Asymptotische Stabilität von Index-2-Algebro-Differentialgleichungen und ihren Diskretisierungen

D I S S E R T A T I O N

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

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät II
Humboldt-Universität zu Berlin

von
Herr Antonio R. Rodríguez S.
geborem am 09.02.1968 in Havanna, Kuba

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Jürgen Mlynek

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

Gutachter:

1. Prof. Dr. Roswitha März
2. Prof. Dr. Claus Führer
3. Prof. Dr. Michael Hanke

eingereicht am: 17. November 2000
Tag der mündlichen Prüfung: 2. Februar 2001

Abstract

The purpose of the present PhD work is the asymptotic stability investigation of numerical methods for index 2 differential algebraic equations. Initial value problems are considered for quasi linear differential algebraic equations (DAEs) that cover the most important applications.

First some stability concepts and related results are presented, which represent the basis for further investigations. This background concerns both, the continuous and the discreet case. Especially contractivity concepts are introduced and the relationship between the asymptotic stability of the DAE and the numerical method applied to it is established. The new contractivity concepts extend or generalize the already known concepts. The most important result in this context is a theorem that establishes general conditions under which the application of an algebraic stable IRK(DAE) method to a DAE is contractive. Well-known assertions for ordinary and differential algebraic equations can be considered as special cases of this general result.

Later on the stability of numerical discretizations applied to index-2 DAEs is investigated. This is made possible by the introduction of new decopling and index reduction techniques. The analysis makes new insights in the asymptotic of numerical methods for DAEs possible. The obtained results state sufficient conditions in order that a BDF or an IRK(DAE) method applying to DAEs shows the same asymptotic stability properties as for ODEs. These results are illustrated by some numerical examples. Moreover, it can be realized that one of the found conditions is sufficient in order to show contractivity of the application of an algebraic stable IRK(DAE) method, supposed the DAE is contractive. This assertion is possible based on the general theorem mentioned in the paragraph above. Further some consequences of the mentioned results for electric network models are shown.

According to both, the above mentioned analysis and the specialized literature of this field, the application of numerical methods to some special DAEs shows asymptotic stability problems. A few approaches are known to manage such difficult equations. Two exponents of these techniques are considered and their chances of success for index-2 DAEs are evaluated with the application to a critical example. A generalization of the Gear-Gupta-Leimkuhler (GGL) approach is proposed for full implicit linear DAEs. This generalization is investigated in detail in the rest of the paper, concerning both the analytical and the numerical asymptotic stability of the GGL equation and the numerical methods applied to it correspondingly. The result is, that, if some conditions are fulfilled, IRK(DAE) and BDF methods for

the GGL equation will produce stable solutions. This result is illustrated by a numerical example. The application of the methods directly to the considered DAE produces unstable solutions. However, the integration of the corresponding GGL formulation is stable. The obtained result opens new possibility for the numerical treatment of instabilities by differential algebraic equations.

Keywords:
Differeintial Algebraic Equations, Numerical Methods, Asymptotical Stability, Contractivity

Zusammenfassung

Ziel dieser Dissertation ist die Untersuchung der asymptotischen Stabilität numerischer Verfahren für Index-2-Algebro-Differentialgleichungen. Es werden Anfangswertaufgaben für quasilineare Algebro-Differentialgleichungen (ADGln). Die meisten anwendungsrelevanten Aufgaben können damit behandelt werden.

Zuerst werden einige Stabilitätsbegrife und Aussagen vorgestellt, die das Fundament für den Rest der Arbeit darstellen. Dies erstreckt sich sowohl auf den kontinuierlichen als auch auf den diskreten Fall. Insbesondere werden Kontraktivitätskonzepte eingeführt und Beziehungen zwischen der Kontraktivität der ADGl und derer der Anwendung eines numerischen Verfahrens. Die eingeführte Kontraktivitätsbegriffe erweitern oder verallgemeinern die bereits bekannten Konzepte. Als wichtigste Aussage in dem Kontraktivitätskontext geht ein Theorem hervor, das allgemeine Bedingungen aufstellt, damit die Anwendung eines IRK(DAE)-Verfahrens auf eine ADGl stabil ist. Bekannte Aussagen für gewöhnliche und Algebro-Differntialgleichungen können als Sonderfälle dieses Ergebnisses gesehen werden.

Im weiteren Verlauf der Arbeit wird anhand von neuartigen Index-2-Entkopplungs- und Indexreduktionstechniken die Stabilität von Diskretisierungsverfahren untersucht. Die durchgeführte Analyse erbringt neue Ergebnisse, die eine Verbesserung des Kenntnissstandes in diesem Gebiet darstellen. Die erzielte Aussagen stellen hinreichende Bedingungen, damit ein BDF- oder IRK-Verfahren für eine ADGl das gleiche Stabilitätsverhalten wie für eine gewöhnliche Differentialgleichung besitzt. Diese Ergebnisse werden durch numerishce Beispiele veranschaulicht. Weiterhin stellt man fest, dass eine der gefundenen Voraussetzungen für die Kontraktivität der Anwendung eines algebraisch stabilen IRK(DAE)-Verfahrens, auf eine ebenfalls kontraktive ADGl, genügt. Dieses Ergebnis wurde durch die Anwendung der im ersten Teil dieser Arbeit erzielten Kontraktivitätsaussagen ermöglicht. Die Konsequenzen der soeben genannten Aussage für bestimmte Modelle der Schaltkreissimulation werden ebenfalls erläutert.

Aus der oben genannten Analyse, ebenso wie aus der Fachliteratur, geht hervor, dass bei manchen ADGl-Aufgaben die Diskretisierungsverfahren Stabilitätsprobleme aufweisen. Um solche Probleme zu behandeln sind bereits einige Ansätze bekannt. Im letzten Teil der Arbeit werden zwei repräsentativen

Ansätze betrachtet und ihre Aussichtschancen für Index-2-Aufgaben anhand eines kritischen Beispieles evaluiert. Des Weiteren wird eine Verallgemeinerung für vollimplizite lineare ADGln des Gear-Gupta-Leimkuhler-Ansatzes (GGL) vorgeschlagen. Der Rest der Arbeit beschäftigt sich mit der Stabilitätsuntersuchung der GGL-Formulierung und der auf sie angewandten numerischen Verfahren. Dafür werden Aussagen dieser Arbeit eingesetzt und man kommt zu der Schlussfolgerung, dass sowohl für die IRK(DAE)- als auch für die BDF-Verfahren die Integration der GGL-Formulierung, natürlich unter bestimmten Voraussetzungen, stabil ist. Dieses Ergebniss wird durch ein numerisches Beispiel belegt. Dabei handelt es um eine Gleichung, die mit einer direkten Anwendung eines Verfahrens Instabilitäten aufweist. Jedoch ist die Integration der entsprechenden GGL stabil.

Sclagwörter:
Algebro-Differentialgleichungen, Numerische Verfahren, Asymptotische Stabilität, Kontraktivität

Vorwort

Die vorliegende Dissertation fand größtenteils im Rahmen des Graduierten-Kollegs "Geometrie und nicht lineare Analysis"der Humboldt-Universität zu Berlin statt.

Ich möchte Frau Prof. R. März für die langjährige Unterstützung, Leitung und Verbesserungsvorschläge auf dem steinigen Weg dieser Arbeit besonders herzlich danken. Auch sehr hoch zu schätzen ist die von Herr Prof. M. Hanke (Royal Institute of Technology Stockholm) geleistete enge Betreuung, vor allen Dingen in der schwierigen Anfangsphase. Ihm gilt mein herzlicher Dank.

Weiterhin möchte ich Herrn Prof. Claus Führer (Lund-University, Schweden) für sein Interesse an meiner Arbeit und die fruchtbaren Diskussionen danken.

Es wäre unverzeihbar, an dieser Stelle nicht die Mitarbeiter des Instituts für Mathematik der Humboldt-Universität zu Berlin, insbesondere der Gruppe Numerische Mathematik", zu nennen. Sie leisteten nicht nur im fachlichen Bereich einen wichtigen Beitrag, sondern schufen durch ihren familiären und vertrauten Umgang eine exzellente Arbeitsatmosphäre.

Zuletzt möchte ich meine Dankbarkeit Frau Prof. Inmaculada Higueras von der Universidad Pública de Navarra Spanien für die nützlichen Kommentare und Diskussionen aussprechen.

Berlin, den 25.10.2001

Antonio R. Rodríguez S.

Inhaltsverzeichnis

Abbildungsverzeichnis

Bezeichnungen und Konventionen

$\mathbb{R}$	Menge der reellen Zahlen
$\mathfrak{I}$	$\subseteq \mathbb{R}$ Intervall
$\mathbb{R}^m$	Menge der m-dimensionalen reellen Vektoren
D	offene Menge in $\mathbb{R}^m$
D_f	$\mathbb{R}^m \times D \times \mathfrak{I}$ (im Zusammenhang mit der Gleichung $f(x', x, t) = 0$)
$L(\mathbb{R}^m)$	Menge der linearen Abbildungen von $\mathbb{R}^m$ in $\mathbb{R}^m$,
	Menge der reellen $m \times m$-Matrizen
GDGl	gewöhnliche Differentialgleichung
ADGl	Algebro-Differentialgleichung
AWA	Anfangswertaufgabe
IRD	Index-Reduktion durch Differentiation
(A, B)	A, $B \in L(\mathbb{R}^m)$; das Matrix-Büschel, das der ADGl $Ax' + Bx = 0$ zugeordnet wird
$\operatorname{im} A$	Bild/Image der Abbildung A
$\ker A$	Nullraum der Matrix A
$\dim R$	Dimension des Unterraumes R
N	$:= \ker A$

S	$:= \{z \in \mathbb{R}^m : Bz \in \operatorname{im} A\}$
I	$\in L(\mathbb{R}^m)$, die Einheitsmatrix
Q	$\in L(\mathbb{R}^m)$, eine Projektor-Matrix/ein Projektor auf N
Q_c	$\in L(\mathbb{R}^m)$, der Projektor auf N längs S (der kanonische Projektor auf N)
P, P_c	$:= I - Q$, $:= I - Q_c$
$Pr(U_1, U_2)$	Projektor auf den Unterraum U_1 längs dem Unterraum U_2
B_1	$:= BP$
G_1	$:= A + BQ$
N_1	$:= \ker G_1$
S_1	$:= \{z \in \mathbb{R}^m : BPz \in \operatorname{im} G_1\}$
$Q^*_{N_1}$	$:= Pr(N_1, *)$ Projektor auf N_1 längs irgendeinem Unterraum
$Q_*^{S_1}$	$:= Pr(*, S_1)$, Projektor auf irgendeinem Unterraum längs S_1
Q_1	$:= Pr(N_1, S_1)$, Der Projektor auf N_1 längs S_1 (der kanonische Projektor)
P_1	$:= I - Q_1$
$P_*^{N_1}$	$:= I - Q^*_{N_1}$
$P^*_{S_1}$	$:= I - Q_*^{S_1}$
G_2	$:= G_1 + B_1 Q^{S_1}_{N_1}$
$G_{2,*}$	$:= G_1 + B_1 Q^*_{N_1}$
T	$:= Pr(N \cap S, *)$
U	$:= I - T$
W_0	$Pr(*, \operatorname{im} A)$
W_1	$Pr(*, \operatorname{im} G_1)$

Z^+	Moore-Penrose Inverse von Z
Z^-	andere reflexive verallgemeinerte Inverse von Z
$f : G \to H$	Abbildung von der Menge G in die Menge H
$C(\Im, I\!R^m)$	Menge der stetigen Funktionen von $\Im$ in $I\!R^m$
$C^1(\Im, I\!R^m)$	Menge der stetig differenzierbaren Funktionen von $\Im$ in $I\!R^m$
$C^1_N(\Im, I\!R^m)$	$:= \{x \in C[\Im, I\!R^m] : Px \in C^1[\Im, I\!R^m]\}$
x'	Ableitung der Funktion x nach t
$f_x,\ f'_x$	partielle Ableitung von f nach x
$\|c\|$	Betrag der reellen Zahl c
$\langle y, x\rangle$	beliebiges Skalarprodukt von zwei Vektoren
$\langle y, z\rangle_2$	euklidisches Skalarprodukt von zwei Vektoren y, z
$\|z\|_2$	euklidische Norm eines Vektors z
$\langle y, z\rangle_D$	Skalarprodukt $y^T Dz$, mit D symmetrisch positiv definit
$\|z\|_D$	durch $\langle \cdot, \cdot \rangle_D$ induzierte Norm von z
$\|z\|_\infty$	$\max\left\{\|z_i\| \,;\, z = (z_1, \ldots, z_m)^T, i = 1, \ldots, m\right\}$
$\|A\|$	$\max\{\|Az\| \,;\, z \in I\!R^m, \|z\| = 1\}$ für $A \in L(I\!R^m)$
$\{I\!R^m, \langle \cdot, \cdot \rangle\}$	euklidischer Raum
M_0	Lösungsmannigfaltigkeit einer Index-1-ADGl
M_1	Lösungsmannigfaltigkeit einer Index-2-ADGl

Kapitel 1

Einleitung

Im Laufe der letzten Jahrzehnte entstand aus der gewöhnlichen Differentialgleichungstheorie ein neues eigenständiges Fachgebiet, das sich mit den sogenannten Algebro-Differentialgleichungen (ADGln) befasst. Solch ein System besteht aus einer impliziten gewöhnlichen Differentialgleichung

$$f(x'(t), x(t), t) = 0, \quad t \in [t_0, \infty) \tag{1.1}$$

mit dem besonderen Merkmal, dass die Jacobi-Matrix $f'_{x'}$ nicht regulär ist. Diese Eigenschaft unterscheidet sie von den gewöhnlichen Differentialgleichungen (GDGln) und bringt radikale Konsequenzen mit sich. Einen wichtigen Sonderfall von (1.1) stellen die semi-expliziten Systeme dar:

$$\begin{aligned} x'_1(t) + b_1(x_1(t), x_2(t), t) &= 0, \\ b_2(x_1(t), x_2(t), t) &= 0, \end{aligned}$$

in denen der Ausdruck "Algebro-Differentialgleichung" ersichtlich wird.

