröffentlich in Internet bei der Electronic Library of mathematics
  http://www.emis.de/proceedings/index.html
  Eine feste Version ist in Vorbereitung bei "Aportaciones Matemáticas", SMM.
- "Vector optimization: Singularities, Regularizations". Mathematical Methods of Operation Research, ZOR, Volume 47, Issue 3, 1998

Preprints : Am Institut für Mathematik der Humboldt Universität zu Berlin.

- "Properties of an interior embedding for solving nonlinear optimization problems", Preprint 96-31.
- "On generic quadratic penalty embeddings for nonlinear optimization problems", Preprint 97-18.