Man führte für (1.1) einen Index-Begriff (im Grunde genommen gibt es mehrere Index-Begriffe, die sich leicht unterscheiden) und damit eine Unterteilung der ADGln ein. In diesem Kontext sind die gewöhnlichen Differentialgleichungen Index-0-ADGln und je höher der Index ist, umso mehr unterscheidet sich die ADGl von einer GDGl. Der Index kann als ein Maß des Schwierigkeitsgrades einer Aufgabe angesehen werden. Index-1-Probleme verhalten sich fast wie GDGln und bereiten im Regelfall keine großen Schwierigkeiten. Ab Index-2 weisen die ADGln einen qualitativen Unterschied auf: Die Gleichungen enthalten "versteckte" Nebenbedingungen, die durch eine

Diskretisierung verloren gehen. Gleichermaßen führen ADGln höheren Indexes (Index≥2) zu Differentiationsaufgaben, die die ADGln als schlecht gestellte Aufgaben im Sinne von Hadamard prägen, wenn man stetige Störungen betrachtet.

Systeme wie (1.1) sind nicht nur aus mathematischer Sicht von Interesse, sie kommen auch bei praktischen Aufgabenstellungen vor. Zu nennen sind die Modellierung von Schaltkreisen, die Mechanik der Mehrkörpersysteme, die Lösung von Optimalsteuerungsproblemen bei den verschiedensten Anwendungen und die Simulation von Prozessen in chemischen Anlagen.

In der Vergangenheit erzielte man auf dem Gebiet der ADGln große Vorschritte sowohl bei der Aufklärung der analytischen Eigenschaften, als auch bei den Diskretisierungs verfahren, und bezüglich vieler Aspekte steht heutzutage ein nützlicher Apparat zur Verfügung. Ein Aspekt, auf den diese Aussage nicht zutrifft, ist zweifellos die asymptotische Stabilität ADGln höheren Indexes, insbesondere bei den numerischen Verfahren. Zwar sind die Durchführbarkeit, die Konsistenzordnung und die Konvergenz der Diskretisierungsverfahren sehr ausführlich untersucht worden ([Griepentrog and März, 1986], [Brenan et al., 1989], [Hairer and Wanner, 1991], [Tischendorf, 1996] unter anderen), aber der Kenntnisstand über das asymptotische Verhalten einer Näherungslösung ist unzureichend. Andererseits bestätigen die numerischen Experimente, dass die Diskretisierungsverfahren ihre Stabilitätseigenschaften verlieren können, [Hanke et al., 1998], [Wensch et al., 1995], [Eich-Soellner and Führer, 1998]. In [Hanke et al., 1998] wurde das folgende Beispiel vorgestellt:

Beispiel 1.0.1

$$\begin{pmatrix} x_1' \\ x_2' \\ 0 \end{pmatrix} + \begin{pmatrix} \lambda & -1 & -1 \\ \eta t(1-\eta t) - \eta & \lambda & -\eta t \\ 1-\eta t & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0,\ t \geq 0.$$

wobei λ, η reelle Parameter sind. Es handelt sich um ein Index-2-System, dessen allgemeine Lösung durch

$$\begin{aligned} x_1(t) &= x_1(0)e^{-\lambda t}, \\ x_2(t) &= (\eta t - 1)x_1(t), \\ x_3(t) &= -(\eta t - 1)x_1(t), \end{aligned}$$

gegeben ist. Offensichtlich sind alle Lösungen exponentiell asymptotisch stabil für $\lambda > 0$. Ebenso kann man erkennen, dass alle Lösungen zu der Ebene

$$M_1(t) = \left\{ z \in I\!R^3 : z = \begin{pmatrix} z_1 \\ (\eta t - 1)z_1 \\ (1 - \eta t)z_1 \end{pmatrix} \right\}$$

gehören. Dadurch kann das ursprünglichen System zu der trivialen gewöhnlichen Differentialgleichung

$$x_1'(t) + \lambda x_1(t) = 0$$

reduziert werden.

Auf der anderen Seite stellt die Anwendung des impliziten Euler-Verfahrens auf dieses Problem die Lösung folgenden linearen Systems in jedem Schritt dar:

$$\begin{pmatrix} 1 + h\lambda & -h & -h \\ h(\eta t_{i+1}(1 - \eta t_{i+1}) - \eta) & 1 + h\lambda & -h\eta t_{i+1} \\ 1 - \eta t_{i+1} & 1 & 0 \end{pmatrix} \begin{pmatrix} x_{1,i+1} \\ x_{2,i+1} \\ x_{3,i+1} \end{pmatrix} = \begin{pmatrix} x_{1,i} \\ x_{2,i} \\ 0 \end{pmatrix},$$

wobei $h > 0$ der Diskretisierungsschritt ist. Aus diesem System folgt für die erste Komponente der Lösung die Rekursion

$$x_{1,i+1} = \frac{1 + h\eta}{1 + h(\eta + \lambda)} x_{1,i},$$

und für die asymptotische Stabilität des Verfahrens in dieser Komponente erhält man die Bedingung

$$\left| \frac{1 + h\eta}{1 + h(\eta + \lambda)} \right| < 1.$$

Diese Bedingung kann für $\lambda > 0$ und $\eta < 0$ verletzt werden, insbesondere für $\eta = -\lambda$, obwohl die analytische Lösung in diesem Fall asymptotisch stabil ist.

Nach diesem Beispiel ist die Notwendigkeit einer Stabilitätsuntersuchung deutlich geworden. Stabilitätsprobleme können prinzipell bei ADGln mit Index höher als 0 auftreten, sind aber im Index-1-Fall eher eine Seltenheit. Der Grund dafür wurde in [Griepentrog and März, 1986] geklärt. An

jener Stelle wurde gezeigt, dass, wenn der Nullraum der Matrix $f_{x'}$ konstant ist, die Diskretisierungsverfahren ihre von den GDGln bekannten Stabilitätseigenschaften behalten. Insbesondere bewies man, dass algebraisch stabile ("stiffly accurate") implizite Runge-Kutta-Verfahren (in dieser Arbeit wird dafür die Bezeichnung IRK(DAE) verwendet) B-stabil sind. Um auf die Seltenheit der Stabilitätsprobleme im Index-1-Fall zurückzukommen, erfüllen die meisten praktisch relevanten Probleme die erwähnte Bedingung auf dem Kernel von $f_{x'}$. Alternative Resultate sind in [Higueras and Garcia-Celayeta, 1997], [Celayeta, 1998] zu finden. Hier wurde durch eine neue Formulierung der Runge-Kutta-Verfahren eine ähnliche Stabilitätsaussage bewiesen, vorausgesetzt das Image von $f_{x'}$ ist konstant.

In der vorliegenden Arbeit erfolgt die Untersuchung des Index-2-Falles. Genauer gesagt wird eine quasilineare ADGl der Art

$$A(x(t),t)x'(t) + b(x(t),t) = 0 \tag{1.2}$$

betrachtet, wobei der Nullraum von $A(x,t)$ als konstant angenommen wird. Die meisten anwendungsrelevanten Aufgaben können in der Form (1.2) geschrieben werden.

Bevor eine Stabilitätsuntersuchung durchgeführt werden kann, müssen die Stabilitätskonzepte klar gestellt werden. Das Kapitel 1 beschäftigt sich vorrangig mit einer allgemeinen Kontraktivitätstheorie. Nach einer Einführung über einige Grundlagen der Index-2-ADGl-Theorie wird der in [Griepentrog and März, 1986] für Index-1-Gleichungen und in [März, 1998] für den Index-2-Fall vorgeschlagene Kontraktivitätsbegriff überarbeitet bzw. erweitert. Einerseits wird er verallgemeinert, indem Kontraktivität statt auf der Lösungsmannigfaltigkeit auf einer möglicherweise größeren Mannigfaltigkeit betrachtet wird. Diese Erweiterung ist auf die Erkenntnisse aus den Diskretisierungsverfahren zurückzuführen. Nach diesen Erkenntnissen ist die Zugehörigkeit der diskreten Lösung zur Lösungsmannigfaltigkeit nicht gesichert, was eine Betrachtung außerhalb dieser Mannigfaltigkeit sinnvoll macht. Dieser neue Kontraktivitätsbegriff ist nicht indexgebunden, und es stellt sich heraus, dass die bekannten Konzepte aus [Griepentrog and März, 1986], [Hanke et al., 1998] und [März, 1998] als seine Sonderfälle betrachtet werden können. Die hier eingeführten Kontraktivitätsbegriffe beziehen sich zum einen, wie die früheren Begriffe, auf die differenzierbaren Komponenten der Lösung (die P-Komponente, wobei $P = I - Q$ und Q ein Projektor auf $\ker f_{x'}$ ist) und zum anderen auf die gesamte Lösung.

Aus der Sicht der numerischen Methoden werden analog zu dem kontinuierlichen Fall P-Kontraktivi -tät und Kontraktivität definiert. Diese Konzepte erinnern an die B-Stablität, beziehen sich aber ausschließlich auf die numerische Rekursion, ohne eine Annahme über die Kontraktivität der Gleichung zu ihr in Verbindung zu setzen. Am Ende des Kapitels 1 wird ein allgemeines Theorem bewiesen, das hinreichende Bedingungen für die (P-)Kontraktivität der Anwendung eines IRK(DAE)-Verfahrens auf eine ADGl stellt. Diese Bedingungen lauten

1. algebraische Stabilität des Verfahrens

2. Kontraktivität der Gleichung auf der Mannigfaltigkeit, auf der die Runge-Kutta-Stufen liegen.

Dieses Ergebnis trägt den allgemeinen Charakter des Kapitels und enthält als Sonderfall die schon bekannten Aussagen für gewöhnliche, Index-1 und Index-2-Gleichungen. Dieses Theorem wird als wichtiges Werkzeug im Verlauf der Arbeit verwendet, um konkretere Ergebnisse zu erzielen.

Ziel des zweiten Kapitels ist eine genauere Untersuchung der BDF- und IRK-Verfahren. Dafür wird das gleiche Prinzip wie in [Griepentrog and März, 1986] für Index-1-Probleme, in [Hanke et al., 1998] für den Index-2-Fall und in [Eich-Soellner and Führer, 1998] für mechanische Systeme verwendet. Die Grundidee ist in der Abbildung 1.1 dargestellt.

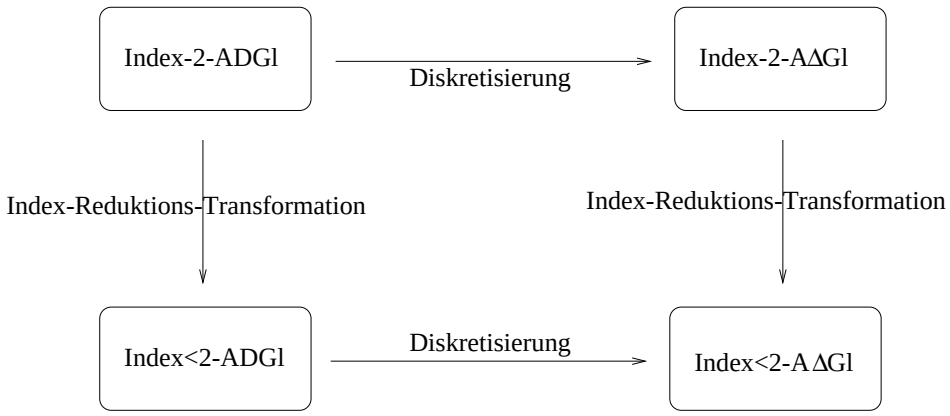

Abbildung 1.1: Die Idee der Analyse im Kapitel 2

Es werden im Rahmen des Gerüstes der Abbildung 1.1 die Entkopplungen des ADGl-Systems in GDGln und algebraischen Gleichungen sowie die Index-Reduktion durch Differentiation betrachtet. Zur Vereinfachung der Darstellung werden diese beiden Prozeduren unter dem Namen Index-Reduktions-Transformation (IRT) zusammengefasst. Wenn das Diagramm 1.1 gilt, also die zwei Wege, die von der Index-2-Gleichung zu der diskretisierten GDGl bzw. Index-1-Gleichung führen, äquivalent sind, dann ist es, als ob die GDGl bzw. Index-1-Gleichung direkt mit dem Verfahren diskretisiert wird. In diesem Fall behält die Diskretisierung ihre für die indexreduzierte Gleichung bekannten Stabilitätseigenschaften.

In [Hanke et al., 1998] wurde bereits mit dem genannten Ansatz der Index-2-Fall untersucht. In dieser Arbeit werden die Entkopplungstechniken von [Hanke et al., 1998], [Tischendorf, 1996] verfeinert und dadurch bessere Ergebnisse erzielt. Mit der Standardprozedur sind die Autoren in [Hanke et al., 1998] zu dem Schluss gekommen, dass, wenn die index-2-
linebreak relevanten Räume N_1 und S_1 (siehe Bezeichnungen und Konventionen) konstant sind, das Kommutativitätsdiagramm 1.1 gilt. Im Gegensatz zu den Standarduntersuchungsmethoden setzen die neuen lediglich auf einen der Räume N_1 oder S_1. Dadurch ist es möglich zu zeigen, dass die Invarianz einer dieser Räume genügt. Die Relevanz der neuen Varianten ist nicht nur auf die Diskretisierungsverfahren begrenzt, sie sind auch ein gutes Mittel für analytische Untersuchungen.

Im zweiten Teil des Kapitels 2 werden die Diskretisierungsverfahren anhand der Indexreduktion durch Differentiation untersucht. Diese Prozedur wurde noch nicht im Rahmen des Untersuchungsgerüstes der Abbildung 1.1 verwendet. Aus ihr geht hervor, dass, damit ein BDF- oder ein IRK-Verfahren seine Stabilitätseigenschaften behält, die Erfüllung der "versteckten" Nebenbedingung seitens der Näherungslösung bzw. der Runge-Kutta-Stufen hinreichend ist. Wie schon erwähnt ist diese Nebenbedingung bei den Index-2-Gleichungen nur implizit vorhanden und dadurch für die Diskretisierungsmethode nicht "sichtbar". Des Weiteren wird eine bemerkenswerte Beziehung zwischen diesem Ergebnis und der Entkopplungsanalyse gezeigt: Wenn der Raum S_1 konstant ist, dann erfüllen die diskrete Lösung bzw. die Runge-Kutta-Stufen die versteckte Nebenbedingung. Hier findet das allgemeine Theorem für IRK(DAE)-Verfahren von Kapitel 1 Anwendung. Man kann garantieren, dass, wenn zusätzlich das Verfahren algebraisch stabil ist und die Gleichung (P-)Kontraktivität auf der Lösungsmannigfaltigkeit aufweist, die Anwendung des Verfahrens ebenfalls (P-)kontraktiv ist.

Die Aussagen von Kapitel 2 finden weiterhin praktische Anwendung in der Modellierung elektrischer Schaltkreise. Hier liefern die klassische und die ladungsorientierte modifizierte Knotenanalyse ADGln-Systeme wie (1.2), die entweder durch die Entkopplungsanalyse oder die Index-Reduktion durch Differentiation untersucht werden können. Unter nicht zu restriktiven Annahmen gilt bei diesen Modellen entweder die Invarianz von S_1 oder jene von N_1, [Estévez Schwarz and Tischendorf, 1998]. So stellen die Ergebnisse von Kapitel 2 eine theoretische Grundlage für die erfolgreiche Diskretisierung dieser Modelle dar.

Wenn die Kriterien von Kapitel 2 nicht erfüllt sind, wie das Beispiel 1.0.1 zeigt (für mehr Einzelheiten siehe das Beispiel 3.0.8 Anfang des zweiten Kapitels), kann es selbst bei algebraisch stabilen Runge-Kutta-Verfahren und kontraktiven Gleichungen zu explodierenden diskreten Lösungen kommen, was sich nur mit einer starken Verkleinerung der Schrittweite beheben lässt. Diese Schwierigkeiten sind relativ lange bekannt, und es gab bereits verschiedene Ansätze, um eine Stabilisierung zu erreichen [Shampine, 1986], [Campbell and Moore, 1995], [Eich-Soellner and Führer, 1998], [Eich et al., 1990]. Bei den Hessenberg-Systemen in der Mechanik versuchte man, die versteckte Nebenbedingung in der Diskretisierung zu berücksichtigen. Dieser Ansatz kann in verschiedener Art und Weise angewendet werden, wodurch mehrere Verfahrensvarianten entstehen.

Das dritte Kapitel befasst sich mit einem der Stabilisierungsansätze, der in [Gear et al., 1985] und [Eich-Soellner and Führer, 1998] für Hessenberg-Systeme aus zwei verschiedenen Perspektiven betrachtet wurde. Dieser Ansatz ist unter dem Namen Gear-Gupta-Leimkuhler-Formulierung (GGL) bekannt und wird in [Eich et al., 1990], [Eich-Soellner and Führer, 1998] auch als Ableitungsprojektionsansatz bezeichnet. Anhand des Beispiels 1.0.1 wird diese Technik mit dem Koordenatenprojektionsansatz, [Eich-Soellner and Führer, 1998], verglichen. Die numerischen Ergebnisse zeigen, dass der zuletzt genannte Ansatz das Stabilitätsproblem nicht löst und dass die GGL-Formulierung zumindest in diesem Fall die Unstabilitäten behebt.

Es wird die Meinung vertreten, dass die GGL-Formulierung der genannten Ansätze die einzige ist, die zu einer Lösung der Stabilitätsprobleme führen kann. Diese These wird im weiteren Verlauf des Kapitels belegt.

Als Nächstes wird die GGL-Formulierung für voll-implizite lineare Aufgaben verallgemeinert und es werden einige gute Eigenschaften gezeigt, die vollkommen mit der Situation im Hessenberg-Fall im Einklang stehen. Aber die

wichtigste Eigenschaft ist zweifellos die Übertragung der (P-)Kontraktivität auf der Lösungsmannigfaltigkeit bei der ursprünglichen Gleichung auf eine (P-)Kontraktivität auf der "sichtbaren" Lösungsmannigfaltigkeit bei der GGL-Formulierung, wenn die GGL adäquat gewählt wird. Dieser Fakt stellt das Erfolgsgeheimnis des Ansatzes dar. Folglich ist zumindest für die IRK(DAE)-Verfahren, laut dem allgemeinen Theorem von Kapitel 1, der Stabilisierungseffekt gesichert. Anschließend werden die BDF-Verfahren wie im Kapitel 2 mittels der Entkopplungstechniken untersucht. Die Analyse zeigt Folgendes: Wenn die Anwendung des BDF-Verfahrens auf die entkoppelten Gleichungen in einem Schritt eine Kontraktion in einer euklidischen Norm darstellt und die GGL-Formulierung die Orthogonalitätsbedingung wie für die Übertragung der (P-)Kontraktivität erfüllt, dann ist die Anwendung des Verfahrens auf die GGL-Formulierung in dem entsprechenden Schritt auch eine Kontraktion. Diese Aussage für die BDF-Diskretisierung kann als das Pendant der IRK(DAE)-Aussage angesehen werden.

Kapitel 2

Stabilitätskonzepte

Ziel dieses Kapitels ist die Einführung der Stabilitätsbegriffe, deren Beziehungen in dieser Dissertation untersucht werden. Es werden zwei parallele Linien verfolgt. Bei der ersten handelt es sich um die Stabilität der analytischen Lösung, und bei der zweiten um die Stabilität einer diskreten Approximation. In den restlichen Kapiteln wird auf die wichtige Beziehung zwischen der analytischen und diskreten Stabilität für eine ADGl eingegangen. Aber bevor dies technisch möglich ist, werden einige grundlegende Konzepte der ADGl-Theorie eingeführt.

2.1 Algebro-Differentialgleichungen

Ein System von ADGln ist eins der Art, [März, 1995],

$$f(x'(t), x(t), t) = 0, \qquad f : \mathbb{R}^m \times D \times \Im \to \mathbb{R}^m, \tag{2.1}$$

wobei $\Im$ ein Intervall in $\mathbb{R}$, $x : \Im \to \mathbb{R}^m$ und D eine offene Menge sind. Es werden stetige Funktionen $f(y, x, t)$ in $D_f := \mathbb{R}^m \times D \times \Im$ betrachtet, bei denen die Jacobi-Matrizen f_y, f_x in dem Definitionsbereich D_f existieren und zugleich stetig sind. Der entscheidende Unterschied zu einem gewöhnlichen Differentialgleichungssystem besteht darin, dass die führende Matrix f_y singulär in D_f ist. Weiterhin schränkt man sich auf Gleichungen ein, bei denen der Raum

$$N(y, x, t) := \ker f_y(y, x, t),\ (y, x, t) \in D_f$$

konstant ist. Die Analyse wird sich auf quasilineare Gleichungen der Form

$$A(x(t), t)x'(t) + b(x(t), t) = 0, \quad t \in \Im, \tag{2.2}$$

konzentrieren. Sowohl die Bedingung an N als auch die betrachtete Familie von ADGln (2.2) sind aus praktischer Sicht vertretbar. Die meisten praktisch relevanten ADGln erfüllen diese Voraussetzung, [Tischendorf, 1996], [Estévez Schwarz and Tischendorf, 1998], [Estévez Schwarz, 2000], [Eich-Soellner and Führer, 1998].

In der Darstellung kennzeichnet Q einen Projektor auf den Raum N, und außerdem wird der ergänzende Projektor $I - Q$ als P gekennzeichnet.

Beispiel 2.1.1 *Ein typischer Fall von (2.2) ist das so genannte semiexplizite System*

$$\begin{aligned} x_1'(t) + b_1(x_1(t), x_2(t), t) &= 0, \\ b_2(x_1(t), x_2(t), t) &= 0, \end{aligned} \tag{2.3}$$

mit $x_1(t)$, $b_1(y, x, t) \in \mathbb{R}^r$ *und* $b_2(y, x, t) \in \mathbb{R}^{m-r}$. *In diesem Fall sieht man direkt, dass das System aus differentialen und algebraischen Gleichungen besteht. Der Notation von (2.1) entsprechend ist in diesem Falle*

$$f_y(y, x, t) = A(x, t) = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}, \qquad N = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^m : x_1 = 0 \right\}$$

und ein möglicher Projektor auf N *ist*

$$Q = \begin{pmatrix} 0 & 0 \\ 0 & I_{m-r} \end{pmatrix}.$$

In dem vorherigen Beispiel kommt die Ableitung von $x_2(t)$ nicht vor. Im Grunde genommen muss sie nicht unbedingt existieren. Im Allgemeinen kann man den führenden Term in (2.2) folgendermaßen umschreiben:

$$A(x(t), t)x'(t) = A(x(t), t)(P + Q)x'(t) = A(x(t), t)(Px)'(t). \tag{2.4}$$

Infolgedessen wird in [Griepentrog and März, 1986], [März, 1995] der Funktionsraum

$$C_N^1(\mathfrak{I}, \mathbb{R}^m) := \left\{ x \in C(\mathfrak{I}, \mathbb{R}^m) : Px \in C^1(\mathfrak{I}, \mathbb{R}^m) \right\}$$

und nicht $C^1(\mathfrak{I}, \mathbb{R}^m)$ als Lösungsraum betrachtet. Diesem Fakt zufolge wird die Formulierung (2.4) übernommen.

2.1.1 Lineare Algebro-Differentialgleichungen

Einen wichtigen Schritt für das Verständnis von ADGln stellen die linearen Systeme dar. Die lineare Variante der ADGl (2.2) kann als

$$A(t)(Px)'(t) + B(t)x(t) = q(t) \tag{2.5}$$

geschrieben werden. Hier ist $f(y, x, t)$ als

$$f(y, x, t) := A(t)y + B(t)x - q(t)$$

definiert, und die Voraussetzungen an (2.1) bedeuten in diesem Kontext, dass A, $B : \Im \to L(\mathbb{R}^m)$, $q : \Im \to \mathbb{R}^m$ stetige Funktionen sind.

In der ADGln-Theorie ist der Index-Begriff unabdingbar. Die Arbeit wird sich auf den Tractability-Index stützen. Dieser Index-Begriff, der von E. Griepentrog und R. März eingeführt wurde ([Griepentrog and März, 1986]), basiert auf linearen Aufgaben wie (2.5) und hat, im Vergleich zu anderen Index-Begriffen (wie der Differentiations-Index [Campbell, 1985], [Campbell, pear], [Brenan et al., 1989] und der geometrische Index [Rheinboldt, 1984], [Reich, 92]), den Vorteil, dass eine niedrigere Glattheit der Gleichung (2.5), bzw. (2.1), verlangt wird.

Neben N ist ein anderer Raum von Bedeutung, nämlich jener, der die Lösungen der homogenen Gleichung

$$A(t)(Px)'(t) + B(t)x(t) = 0 \tag{2.6}$$

enthält. Sei

$$S(t) := \{z \in \mathbb{R}^m : B(t)z \in \operatorname{im} A(t)\}\,.$$

Offensichtlich gilt für jede Lösung der homogenen Gleichung (2.6) $x(t) \in S(t)$.

Definition 2.1.2 *Die ADGl (2.5), mit singulärer Führungsmatrix $A(t)$ und konstantem* $\ker A(t)$*, besitzt den Tractability-Index 1 auf dem Intervall $\Im$, g.d.w. eine der folgenden äquivalenten Bedingungen erfüllt ist:*

- $N \cap S(t) = \{0\} \qquad \forall t \in \Im$
- $N \oplus S(t) = \mathbb{R}^m \qquad \forall t \in \Im$
- die Matrix $G_1(t) := A(t) + B(t)Q$ ist regulär in $\Im$ für einen beliebigen Projektor Q auf N.

Beispiel 2.1.3 *Man betrachtet die lineare Variante von Beispiel 2.1.1*

$$\begin{aligned} x_1'(t) + B_{11}(t)x_1(t) + B_{12}(t)x_2(t) &= q_1(t), \\ B_{21}(t)x_1(t) + B_{22}(t)x_2(t) &= q_2(t). \end{aligned}$$

Hier sind die Räume N und Q durch

$$N = \left\{ \begin{pmatrix} 0 \\ z_2 \end{pmatrix} : z_2 \in I\!R^{m-r} \right\} \quad \text{und} \quad Q = \begin{pmatrix} 0 & 0 \\ 0 & I_{m-r} \end{pmatrix}$$

gegeben, und

$$S(t) = \left\{ \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in I\!R^m : B_{21}(t)z_1 + B_{22}(t)z_2 = 0 \right\}.$$

Die Schnittmenge $N \cap S(t)$ ist dann

$$N \cap S(t) = \left\{ \begin{pmatrix} 0 \\ z_2 \end{pmatrix} : z_2 \in I\!R^{m-r} \wedge B_{22}(t)z_2 = 0 \right\},$$

und die Index-1-Bedingung lautet: $B_{22}(t)$ ist regulär in $\Im$. Wenn dies der Fall ist, kann man die algebraische Gleichung nach x_2 lösen, diese Variable in die erste Gleichung einsetzen und eine GDGl für x_1 aufstellen.

Andererseits ist die Matrix $G_1(t)$

$$G_1(t) = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} B_{11}(t) & B_{12}(t) \\ B_{21}(t) & B_{22}(t) \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & I_{m-r} \end{pmatrix} = \begin{pmatrix} I_r & B_{12}(t) \\ 0 & B_{22}(t) \end{pmatrix},$$

und man erhält die gleiche Bedingung für $B_{22}(t)$.

Bemerkung 2.1.4 *Für die Äquivalenz der drei Bedingungen wird auf das Theorem A.13 in [Griepentrog and März, 1986] verwiesen.*

Es gibt eine Sonderwahl des Projektors Q, die sich von den anderen abhebt. In dem Index-1-Fall gilt nach der Definition die $I\!R^m$-Zerlegung

$$I\!R^m = N \oplus S(t).$$

So kann man den Projektor Q längs $S(t)$ wählen. Dieser Projektor - in der Literatur als kanonisch bezeichnet, [Griepentrog and März, 1986], [März, 1993]

- wird üblicherweise durch $Q_c(t)$ repräsentiert, und nach dem Lemma A.14 von [Griepentrog and März, 1986] gilt

$$Q_c(t) = QG_1^{-1}(t)B(t),$$

für einen beliebigen Projektor Q auf N. An dieser Stelle ist zu betonen, dass, obwohl N konstant ist, der kanonische Projektor $Q_c(t)$ im Allgemeinen zeitabhängig ist. Das kann ein Nachteil werden, aber dafür vereinfacht sich die Entkopplung [Griepentrog and März, 1986]. Zusätzlich steht dieser Projektor in enger Beziehung zu dem Lösungsraum im Index-1-Fall, [Griepentrog and März, 1986].

Falls $G_1(t)$ immer singulär ist, dann werden neue ähnlich gebaute Räume und Projektoren eingeführt. Die Gleichung (2.6) kann folgendermaßen umgeschrieben werden:

$$\begin{aligned} (A+BQ)(t)\{P(Px)' + Qx\}(t) + B(t)Px(t) &= 0, \\ G_1(t)\{P(Px)' + Qx\}(t) + B_1(t)x(t) &= 0, \end{aligned} \tag{2.7}$$

wobei $B_1(t) := B(t)P$.

Seien

$$\begin{aligned} N_1(t) &:= \ker G_1(t), \\ S_1(t) &:= \{z \in I\!R^m : B(t)Pz \in \operatorname{im} G_1(t)\}, \end{aligned}$$

$Q^*_{N_1}(t) := Pr(N_1, *)$ ein Projektor auf $N_1(t)$ und $P^{N_1}_*(t) = I - Q^*_{N_1}(t)$. Analog zum Index-1-Fall sieht man, dass auch $S_1(t)$ die Lösungen der homogenen Gleichung enthält.

Das folgende Lemma wird mehrmals in dieser Arbeit verwendet:

Lemma 2.1.5 *Eine Matrix der Form $I + MN$, wobei $NM = 0$, ist immer regulär, und ihre Inverse ist durch $I - MN$ gegeben.*

Beweis: E s genügt zu überprüfen, dass

$$\begin{aligned} (I+MN)(I-MN) &= I, \\ (I-MN)(I+MN) &= I, \end{aligned}$$

gelten.
Die folgende Aussage ist sehr wichtig für ein Verständnis des Tractability-Indexes.

Lemma 2.1.6 *Für $N_1(t)$ gelten die Identitäten*

1. $N_1(t) = [I - PA^+(t)B(t)Q]\,(N \cap S(t))$,
2. $\dim N_1(t) = \dim(N \cap S(t))$.

Beweis: Z uerst wird für 1. die $\supset$ Inklusion gezeigt. Sei $y \in N \cap S(t)$. Sei $x \in [I - PA^+(t)B(t)Q]\,(N \cap S(t))$, dann gibt es ein $y \in N \cap S(t)$, so dass $x = [I - PA^+(t)B(t)Q]\,y$. Man berechnet $G_1(t)x$:

$$\begin{aligned} G_1(t)x &= G_1(t)\left[I - PA^+(t)B(t)Q\right]y = G_1(t)y - A(t)A^+(t)B(t)Qy \\ &= By - A(t)A^+(t)B(t)y = (I - AA^+)(t)B(t)y = 0. \end{aligned}$$

Hier wurde die Tatsache verwendet, dass $y \in N \cap S(t)$, also $y = Qy$ und $B(t)y \in \operatorname{im} A(t)$.

Für die verbleibende Inklusion wird folgendermaßen fortgefahren. Sei $x \in N_1(t)$, das heißt:

$$G_1(t)x = 0 \Leftrightarrow (A(t) + B(t)Q)x = 0.$$

Die letzte Gleichung wird mit $PA^+(t)$ skaliert und man bekommt

$$\begin{aligned} Px + PA^+(t)B(t)Qx &= 0, \\ P(I + PA^+BQ)(t)x &= 0. \end{aligned}$$

Folglich $y := (I + PA^+BQ)(t)x \in N$ und $x = (I - PA^+BQ)(t)y$. Jetzt muss nur noch gezeigt werden, dass y auch zu $S(t)$ gehört. Dafür wird wieder die Gleichung $G_1(t)x = 0$ benutzt, woraus folgt

$$\begin{aligned} (A(t) + B(t)Q)(I - PA^+BQ)(t)y &= 0, \\ B(t)y + A(t)(I - PA^+BQ)(t)y &= 0, \end{aligned}$$

und damit ergibt sich die Zugehörigkeit von y auch zu $S(t)$.

Die zweite Aussage des Lemmas folgt nun aus der Regularität von $(I - PA^+BQ)(t)$ (Lemma 2.1.5).

Definition 2.1.7 *Die ADGl (2.5) besitzt den Tractability-Index 2, gdw.* $\dim(N \cap S(t)) = \nu > 0$ *konstant in* $\Im$ *ist und eine der folgenden äquivalenten Bedingungen erfüllt ist:*

- $N_1(t) \cap S_1(t) = \{0\}, \qquad \forall t \in \Im$

- $N_1(t) \oplus S_1(t) = I\!R^m, \qquad \forall t \in \Im$
- die Matrix $G_{2,*}(t) := G_1(t) + B_1(t)Q^*_{N_1}(t)$ ist regulär in $\Im$, für einen beliebigen Projektor $Q^*_{N_1}(t)$ auf $N_1(t)$.

Bemerkung 2.1.8 *Wie in dem Index-1-Fall folgt die Äquivalenz zwischen den Bedingungen durch das Theorem A.13 von [Griepentrog and März, 1986]. Hier wird das Theorem auf das Matrix-Büschel (G_1, B_1) angewendet.*

So ein allgemeiner Projektor $Q^*_{N_1}(t)$ wie in der Definition 2.1.7 reicht für die Zwecke dieser Arbeit nicht aus. Es wird in erster Linie gefordert, dass die Identität $Q^*_{N_1}(t)Q = 0\,\forall t \in \Im$ gilt. Es ist möglich, den Projektor $Q^*_{N_1}$ so zu wählen, weil

$$N \subset S_1(t),\ S_1(t) \cap N_1(t) = \{0\},\ N \cap N_1(t) = \{0\}.$$

Diese Voraussetzung macht die Ausdrücke $PP^{N_1}_*(t)$ und $PQ^*_{N_1}(t)$ wiederum zu Projektoren.

Beispiel 2.1.9 *Die am besten verstandenen Index-2-Systeme sind womöglich die in Hessenberg-Form*

$$\begin{aligned} x_1'(t) + B_{11}(t)x_1(t) + B_{12}(t)x_2(t) &= q_1(t), \\ B_{21}(t)x_1(t) &= q_2(t), \end{aligned}$$

wobei die Matrix $B_{21}(t)B_{12}(t)$ in $\Im$ regulär ist. Man berechnet jetzt die relevanten Unterräume und Projektoren für dieses System.

$$N = \left\{ \begin{pmatrix} 0 \\ z_2 \end{pmatrix} : z_2 \in I\!R^{m-r} \right\}, \quad Q = \begin{pmatrix} 0 & 0 \\ 0 & I_{m-r} \end{pmatrix},$$

$$S(t) = \left\{ \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in I\!R^m : B_{21}(t)z_1 = 0 \right\} \quad \textit{und} \quad N \cap S(t) = N.$$

Jetzt wird die Index-2-Bedingung überprüft,

$$G_1(t) = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} B_{11}(t) & B_{12}(t) \\ B_{21}(t) & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & I_{m-r} \end{pmatrix} = \begin{pmatrix} I_r & B_{12}(t) \\ 0 & 0 \end{pmatrix},$$

$$N_1(t) = \left\{ \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in I\!R^m : z_1 = -B_{12}(t) z_2 \right\},$$

und als $Q^*_{N_1}(t)$ *kann man z.B.*

$$Q^*_{N_1}(t) = \begin{pmatrix} 0 & -B_{12}(t) \\ 0 & I_{m-r} \end{pmatrix},$$

nehmen.

Auf der anderen Seite ist der Raum $S_1(t)$ *durch*

$$\begin{aligned} S_1(t) &= \left\{ \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in I\!R^m : \begin{pmatrix} B_{11}(t) & B_{12}(t) \\ B_{21}(t) & 0 \end{pmatrix} \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \right. \\ &\quad \left. \in \operatorname{im} \begin{pmatrix} I_r & B_{12}(t) \\ 0 & 0 \end{pmatrix} \right\}, \\ &= \left\{ \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in I\!R^m : \begin{pmatrix} B_{11}(t) z_1 \\ B_{21}(t) z_1 \end{pmatrix} \in \operatorname{im} \begin{pmatrix} I_r & B_{12}(t) \\ 0 & 0 \end{pmatrix} \right\}, \\ &= \left\{ \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in I\!R^m : B_{21}(t) z_1 = 0 \right\}, \end{aligned}$$

gegeben. Schließlich ist die entscheidende Schnittmenge

$$N_1(t) \cap S_1(t) = \left\{ \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in I\!R^m : z_1 = -B_{12}(t) z_2 \wedge B_{21}(t) z_1 = 0 \right\},$$

und wegen der Regularität von $B_{21}(t)B_{12}(t)$ *folgt* $N_1(t) \cap S_1(t) = \{0\}$.

2.1.2 Nicht-lineare Algebro-Differentialgleichungen

Es wird hier der Darstellung von [Griepentrog and März, 1986] und [März, 1995] in weitem Sinne gefolgt und nun erneut die nicht-linearen AD-Gln aufgenommen. Man betrachtet zunächst ein nicht-lineares ADGl-System wie (2.1), also

$$f(x'(t), x(t), t) = 0, \quad t \in \Im, \tag{2.8}$$

unter den dort genannten Voraussetzungen. Man ist eigentlich an einem System wie (2.2) interessiert, aber für die Definition des Tractabilitiy-Indexes ist es günstiger, die Gleichung (2.8) zu betrachten.

Der Tractability-Index

Dieser Index-Begriff für (2.8) basiert auf der Linearisierung dieser Gleichung. Der erste relevante Unterraum N und die jeweiligen Projektoren Q und $P := I - Q$ wurden schon eingeführt. Der andere Index-1-relevante Unterraum S und die Matrix G_1 werden in einem Punkt $(y, x, t) \in D_f$ als

$$\begin{aligned} S(y,x,t) &:= \{z \in I\!R^m : f_x(y,x,t)z \in \operatorname{im} f_y(y,x,t)\}, \\ G_1(y,x,t) &:= f_y(y,x,t) + f_x(y,x,t)Q, \end{aligned}$$

definiert.

Definition 2.1.10 *([März, 1995]) Die ADGl (2.8) besitzt den Tractability-Index-1 in einer offenen Menge $G \subset D_f$, falls*

$$N \oplus S(y,x,t) = I\!R^m, \ \forall (y,x,t) \in G. \tag{2.9}$$

Bemerkung 2.1.11 *Die Bedingung ist wie in dem linearen Fall äquivalent zu der Regularität von $G_1(y,x,t)$ (Theorem A. 13 von [Griepentrog and März, 1986]).*

Wenn die Aufgabe die Index-1-Bedingung nicht erfüllt, geht man folgendermaßen vor: Es werden für $(y, x, t) \in D_f$ die Unterräume

$$\begin{aligned} N_1(y,x,t) &:= \ker G_1(y,x,t), \\ S_1(y,x,t) &:= \{z \in I\!R^m : f_x(y,x,t)Pz \in \operatorname{im} G_1(y,x,t)\}, \end{aligned}$$

und die Matrix $G_{2,*}$ als

$$G_{2,*}(y,x,t) := G_1(y,x,t) + f_x(y,x,t)PQ^*_{N_1}(y,x,t)$$

definiert. Dabei ist $Q^*_{N_1}(y,x,t) := Pr(N_1, *)$ ein Projektor auf $N_1(y,x,t)$ und $P^{N_1}_*(y,x,t) := I - Q^*_{N_1}(y,x,t)$.

Definition 2.1.12 *([März, 1995]) Die ADGl (2.8) besitzt den Tractability-Index-2 in einer offenen Menge $G \subset D_f$, wenn die Bedingungen*

1. $\dim N_1(y,x,t)$ ist konstant und größer als Null,
2. $N_1(y,x,t) \oplus S_1(y,x,t) = I\!R^m, \ \forall\, (y,x,t) \in G$

erfüllt sind.

Bemerkung 2.1.13 *Wieder ist die Regularität der Matrix $G_2(y,x,t)$ in G eine äquivalente Bedingung zu 2. in der Definition 2.1.12.*

Beispiel 2.1.14 *Man betrachtet das semiexplizite System*

$$\begin{aligned} x_1'(t) &= g_1(x_1(t), x_2(t), t), \\ 0 &= g_2(x_1(t), t), \end{aligned} \tag{2.10}$$

und berechnet die relevanten Matrizen. Die partiellen Ableitungen f_y und f_x sind in diesem Fall

$$f_y(y,x,t) = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}, \quad f_x(y,x,t) = \begin{pmatrix} -g_{1x_1}(x_1,x_2,t) & -g_{1x_2}(x_1,x_2,t) \\ -g_{2x_1}(x_1,t) & 0 \end{pmatrix},$$

und die erste relevante Matrix $G_1(y,x,t)$ ist durch

$$\begin{aligned} G_1(y,x,t) &= \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} g_{1x_1}(x_1,x_2,t) & g_{1x_2}(x_1,x_2,t) \\ g_{2x_1}(x_1,t) & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix} \\ &= \begin{pmatrix} I & -g_{1x_2}(x_1,x_2,t) \\ 0 & 0 \end{pmatrix} \end{aligned}$$

gegeben. Der Nullraum von $G_1(x,t)$ hat für alle (x,t) eine konstante positive Dimension. Folglich besitzt die Gleichung einen höheren Index. Eine Fortsetzung der Berechnung mit dem Projektor

$$Q_{N_1}^*(y,x,t) = \begin{pmatrix} 0 & g_{1x_2}(x_1,x_2,t) \\ 0 & I \end{pmatrix}$$

auf $N_1(x,t)$ ergibt

$$\begin{aligned} G_{2,*}(y,x,t) &= \begin{pmatrix} I & -g_{1x_2}(x_1,x_2,t) \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} g_{1x_1}(x_1,x_2,t) & g_{1x_2}(x_1,x_2,t) \\ g_{2x_1}(x_1,t) & 0 \end{pmatrix} \times \\ &\quad \begin{pmatrix} 0 & g_{1x_2}(x_1,x_2,t) \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} I & -(I + g_{1x_1}(x_1,x_2,t))g_{1x_2}(x_1,x_2,t) \\ 0 & -(g_{2x_1}g_{1x_2})(x_1,x_2,t) \end{pmatrix}, \end{aligned}$$

woraus die bekannte Hessenberg-Index-2-Bedingung $(g_{2x_1}g_{1x_2})(x_1,x_2,t)$ regulär folgt.

Die Index-2-Lösungsmannigfaltigkeit

Als Letztes, bevor man die Stabilitätsproblematik angehen kann, sind die Kenntnisse über Index-2-ADGln zu vertiefen. Die folgenden Fakten sind vor allen Dingen für das Kontraktivitätskonzept unabdingbar.

Es wird eine Gleichung wie (2.2) betrachtet,

$$A(x(t),t)(Px)'(t) + b(x(t),t) = 0, \quad t \in \Im. \tag{2.11}$$

Die Voraussetzungen bedeuten jetzt, dass $A(x,t)$, $b(x,t)$, $A_x(x,t)$ und $b_x(x,t)$ stetige Funktionen in D_f sind. Für (2.11) wird angenommen, dass der Tractability-Index-2 in einer offenen Menge $G \subset D_f$ vorliegt.

Es wird darauf aufmerksam gemacht, dass sich aus (2.11) der Ausdruck

$$(Px)'(t) = -(PA^+b)(x(t),t) \tag{2.12}$$

ergibt.

Weiterhin wird die Matrix

$$B(y,x,t) := f_x(y,x,t) = b_x(x,t) + (A(x,t)y)_x$$

definiert, wodurch

$$G_1(y,x,t) = A(x,t) + B(y,x,t)Q$$

gilt. Außerdem wird $W_0(x,t)$ einen Projektor längs im $A(x,t)$ repräsentieren, zum Beispiel $I - A(x,t)A^+(x,t)$.

Jetzt kann man die so genannte sichtbare Lösungsmannigfaltigkeit aufschreiben. Da für alle Lösungen $b(x(t),t) \in \operatorname{im} A(x(t),t)$ gelten muss, bekommt man

$$M_0(t) = \{x \in D : b(x,t) \in \operatorname{im} A(x,t)\} = \{x \in D : W_0(x,t)b(x,t) = 0\}. \tag{2.13}$$

Das erinnert an den Unterraum $S(y,x,t)$, der eigentlich der Tangentialraum zu $M_0(t)$ in dem Punkt x ist.

Ab hier wird angenommen, dass der Unterraum $\operatorname{im} A_1(y,x,t)$ nur von Px und t abhängt und eine in beiden Argumenten stetig differenzierbare Basis dieses Raumes existiert. Unter diesen Voraussetzungen gibt es einen C^1-Projektor W_1 längs $\operatorname{im} A_1(Px,t)$, der nur von (Px,t) abhängt.

Durch die Multiplikation von (2.2) mit $W_0(x(t),t)$ erhält man eine ableitungsfreie Gleichung. Oft wird diese Gleichung nach t differenziert, um die

Index-Reduktion zu erreichen. Es wird aber gezeigt, dass noch weitere Teile aus der Gleichung (2.2) ausgeschnitten werden können. Das ist mit dem Projektor W_1 möglich:

$$\ker W_1(Px,t) = \operatorname{im} A_1(Px,t) \supset \operatorname{im} A(x,t) = \ker W_0(x,t).$$

Die Gleichung (2.11) ist äquivalent zu

$$\begin{aligned} A(x(t),t)(Px)'(t) + (I - W_1(x(t),t))b(x(t),t) &= 0, \qquad (2.14)\\ W_1(x(t),t)b(x(t),t) &= 0. \qquad (2.15) \end{aligned}$$

An dieser Stelle muss man annehmen, dass der Teil $W_1(x,t)b(x,t)$ nach t differenzierbar ist. Für ein lineares System bedeutet dies, dass $W_1(t)B(t)$ und $W_1(t)q(t)$ differenzierbar sind und das sind hinreichende Bedingungen für die Lösbarkeit der Gleichung, [März and Rodríguez Santiesteban, 1999].

Lemma 2.1.15 *Wenn* $\operatorname{im} A_1$ *nur von* Px *und* t *glatt abhängt, dann ist* $W_1 b$ *auch nur von* Px *und* t *abhängig.*

Beweis: A us $W_1(Px,t)G_1(y,x,t) = 0$ folgt $W_1(Px,t)B(y,x,t)Q = 0$, und das bedeutet

$$W_1(Px,t)b_x(x,t)Q = -W_1(Px,t)\left(A(x,t)y\right)_x Q.$$

Außerdem gilt für ein beliebiges $z \in N$

$$W_1(Px,t)\left(A(x,t)y\right)_x z = \left[W_1(Px,t)A(x,t)y\right]_x - W_{1x}(Px,t)zA(x,t)y = 0,$$

was mit sich bringt, dass

$$\left[W_1(Px,t)b(x,t)\right]_{Qx} = W_1(Px,t)b_x(x,t)Q = 0.$$

Die Aussage dieses Lemmas ist, dass man unter den Voraussetzungen (2.15) für $x(t) \in C_N^1(\Im, I\!R^m)$ nach t ableiten kann, woraus man eine weitere Bedingung für jede Lösung von (3.89) erhält,

$$(W_1 b)_x(x(t),t)(Px)'(t) + (W_1 b)_t(x(t),t) = 0.$$

Unter Betrachtung von (2.15) folgt aus der letzten Gleichung

$$W_1(x(t),t)\left\{(W_1 b)_x(x(t),t)(Px)'(t) + (W_1 b)_t(x(t),t)\right\} = 0. \qquad (2.16)$$

Nach Einsetzen von $(Px)'(t)$ durch (2.12) erhält man schließlich

$$W_1(x(t),t)\left\{-(W_1b)_x(x(t),t)(A^+b)(x(t),t)+(W_1b)_t(x(t),t)\right\}=0. \quad (2.17)$$

Es wird die versteckte Nebenbedingungsmannigfaltigkeit als

$$H(t):=\left\{x\in D: W_1(x,t)\left\{-(W_1b)_x(x,t)(A^+b)(x(t),t)+(W_1b)_t(x,t)\right\}=0\right\} \quad (2.18)$$

definiert und die Lösungsmannigfaltigkeit ist dann durch

$$M_1(t):=M_0(t)\cap H(t) \quad (2.19)$$

gegeben.

2.2 Stabilität im kontinuierlichen Fall

Ziel dieser Arbeit ist die Untersuchung des asymptotischen Verhaltens numerischer Verfahren für Index-2-ADGln. Dabei stellt man sich die Frage, ob ein numerisches Verfahren gewisse analytische Stabilitätseigenschaften der Lösung widerspiegelt. In diesem Abschnitt werden die asymptotischen Eigenschaften definiert, an denen man interessiert ist.

2.2.1 Lyapunov-Stabilität

Sowohl für die Anwendungen als auch in der Theorie gewöhnlicher Differentialgleichungen ist die Lyapunov-Stabilitätstheorie von großer Bedeutung, [Hale, 1980], [Hairer et al., 1987]. Dabei geht es um das asymptotische Verhalten der Lösung, wenn die Anfangswerte in einer Umgebung einer gegebenen Trajektorie gestört werden. Für ADGln stellt sich natürlich die gleiche Problematik, und man kann die Lyapunov-Begriffe erweitern, [Griepentrog and März, 1986], [März, 1994], [Tischendorf, 1994].

Man kennzeichnet eine Lösung, die den Anfangswert $x(t_0)=x_0$ erfüllt, mit $x(t;t_0,x_0)$, außerdem repräsentiert $\Im_0$ ein Intervall der Form $[t_0,\infty)$.

Definition 2.2.1 *([Griepentrog and März, 1986]) Sei $x\in C^1_N(\Im_0,\mathbb{R}^m)$ die Lösung des Anfangswertproblems*

$$\begin{aligned} &f(x'(t),x(t),t)=0, \quad t\in\Im_0 \\ &x(t_0)=x_0\in M(t_0), \end{aligned} \quad (2.20)$$

wobei $M(t_0)$ die Lösungsmannigfaltigkeit für t_0 ist. Sei außerdem $U(x_0)$ eine Umgebung von x_0 auf $M(t_0)$. Dann ist die Lösung x stabil (im Lyapunovschen Sinne) in $U(t_0)$, wenn

1. *es ein $\tau > 0$ gibt, so dass jedes Anfangswertproblem*

$$\begin{aligned} &f(x'(t), x(t), t) = 0, \quad t \in \Im_0 \\ &x(t_0) = z \in U(t_0),\ \|z - x_0\| \leq \tau, \end{aligned}$$

   eindeutig lösbar in $\Im_0$ ist,

2. *es für alle $\epsilon > 0$ ein $\delta(\epsilon)$, mit $\tau \geq \delta(\epsilon) > 0$ gibt, so dass $z \in U(t_0)$ und $\|z - x_0\| \leq \delta(\epsilon)$ die Ungleichung*

$$\|x(t; t_0, z) - x(t; t_0, x_0)\| \leq \epsilon,\ \forall t \in \Im_0$$

   implizieren.

Falls dazu eine andere Konstante $\sigma > 0$ existiert, so dass für alle $z \in U(t_0)$ mit $\|z - x_0\| \leq \sigma$ die Bedingung

$$\lim_{t \to \infty} \|x(t; t_0, z) - x(t; t_0, x_0)\| = 0$$

gilt, nennt man x asymptotisch stabil (im Lyapunovschen Sinne).

Bemerkung 2.2.2

1. Ein wichtiger Sonderfall der Lyapunovschen asymptotischen Stabilität ist die so genannte exponentielle asymptotische Stabilität. In diesem Falle gilt die stärkere Bedingung

$$\|x(t; t_0, z) - x(t; t_0, x_0)\| \leq \|z - x_0\|\, e^{-\lambda(t - t_0)},\ \forall t > t_0$$

   für ein positives $\lambda \in I\!R$.

2. Für ein explizites System stimmt diese Definition mit dem gewöhnlichen Begriff überein.

3. Es ist zu unterstreichen, dass in der Definition 2.2.1 nichts über den Index der Gleichung angenommen wird. Die genannte Lösungsmannigfaltigkeit $M(t)$ kann beispielsweise im Index-2-Fall $M_1(t)$ sein, Abschnitt 2.1.2.

Besondere Aufmerksamkeit in diesem Kontext widmete man in der gewöhnlichen Differentialgleichungstheorie den autonomen Systemen. Als wichtiges Werkzeug hat man das bekannte Lyapunov-Resultat für die Stabilität einer stationären Lösung zu Verfügung, [Hale, 1980]. Für ADGln kann man analoge Ergebnisse erzielen, siehe [Griepentrog and März, 1986], [Tischendorf, 1994], [März, 1994]. Da sich die Arbeit mit Index-2-Gleichungen befasst, wird das Lyapunov-Theorem für diesen Fall vorgestellt. Im Vergleich zu dem Index-1-Fall ist bei den Index-2-Aufgaben mit zusätzlichen Schwierigkeiten zu rechnen. Der Fakt, dass die ADGl in einem Punkt die Index-2-Bedingung erfüllt, garantiert im Allgemeinen nicht die Erhaltung dieser Bedingung in einer Umgebung des Punktes. Um dies zu gewährleisten, wird eine weitere Strukturbedingung vorausgesetzt, [März, 1995], [Tischendorf, 1996].

Man betrachtet wieder eine quasilineare ADGl der Form:

$$A(x(t),t)(Px)'(t) + b(x(t),t) = 0, \tag{2.21}$$

deren Linearisierung längs einer Trajektorie $(x_*(t),t)$ den Tractability-Index-2 besitzt. Außerdem seien

$$\begin{aligned}
G_2(t) &:= G_2(x_*'(t), x_*(t), t),\\
Q_{N_1}^{S_1}(t) &:= \Pr(N_1((Px_*)'(t), x_*(t), t), S_1(x_*'(t), x_*(t), t)),\\
T(t) &:= \Pr(N \cap S(x_*(t), t), *),\\
U(t) &:= I - T(t),\\
\hat{b}(x,t) &:= (U(t)Q + PQ_{N_1}^{S_1}(t))G_2^{-1}(t)b(x,t).
\end{aligned}$$

Die Strukturbedingung fordert dann

$$Q_{N_1}^{S_1}(t)(I + \hat{b}_x'(x,t) - \hat{b}_x'(x_*(t),t))^{-1}T(t)Q = 0, \tag{2.22}$$

für alle x in einer Umgebung der Trajektorie $(x_*(t),t)$. Auf die Hintergründe dieser Bedingung wird im Abschnitt 3.1 näher eingegangen.

Man betrachtet jetzt den autonomen Fall, also das System

$$A(Px)' + b(x) = 0,\ t \in \Im_0, \tag{2.23}$$

das eine stationäre Lösung x_* besitzen soll. Systeme, für die die führende Matrix A nur von der Lösung abhängt und die konstanten führenden Nullraum aufweisen, kann man in die Form (2.23) überführen, indem die neue Variable $y := (Px)'$ eingeführt wird:

$$\begin{aligned}
(Px)' - y &= 0,\\
A(x)y + b(x) &= 0.
\end{aligned}$$

Theorem 2.2.3 *([März, 1994]) Sei $b \in C^2(D, \mathbb{R}^m)$, $D \subset \mathbb{R}^m$ offen, $x_* \in D$, $b(x_*) = 0$, $B := b_x(x_*)$. Außerdem sei das Matrix-Büschel $\lambda A + B$ regulär mit Index 2 und alle seine Eigenwerte in der linken komplexen Halbebene enthalten. Wie schon erwähnt setzt man voraus, dass die Strukturbedingung (2.22) in einer Umgebung der Trajektorie $B(x_*, \sigma)$ erfüllt ist.*

Dann gibt es für alle $\epsilon > 0$ ein $\tau > 0$ und ein $\delta(\epsilon) > 0$, so dass

- alle AWAn für (2.23) mit

$$PP_*^{N_1}(t_0)(x(t_0) - x^0) = 0, \quad \left\| PP_*^{N_1}(x^0 - x_*) \right\| < \tau$$

  auf $[t_0, \infty)$ eindeutig lösbar sind,

- $\left\| PP_*^{N_1}(t_0)(x^0 - x_*) \right\| < \delta(\epsilon)$ impliziert

$$\left\| x(t; t_0, x^0) - x_* \right\| \leq \epsilon, \quad t \geq t_0, \text{ und}$$

- $\|x(t; t_0, x^0) - x_*\| \to 0$ $(t \to \infty)$.

Bemerkung 2.2.4 *Die Aussage des Theorems ist gleichbedeutend mit der asymptotischen Stabilität der Lösung x_*. Die Anfangsbedingung*

$$PP_*^{N_1}(t_0)(x(t_0) - x^0) = 0, \quad \left\| PP_*^{N_1}(x^0 - x_*) \right\| < \tau$$

impliziert, dass die Umgebung von x_* auf M_1 liegt (siehe Abschnitt 3.1.2).

2.2.2 Kontraktivität

Der Kontraktivitätsbegriff ist in der gewöhnlichen Differentialgleichungstheorie geläufig. Er verkörpert eine strengere Bedingung als die Lyapunov-Stabilität und spielt eine wichtige Rolle in der asymptotischen Stabilitätsuntersuchung numerischer Verfahren. Diese Notion wurde bei Dahlquist eingeführt, um die A-Stabilitätsnotion für nicht-lineare Aufgaben zu verallgemeinern, [Dekker and Verwer, 1984]. Ebenso wie bei der A-Stabilität legt man damit eine Klasse von Aufgaben fest, die als Testaufgaben dienen. Die Qualität eines Verfahrens wird dann für diese Klasse von Problemen untersucht und mit anderen Verfahren für die gleiche Aufgabenklasse verglichen.

Für ADGln versuchte man, einen ähnlichen Weg zu gehen. Es gab Versuche, den A-Stabilitätsbegriff für ADGln zu definieren, [Wensch et al., 1995].

Diese Arbeiten bestätigten die schon bekannten Stabilitätsprobleme bei den ADGln. Allerdings ist die Bedeutung der geforderten Voraussetzungen, die die Testproblemklasse definieren, unklar. Nichtsdestotrotz scheint der Kontraktivitätsbegriff für ADGln vorstellbarer zu sein. In [Griepentrog and März, 1986] wurde eine Erweiterung des Kontraktivitätsbegriffes für allgemeine nicht-lineare Index-1-Aufgaben eingeführt. Dieser Begriff stellt sich, wie in dem gewöhnlichen Fall, als stärker als die Lyapunov-Stabilität heraus. Er entspricht der Kontraktivität des Zustandsvariablensystems in einem invarianten linearen Raum im gewöhnlichen Sinne, [Griepentrog and März, 1986]. Aus der Sicht der Diskretisierungsverfahren wurde ein Satz der Art "Kontraktivität plus algebraische Stabilität implizieren B-Stabilität" für bestimmte IRK-Verfahren gezeigt (die so genannten IRK(DAE)-Verfahren).

Ein erster Versuch für Index-2-Aufgaben wurde in [Hanke et al., 1998] und [Hanke et al., 1998] unternommen. In diesen Arbeiten wurde der lineare Fall betrachtet und ein entsprechendes B-Stabilitätsergebnis erzielt. Dabei musste man zusätzlich annehmen, dass die Unterräume $S_1(t)$ und $N_1(t)$ konstant sind. Im Kapitel 3 werden mit der Hilfe neuer Entkopplungsvarianten die Ergebnisse von [Hanke et al., 1998] verbessert. Eine Erweiterung des Kontraktivitätsbegriffes auf nicht-lineare Index-2-Aufgaben ist in [März, 1998] zu finden. Dieser Begriff kann als eine Weiterentwicklung von [Hanke et al., 1998] und [Hanke et al., 1998] angesehen werden und schließt Gleichungen der Form (2.2) ein. Diese Kontraktivitätsnotion wird jetzt als Ausgangspunkt der Darstellung verwendet. Die Definition, nach [März, 1998], lautet:

Definition 2.2.5 *([März, 1998]) Man betrachtet eine Index-2-Gleichung der Form (2.2) unter den Voraussetzungen von Abschnitt 2.1.2. Die Gleichung (2.2) heißt kontraktiv in $D \subset \mathbb{R}^m$, wenn ein Skalarprodukt $\langle \cdot, \cdot \rangle_S$ und eine Konstante $c \geq 0$ existieren, so dass für alle Vektoren $(y_1, x_1, t), (y_2, x_2, t) \in \mathbb{R}^m \times D \times [t_0, \infty)$ mit*

$$\begin{aligned} Qy_1 = Qy_2 = 0, \\ A(x_1,t)y_1 + b(x_1,t) = A(x_2,t)y_2 + b(x_2,t) = 0, \\ W_1(x_1,t)\left\{(W_1b)_x(x_1,t)y_1 + (W_1b)_t(x_1,t)\right\} = 0, \\ W_1(x_2,t)\left\{(W_1b)_x(x_2,t)y_2 + (W_1b)_t(x_2,t)\right\} = 0, \end{aligned}$$

die Ungleichung

$$\langle y_1 - y_2, P(x_1 - x_2)\rangle_S \leq -c \left\|P(x_1 - x_2)\right\|_S^2, \tag{2.24}$$

gilt.

Bemerkung 2.2.6

1. Die Definition ist unabhängig von der Wahl des Projektors P (siehe Lemma 2.2.9). Vom Projektor W_1 ist die Definition ebenfalls unabhängig. Die durch die sichtbare und versteckte Nebenbedingung definierte Schnittmenge hängt nicht von W_1 ab.

2. Die linearen Kontraktivitätsdefinitionen in [Hanke et al., 1998] und [März and Rodríguez Santiesteban, 1999] sind in dieser Definition enthalten.

3. In anderen Worten fordert diese Definition die Erfüllung von (2.24) auf der Lösungsmannigfaltigkeit $M_1(t)$. Allerdings gehört eine approximierte Lösung nicht zwangsläufig zu M_1. Aus diesem Grund ist es sinnvoll, Kontraktivität auf einer Obermenge von M_1 zu betrachten.

4. Die Kontraktivität bezieht sich nur auf die P-Komponente von x. Diese Tatsache hat ihren Ursprung in den Index-1-Aufgaben. Da definierte man den Kontraktivitätsbegriff nach der Kontraktivität der inhärenten Differentialgleichung für Px auf dem invarianten Unterraum im P. Für den linearen homogenen Index-2-Fall gilt die gleiche Grundidee bis auf die Tatsache, dass die inhärente Differentialgleichung jetzt für $PP_*^{N_1}x$ auf dem invarianten Unterraum im $PP_*^{N_1}$ ist. Als gemeinsamen Nenner kann man hier erkennen, dass die Kontraktivität nur für die auf jeden Fall differenzierbaren Komponenten der Lösung verlangt wird. Demzufolge beziehen sich die resultierenden Aussagen ausschließlich auf diese Komponenten.

In dieser Arbeit wird den letzten zwei Bemerkungen nachgegangen. Es wird zunächst die folgende Definition angestrebt:

Definition 2.2.7 *Man betrachtet eine Differentialgleichung der Form*

$$f(x'(t), x(t), t) = 0, \quad t \in \Im_0. \tag{2.25}$$

Die Gleichung (2.25) heißt P-kontraktiv auf der Mannigfaltigkeit $\Gamma(t) \subseteq M_0(t) \subset \mathbb{R}^m$, wenn ein Skalarprodukt $\langle \cdot, \cdot \rangle_S$ und eine Konstante $c \geq 0$ existieren, so dass für alle Vektoren $(y_1, x_1, t), (y_2, x_2, t) \in \mathbb{R}^m \times \Gamma(t) \times \Im_0$ mit

$$\begin{aligned} Qy_1 = Qy_2 = 0, \\ f(y_1, x_1, t) = f(y_2, x_2, t) = 0, \end{aligned}$$

wobei Q ein konstanter Projektor auf N ist, die Ungleichung

$$\langle y_1 - y_2, P(x_1 - x_2)\rangle_S \leq -c \left\| P(x_1 - x_2)\right\|_S^2, \tag{2.26}$$

gilt.

Bemerkung 2.2.8

1. Für eine gewöhnliche Differentialgleichung gilt $N = \{0\}$. Demzufolge, wenn $\Gamma(t) = M_0(t) = I\!R^m$, erhält man den Kontraktivitätsbegriff aus der GDGl-Theorie.

2. Es ist zu unterstreichen, dass nur Mannigfaltigkeiten Sinn machen, die in $M_0(t)$ enthalten sind. Die Erfüllung der Gleichung $f(y, x, t) = 0$ zwingt x, zumindest zu $M_0(t)$ zu gehören.

3. Diese Definition erinnert teilweise an die logarithmische Norm in [Higueras and Celayeta, 1999] und [Celayeta, 1998]. In diesen Arbeiten wird die logarithmische Norm in einem beliebigen "zulässigen" Unterraum, der insbesondere der Lösungsunterraum sein kann, betrachtet. Obwohl aus analytischer Sicht nur die Lösungsmannigfaltigkeit von Bedeutung ist, ist die Situation bei den numerischen Verfahren anders. Wie man im Kapitel 3 sehen wird, liegt die numerische Lösung für ein Index-2-Problem in einem bestimmten Zeitpunkt t_i nicht unbedingt auf $M_1(t_i)$, aber immer auf $M_0(t_i)$. Deswegen ist es entscheidend, ob Kontraktivität außerhalb von $M_1(t)$ vorliegt.

An dieser Stelle soll eine Aussage, die in den Bemerkungen der Definition 2.2.5 erwähnt wurde, bewiesen werden.

Lemma 2.2.9 *Die folgende Aussage ist unabhängig von der Wahl des konstanten Projektors Q: Es gibt eine symmetrische, positiv definite Matrix S und eine Konstante $c \geq 0$, so dass die Erfüllung von*

$$\begin{aligned} Qy_1 = Qy_2 = 0, \\ A(x_1, t)y_1 + b(x_1, t) = A(x_2, t)y_2 + b(x_2, t) = 0, \end{aligned} \tag{2.27}$$

immer

$$\langle y_1 - y_2, P(x_1 - x_2)\rangle_S \leq -c \left\| P(x_1 - x_2)\right\|_S^2 \tag{2.28}$$

impliziert.

Beweis: S ei die Aussage mit dem Projektorpaar P, Q gültig, und außerdem seien $\overline{P}$, $\overline{Q}$ ein auch mögliches Projektorpaar. Es wird angenommen, dass die Bedingungen (2.27) für $\overline{Q}$ erfüllt sind. Man versucht, die Ungleichung (2.28) für ein $\bar{c} > 0$ und eine Matrix $\overline{S}$ abzuleiten.

Als Nächstes definiert man $y_1^0 := Py_1$ und $y_2^0 := Py_2$, dann gelten die Gleichungen

$$\begin{aligned} Qy_1^0 = Qy_2^0 = 0, \\ A(x_1,t)y_1^0 + b(x_1,t) = A(x_2,t)y_2^0 + b(x_2,t) = 0, \end{aligned}$$

und auf Grund der Annahme bezüglich der Projektoren P, Q gilt

$$\begin{aligned} \left\langle y_1^0 - y_2^0, P(x_1-x_2)\right\rangle_S &\leq -c\left\|P(x_1-x_2)\right\|_S^2, \\ (y_1^0-y_2^0)^T SP(x_1-x_2) &\leq -c(x_1-x_2)^T P^T SP(x_1-x_2), \\ (y_1-y_2)^T P^T SP(x_1-x_2) &\leq -c(x_1-x_2)^T P^T SP(x_1-x_2). \end{aligned}$$

Zwischen den Projektoren P und $\overline{P}$ gilt die Beziehung

$$P = R\overline{P}R^{-1}, \quad R := \overline{Q} + P.$$

Nach dem Einsetzen dieses Ausdrucks für P in die letzte Ungleichung erhält man

$$(y_1-y_2)^T R^{-T}\overline{P}^T R^T SR\overline{P}R^{-1}(x_1-x_2) \leq -c(x_1-x_2)^T R^{-T}\overline{P}^T R^T SR\overline{P}R^{-1}(x_1-x_2).$$

Jetzt wird eine neue symmetrische positiv definite Matrix $\overline{S} := R^T SR$ definiert, und es folgt

$$\begin{aligned} \left\langle \overline{P}R^{-1}(y_1-y_2), \overline{P}R^{-1}(x_1-x_2)\right\rangle_{\overline{S}} &\leq -c\left\|\overline{P}R^{-1}(x_1-x_2)\right\|_{\overline{S}}^2, \\ \left\langle \overline{P}(y_1-y_2), \overline{P}(x_1-x_2)\right\rangle_{\overline{S}} &\leq -c\left\|\overline{P}(x_1-x_2)\right\|_{\overline{S}}^2, \\ \left\langle y_1-y_2, \overline{P}(x_1-x_2)\right\rangle_{\overline{S}} &\leq -c\left\|\overline{P}(x_1-x_2)\right\|_{\overline{S}}^2, \end{aligned}$$

und damit die Gültigkeit mit dem Projektorpaar $\overline{Q}$, $\overline{P}$.
Mit der Definition 2.2.7 ist man der dritten Bemerkung der Definition 2.2.5 entgegengekommen. Es soll sich nun der vierten Bemerkung zugewandt werden.

Definition 2.2.10 *Man betrachtet eine Differentialgleichung der Form*

$$f(x'(t), x(t), t) = 0, \quad t \in \Im_0. \tag{2.29}$$

Die Gleichung (2.25) heißt kontraktiv auf der Mannigfaltigkeit $\Gamma(t) \subseteq M_0(t) \subset \mathbb{R}^m$*, wenn ein Skalarprodukt* $\langle \cdot, \cdot \rangle_S$ *und eine Konstante* $c \geq 0$ *existieren, so dass für alle Vektoren* $(y_1, x_1, t), (y_2, x_2, t) \in \mathbb{R}^m \times \Gamma(t) \times \Im_0$ *mit*

$$\begin{aligned} Qy_1 = Qy_2 = 0, \\ f(y_1, x_1, t) = f(y_2, x_2, t) = 0, \end{aligned}$$

wobei Q *der Orthoprojektor auf* N *in dem euklidischen Raum* $\{\mathbb{R}^m, \langle \cdot, \cdot \rangle_S\}$ *ist, die Ungleichung*

$$\langle y_1 - y_2, x_1 - x_2 \rangle_S \leq -c \left\| x_1 - x_2 \right\|_S^2, \tag{2.30}$$

gilt.

Bemerkung 2.2.11

1. *Bei allen drei Kontraktivitätsbegriffen unterscheidet man auch zwischen den Fällen* $c > 0$ *und* $c = 0$*. Man spricht zum einen von starker, zum anderen von schwacher Kontraktivität.*

2. *Es ist im Moment nicht eindeutig geklärt, ob die Forderung nach einer Kontraktivität für den ganzen Vektor* x *zu restriktiv ist. Zumindest bekommt man diesen Eindruck, wenn die Beispiele 3.0.8, 3.2.1 betrachtet werden. Andererseits kann man mit der Definition 2.2.10, wie es in den kommenden Seiten festgestellt wird, mehr erreichen.*

Das folgende Beispiel veranschaulicht, welche Bedeutung die Betrachtung des ganzen Vektors x in (2.30) hat.

Beispiel 2.2.12 *Man betrachtet die Index-1-Gleichung*

$$\begin{aligned} x_1'(t) &= -x_1(t), \\ x_2(t) &= e^{\alpha t} x_1(t), \end{aligned}$$

wobei die Lösung durch

$$\begin{aligned} x_1(t) &= x_1(t_0) e^{-t} \\ x_2(t) &= x_1(t_0) e^{(\alpha - 1)t} \end{aligned}$$

gegeben ist.

Angesichts der Kontraktivitäts-Definition aus [Griepentrog and März, 1986] müssen eine symmetrische positiv definite Matrix S und eine Konstante $c > 0$ existieren, so dass die Gleichungen

$$\begin{aligned} y_1 &= -x_1, \\ x_2 &= e^{\alpha t} x_1, \\ y_2 &= 0, \end{aligned}$$

die Ungleichung

$$\left\langle \begin{pmatrix} y_1 \\ 0 \end{pmatrix}, \begin{pmatrix} x_1 \\ 0 \end{pmatrix} \right\rangle_S \leq -c \left\| \begin{pmatrix} x_1 \\ 0 \end{pmatrix} \right\|_S^2,$$

implizieren. Nimmt man zum Beispiel die euklidische Norm, dann folgt

$$\begin{aligned} y_1 x_1 &\leq -c x_1^2, \\ -x_1^2 &\leq -c x_1^2, \end{aligned}$$

was für $0 \leq c \leq 1$ gilt. Demzufolge ist diese ADGl kontraktiv, ungeachtet des asymptotischen Verhaltens der zweiten Komponente der Lösung. Diese Komponente ist nämlich exponentiell wachsend für $\alpha > 1$.

Wenn man jetzt die Definition 2.2.10 für die Lösungsmannigfaltigkeit betrachtet, dann muss man die Erfüllung der Ungleichung

$$\left\langle \begin{pmatrix} y_1 \\ 0 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\rangle_S \leq -c \left\| \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\|_S^2$$

verlangen. Das bedeutet in der euklidischen Norm

$$\begin{aligned} y_1 x_1 &\leq -c(x_1^2 + x_2^2), \\ -x_1^2 &\leq -c(1 + e^{2\alpha t}) x_1^2, \\ 1 &\geq c(1 + e^{2\alpha t}). \end{aligned}$$

Nun ist α von Bedeutung. Die letzte Ungleichung kann nur erfüllt werden, entweder wenn $\alpha < 0$ (also die gesamte Lösung abfallend ist) oder wenn $c = 0$ (schwache Kontraktivität).

Eine genauere Betrachtung der Index-1-Bedingung,

$$N \cap S(t) = \{0\},$$

liefert ebenso interessante Einblicke. Die relevanten Räume sind durch

$$N = \left\{ z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in I\!R^2 : z_1 = 0 \right\},$$

$$S(t) = \left\{ z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in I\!R^2 : z_1 = e^{-\alpha t} z_2 \right\},$$

gegeben. Offensichtlich besitzt die Gleichung den Tractability-Index-1, aber wenn $\alpha > 1$, also wenn x_2 wächst, nähern sich die Unterräume N und $S(t)$ mit wachsendem t und es wächst die Norm von $G_1^{-1}(t)$.

Das folgende Theorem erfasst die Beziehung zwischen den Kontraktivitätsbegriffen für Index-2-ADGln der Form (2.21).

Theorem 2.2.13 *Kontraktivität auf $\Gamma(t) \subseteq M_0(t)$ impliziert P-Kontraktivität auf $\Gamma(t)$.*

Beweis: E s genügt zu zeigen, dass unter den Voraussetzungen der Definition 2.2.10 die Ungleichung (2.30) (2.24) für den Orthoprojektor P mit sich bringt. Sei

$$\langle y_1 - y_2, x_1 - x_2 \rangle_S \leq -c \, \|x_1 - x_2\|_S^2$$

erfüllt. Wenn wir die Vektoren in die $\langle \cdot, \cdot \rangle_S$-orthogonalen Komponenten P und Q zerlegen, folgt

$$\begin{aligned} \langle P(y_1 - y_2), P(x_1 - x_2) \rangle_S &\leq -c(\|P(x_1 - x_2)\|_S^2 + \|Q(x_1 - x_2)\|_S^2) \\ \langle y_1 - y_2, P(x_1 - x_2) \rangle_S &\leq -c \, \|P(x_1 - x_2)\|_S^2 . \end{aligned}$$

Bemerkung 2.2.14

1. Dass die Umkehrung dieses Theorems nicht gilt, zeigt das Beispiel 2.2.12.

2. Das Theorem sagt aus, dass (2.30) eine stärkere Anforderung als (2.24) an die Gleichung stellt.

Wie aus der gewöhnlichen Differentialgleichungstheorie bekannt, ist eine der wichtigsten Eigenschaften einer kontraktiven Gleichung der exponentielle Abfall des Abstandes zwei beliebiger Lösungen. Eine ähnliche Aussage ist mit der Definition 2.2.10 auch für ADGln gültig.

Theorem 2.2.15 *Sei eine ADGl der Form (2.2) im starken Sinne kontraktiv auf einer invarianten Mannigfaltigkeit* $\Gamma(t) \subseteq M_0(t)$ *und* $x_1(\cdot)$, $x_2(\cdot)$ *zwei Lösungen, die in* $\Gamma(t)$ *enthalten sind. Dann gilt*

$$\begin{aligned} &\|P(x_1(t) - x_2(t))\|_S \leq \|P(x_1(t_0) - x_2(t_0))\|_S \, e^{-2c(t-t_0)}, \quad \forall t \geq t_0. \\ &\lim_{t\to\infty} \|Q(x_1(t) - x_2(t))\|_S = 0. \end{aligned}$$

Beweis: M an definiert die Skalarfunktion

$$\alpha(t) := \|P(x_1(t) - x_2(t))\|_S^2 = \langle P(x_1(t) - x_2(t)), P(x_1(t) - x_2(t))\rangle_S \,,$$

und als Nächstes differenziert man diese Funktion:

$$\alpha'(t) = 2\,\langle (Px_1)'(t) - (Px_2)'(t), P(x_1(t) - x_2(t))\rangle_S \,.$$

Man führt die Notation

$$\begin{aligned} y_1(t) &:= (Px_1)'(t) \\ y_2(t) &:= (Px_2)'(t) \end{aligned}$$

ein und erhält

$$\alpha'(t) = 2\,\langle (y_1(t) - y_2(t)), P(x_1(t) - x_2(t))\rangle_S \,.$$

Die Vektoren $y_1(t)$ und $y_2(t)$ erfüllen offensichtlich die Bedingung

$$Qy_1(t) = Qy_2(t) = 0.$$

Außerdem sind $x_1(\cdot)$, $x_2(\cdot)$ Lösungen der ADGl und gehören zu $\Gamma(t)$. Deswegen sind die Ungleichung (2.30) und demzufolge auch

$$\langle y_1(t) - y_2(t), P(x_1(t) - x_2(t))\rangle_S \leq -c\,\|P(x_1(t) - x_2(t))\|_S^2 \,,$$

gültig. Man erhält dann

$$\begin{aligned} \alpha'(t) &\leq -2c\,\|P(x_1(t) - x_2(t))\|_S^2 \,, \\ \alpha'(t) &\leq -2c\alpha(t). \end{aligned}$$

Nun braucht man nur Peanos Lemma auf diese Differentialungleichung anzuwenden und es folgt

$$\alpha(t) \leq \alpha(t_0)e^{-2c(t-t_0)}.$$

Schließlich benutzt man für die zweite Aussage die Ungleichung

$$\langle y_1(t) - y_2(t), P(x_1(t) - x_2(t))\rangle_S \leq -c \left\|Q(x_1(t) - x_2(t))\right\|_S^2,$$

woraus folgt

$$\begin{aligned} \alpha'(t) &\leq -2c \left\|Q(x_1(t) - x_2(t))\right\|_S^2 \leq 0, \\ \alpha(t) - \alpha(t_0) &\leq -2c \int_{t_0}^{t} \left\|Q(x_1(\tau) - x_2(\tau))\right\|_S^2 d\tau \leq 0. \end{aligned}$$

Die letzte Ungleichung impliziert, dass das Integral

$$\int_{t_0}^{\infty} \left\|Q(x_1(\tau) - x_2(\tau))\right\|_S^2 d\tau$$

konvergiert, und damit ist zwangläufig die zweite Aussage gültig.

Bemerkung 2.2.16

1. Dieses Theorem, dass die Lösungen in der P-Komponente exponentiell asymptotisch stabil und in der Q-Komponente asymptotisch stabil sind.

2. Wenn statt Kontraktivität P-Kontraktivität angenommen wird, dann kann man nur die exponentielle asymptotische Stabilität der P-Komponente beweisen. In diesem Falle kann außerdem schwache P-Kontraktivität vorausgesetzt werden. Der Fakt $c > 0$ wird im Beweis nur für die Aussage der Q-Komponente gebraucht.

3. Wie schon gezeigt wurde, ist die Definition 2.2.10 restriktiver als die Definiton 2.2.7. Aber mit diesem Ergebnis wird deutlich, dass man in manchen Fällen aus der restriktiveren Bedingung (2.30) Vorteile ziehen kann.

4. Eine Aussage für die Q-Komponente der Lösung, wenn ausschließlich die Ungleichung (2.24) gilt, kann auch bei den linearen Aufgaben erzielt werden, wenn der Projektor auf den Lösungsraum beschränkt ist, [Griepentrog and März, 1986], [Hanke et al., 1998], [März and Rodríguez Santiesteban, 1999].

2.3 Stabilität im diskreten Fall

In diesem Abschnitt wird einen Stabilitätsbegriff für Diskretisierungsverfahren im ADGln-Kontext eingeführt. Aus der gewöhnlichen Differentialgleichungstheorie kennt man mehrere numerische Stabilitätsbegriffe, [Hairer et al., 1987], [Dekker and Verwer, 1984]. Die A-Stabilität und ihre Verallgemeinerung auf nicht-lineare Systeme, die B-Stabilität, sind womöglich dabei die allerwichtigsten.

Das A-Stabilitätskonzept für ADGln ist ebenfalls in der Fachliteratur zu finden, [Griepentrog and März, 1986], [Wensch et al., 1995]. Die Schwierigkeiten bestehen im Wesentlichen darin, eine sinnvolle Testproblemklasse zu definieren, die in der ADGln-Theorie die Rolle der Gleichung

$$x'(t) = Bx(t)$$

übernehmen soll. Beispielsweise erweist sich die Betrachtung der Testgleichung

$$Ax'(t) + Bx(t) = 0 \tag{2.31}$$

als nicht ausreichend. Wie aus dem Beispiel 1.0.1 und [Hanke et al., 1998] hervorgeht, ist die Stabilität eines impliziten Euler-Verfahrens für (2.31) für Aufgaben mit variablen Koeffizienten irrelevant. Wie in Kapitel 3 ersichtlich wird, sind letztere qualitativ komplexer als die Gleichungen mit konstanten Koeffizienten.

Dennoch stellen die (P-)kontraktiven nicht-linearen ADGln der Art (2.2) eine Klasse von Aufgaben dar, die als Testklasse sinnvoll ist und den Weg für B-stabilitätsähnliche Konzepte öffnet.

Analog zum Abschnitt 2.2.2 sollen in dem diskreten Fall zwei Kontraktivitätskonzepte betrachtet werden.

Definition 2.3.1 *Die Anwendung eines Einschrittverfahrens $x_{j+1} = \Phi(x_j, t_j, h_j)$ auf eine Differentialgleichung (ADGl oder gewöhnliche Differentialgleichung) heißt P-kontraktiv, wenn eine Norm $\|\cdot\|_S$ existiert, so dass die Ungleichung*

$$\left\|Px_{j+1}^{(1)} - Px_{j+1}^{(2)}\right\|_S \leq \left\|Px_j^{(1)} - Px_j^{(2)}\right\|_S$$

für konsistente Anfangswerte $x_0^{(1)}$, $x_0^{(2)}$ und $\forall j > 0$, erfüllt ist. Wenn außerdem die Ungleichung

$$\left\|Qx_{j+1}^{(1)} - Qx_{j+1}^{(2)}\right\|_S \leq K \left\|Px_j^{(1)} - Px_j^{(2)}\right\|_S,$$

wobei K eine positive Konstante ist, gilt, dann nennt man die Verfahrensanwendung kontraktiv.

Bemerkung 2.3.2 *Die Verwandtschaft mit dem B-Stabilitätsbegriff von [Griepentrog and März, 1986], [Hanke et al., 1998] und [März and Rodríguez Santiesteban, 1999] ist nicht zu übersehen. Der Unterschied besteht darin, dass, statt die Kontraktion des Verfahrens für alle kontraktiven Gleichungen zu fordern, Kontraktion bezogen auf die Anwendung des Verfahrens bei einer bestimmten Aufgabe verlangt wird.*

Als Abschluss für dieses Kapitel soll eine Aussage bewiesen werden, die in den kommenden Kapiteln eine zentrale Rolle spielen wird. Die Aussage bezieht sich auf die oben genannten impliziten Runge-Kutta-DAE-Verfahren.

Ein implizites Runge-Kutta-Verfahren kann folgendermaßen für die ADGl (2.11) realisiert werden, [Petzold, 1986]: Seien auf dem Intervall $\Im$ ein Gitter $t_l = t_{l-1} + h$, $l \geq 1$, und eine Approximation x_{l-1} für $x(t_{l-1})$ vorgegeben. Eine Approximation x_l für die Lösung in dem Punkt t_l folgt dann nach dem s-stufigen Runge-Kutta-Schema

$$\begin{array}{c|c} c & A_\alpha \\ \hline & \beta \end{array}$$

mit A_α als die Matrix $(\alpha_{ij})_{i,j=1}^s$, $\beta = (\beta_1, \ldots, \beta_s)$ und $c = (c_1, \ldots, c_s)^T$, durch

$$x_l = x_{l-1} + h \sum_{i=1}^{s} \beta_i X'_{li}, \tag{2.32}$$

wobei die Vektoren X'_{li}, X_{li} das Gleichungssystem

$$\begin{aligned} A(X_{li}, t_{li}) X'_{li} + b(X_{li}, t_{li}) &= 0,\ i = 1, \ldots, s, && (2.33) \\ X_{li} &= x_{l-1} + h \sum_{j=1}^{s} \alpha_{ij} X'_{lj},\ i = 1, \ldots, s, && (2.34) \\ t_{li} &= t_{l-1} + c_i h, \end{aligned}$$

erfüllen.

Eine hinreichende Bedingung, damit die Gleichungen (2.33), (2.34) die Unbekannten X'_{li}, X_{li} eindeutig bestimmen, ist die Regularität der Matrix A_α, [Macana, 1993]. Es wird im Weiteren die Erfüllung dieser Bedingung angenommen werden.

Sei $\widehat{A}_\alpha := A_\alpha^{-1} = (\hat{\alpha}_{ij})_{i,j=1}^s$ und $\rho := 1 - \sum_{i=1}^s \sum_{j=1}^s \beta_i \hat{\alpha}_{ij}$. Das IRK-Verfahren kann dann folgendermaßen umgeschrieben werden:

$$x_l = \rho x_{l-1} + \sum_{i=1}^{s} \sum_{j=1}^{s} \beta_i \hat{\alpha}_{ij} X_{lj}, \qquad (2.35)$$

$$A_{li} \sum_{j=1}^{s} \hat{\alpha}_{ij}(X_{lj} - x_{l-1}) + h b_{li} = 0,\ i = 1, \ldots, s. \qquad (2.36)$$

Hier wurden die Bezeichnungen $A_{li} := A(X_{li}.t_{li})$ und $b_{li} := b(X_{li}, t_{li})$ benutzt.

Bemerkenswert ist, dass auf Grund von (2.36) alle Stufen X_{li} auf der Mannigfaltigkeit $M_0(t_{li})$ liegen. Wenn das Verfahren ein so genanntes IRK(DAE) ist, also wenn

$$\beta_i = \alpha_{si}, \quad i = 1, \ldots, s, \quad c_s = 1,$$

dann ist $\rho = 0$, $t_{ls} = t_l$ und folglich $x_l = X_{ls}$. So gehört die Approximation für die Lösung auch zu $M_0(t_l)$. Aus diesem Grund sind die IRK(DAE)-Verfahren (in der gewöhnlichen Differentialgleichungstheorie als "stiffly accurate" bekannt, [Hairer et al., 1987]) für ADGl besonders geeignet.

Bevor die erwähnte Aussage formuliert wird, benötigt man noch das Konzept der algebraischen Stabilität.

Definition 2.3.3 *([Hairer et al., 1987]) Ein Runge-Kutta-Verfahren heißt algebraisch stabil, wenn die folgenden Bedingungen erfüllt sind:*

1. $\beta_i \geq 0$, für $i = 1, \ldots, s$
2. Die Matrix $(m_{ij}) = (\beta_i \alpha_{ij} + \beta_j \alpha_{ji} - \beta_i \beta_j)$ ist positiv semi definit.

Bemerkung 2.3.4 *Sehr wichtige Vertreter dieser Verfahren sind die Gauß-, Radau-IA-, Radau-IIA- und Lobatto-IIIC-Runge-Kutta-Verfahren, [Hairer et al., 1987]. Allerdings sind davon nur die Radau-IIA-Verfahren auch IRK(DAE).*

Theorem 2.3.5 *Sei $x_{l+1} = \Phi(x_l, t_l, h_l)$ ein algebraisch stabiles IRK(DAE), mit $\beta_s > 0$. Seine Anwendung auf die AWA*

$$\begin{aligned} A(x(t), t)x'(t) + b(x(t), t) &= 0,\ t \in \Im_0 \qquad (2.37) \\ x(t_0) &= x_{0,} \end{aligned}$$

ist (P-)kontraktiv (wobei es sich bei x_0 um einen konsistenten Anfangswert handelt), wenn

1. *die Gleichung (2.37) (P-)kontraktiv auf einer Mannigfaltigkeit $\Gamma(t) \subseteq M_0(t)$ ist und*

2. *die Stufen X_{li} des Verfahrens immer zu $\Gamma(t_{li})$ gehören.*

Beweis: D er Beweis verläuft nach dem klassischen Schema. Es wird zuerst die Kontraktivität für die P-Komponente der Lösung gezeigt. Man betrachtet die Differenzen

$$\begin{aligned} \Delta x_l &:= x_l^{(1)} - x_l^{(2)}, \\ \Delta X_{li} &:= X_{li}^{(1)} - X_{li}^{(2)}, \\ \Delta X'_{li} &:= X_{li}^{(1)'} - X_{li}^{(2)'}. \end{aligned}$$

Aus (2.32) und (2.34) folgt für diese Differenzen

$$\begin{aligned} P\Delta x_l &= P\Delta x_{l-1} + h\sum_{i=1}^{s} \beta_i P\Delta X'_{li}, \\ P\Delta X_{li} &= P\Delta x_{l-1} + h\sum_{j=1}^{s} \alpha_{ij} P\Delta X'_{lj}. \end{aligned} \tag{2.38}$$

Als Nächstes wird das Quadrat der Norm von $P\Delta x_{l+1}$ genommen,

$$\|P\Delta x_l\|_S^2 = \|P\Delta x_{l-1}\|_S^2 + 2h\sum_{i=1}^{s} \beta_i \langle P\Delta X'_{li}, P\Delta x_{l-1}\rangle_S + h^2 \sum_{i=1}^{s}\sum_{j=1}^{s} \beta_i\beta_j \left\langle P\Delta X'_{li}, P\Delta X'_{lj}\right\rangle_S,$$

und man setzt den Ausdruck für $P\Delta x_l$, der sich aus (2.38) ergibt, in die letzte Gleichung ein:

$$\|P\Delta x_l\|_S^2 = \|P\Delta x_{l-1}\|_S^2 + 2h\sum_{i=1}^{s} \beta_i \langle P\Delta X'_{li}, P\Delta X_{li}\rangle_S - h^2 \sum_{i=1}^{s}\sum_{j=1}^{s} m_{ij} \left\langle P\Delta X'_{li}, P\Delta X'_{lj}\right\rangle_S.$$

Da die Stufen X_{li} zu der Mannigfaltigkeit $\Gamma(t_{li})$ gehören und Kontraktivität von (2.37) auf dieser Mannigfaltigkeit vorliegt, gilt, dass

$$\langle P\Delta X'_{li}, P\Delta X_{li}\rangle_S \leq 0, \ \forall i.$$

Durch die algebraische Stabilität des Verfahrens ist der zweite und dritte Term in dem Ausdruck von $\|P\Delta x_l\|_S^2$ kleiner oder gleich Null, und somit folgt die Ungleichung

$$\|P\Delta x_l\|_S^2 \leq \|P\Delta x_{l-1}\|_S^2 .$$

Die Ungleichung

$$\begin{aligned} \|P\Delta x_l\|_S^2 &\leq \|P\Delta x_{l-1}\|_S^2 + 2h\sum_{i=1}^{s}\beta_i \langle P\Delta X'_{li}, P\Delta X_{li}\rangle_S \\ &\leq \|P\Delta x_{l-1}\|_S^2 + 2h\beta_s \langle P\Delta X'_{ls}, P\Delta X_{ls}\rangle_S \end{aligned}$$

ist ebenfalls gültig und stellt den Ausgangspunkt für den Beweis der Q-Komponente dar. Die Kontraktivität der ADGl bringt auch

$$\langle P\Delta X'_{ls}, P\Delta X_{ls}\rangle_S \leq -c\,\|Q\Delta X_{ls}\|_S^2$$

mit sich und die zwei letzten Ungleichungen zusammen ergeben

$$\begin{aligned} \|P\Delta x_l\|_S^2 &\leq \|P\Delta x_{l-1}\|_S^2 - 2hc\beta_s\,\|Q\Delta X_{ls}\|_S^2 , \\ 2hc\beta_s\,\|Q\Delta X_{ls}\|_S^2 &\leq \|P\Delta x_{l-1}\|_S^2 , \\ \|Q\Delta x_l\|_S = \|Q\Delta X_{ls}\|_S &\leq \frac{1}{\sqrt{2hc\beta_s}}\,\|P\Delta x_l\|_S . \end{aligned}$$

Bemerkung 2.3.6

1. Dieses Theorem hätte man noch allgemeiner, nämlich für eine Differentialgleichung der Art (2.25), formulieren können. Die Aussagen für gewöhnliche Differentialgleichungen sind Sonderfälle dieses allgemeinen Theorems.

2. Wie aus dem Beweis hervorgeht, genügt für die P-Kontraktivität des IRK(DAE) die schwache (P-)Kontraktivität der ADGl auf $\Gamma(t)$.

3. Dieses Ergebnis ist natürlich nur für theoretische Untersuchungen von Bedeutung. Die Bedingungen sind im Allgemeinen sehr schwer zu überprüfen. Allerdings kann Theorem 2.3.5 als Werkzeug für die Untersuchung von bestimmten Klassen von Aufgaben verwendet werden. In den Kapiteln 3 und 4 wird jeweils eine Anwendung dieses Theorems vorgestellt.

Korollar 2.3.7 *Die algebraisch stabilen IRK(DAE) sind (P-)kontraktiv für die Klasse der Index-2-ADGln, die auf der Mannigfaltigkeit $M_0(t)$ (schwach P-)kontraktiv sind.*

Beweis: U m das Theorem 2.3.5 anzuwenden, ist es nur noch notwendig, dass alle Stufen des Verfahrens zu $M_0(t_{li})$ gehören. Dieser Fakt folgt unmittelbar aus der Gleichung (2.36).

Kapitel 3

Stabilitätserhaltungsfälle

Das Theorem 2.3.5 des vorhergehenden Kapitels stellt allgemeine Bedingungen, damit bestimmte IRK-Verfahren (P-)kontraktiv sind. In diesem Kapitel werden verschiedene Fälle untersucht, in denen allgemeine Runge-Kutta- und BDF-Verfahren eine bestimmte Kommutativitätsregel erfüllen. Die Grundidee der Analyse kann wie im Diagramm 3.1 dargestellt werden.

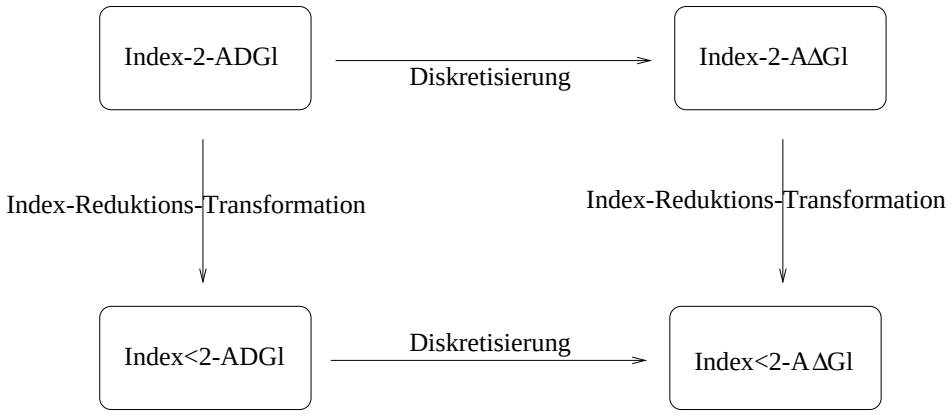

Abbildung 3.1: Kommutativität zwischen einer Index-Reduktions-Transformation und einer Diskretisierung

Der Term AΔGl steht für eine diskretisierte ADGl. Das Diagramm repräsentiert die Kommutativität der Operationen Diskretisierung und IRT (Index-Reduktions-Transformation). Wenn es gilt, kann man davon ausgehen, dass sich die Diskretisierung der Index-2-ADGl genau so wie bei der

Index¡2-ADGl verhält. Als IRT werden zwei Möglichkeiten betrachtet: zum einen die Index-Reduktion durch Differentiation und zum anderen die Entkopplungstechniken. Im ersten Fall ergibt sich eine Index-1-Gleichung, die unter den Voraussetzungen dieser Arbeit eine führende Matrix mit konstantem Nullraum N besitzt. Damit verhält sich die Anwendung eines Diskretisierungsverfahrens wie für eine gewöhnliche Differentialgleichung, [Griepentrog and März, 1986]. Im zweiten Fall folgt aus der IRT, unter anderem, eine gewöhnliche Differentialgleichung. Als mögliche Diskretisierung werden IRK-Verfahren, insbesondere IRK(DAE), und BDF-Verfahren in der Analyse betrachtet. Als Nebenprodukt dieser Untersuchung fällt eine konkrete Anwendung von Theorem 2.3.5 auf Modelle der Schaltkreissimulation ab.

Die oben genannte Kommutativitätsregel als Untersuchungmethode wurde bereits in [Hanke et al., 1998] angewandt. Aus dieser Arbeit geht hervor, dass eine hinreichende Bedingung, damit die numerischen Verfahren für lineare ADGl aus der Stabilitätssicht sich wie für GDGl verhalten, $Q_{N_1}^{S_1,}(t) \equiv 0$ ist. Die Bedingung von [Hanke et al., 1998] ist aus der Sicht der Anwendungen in gewissem Maße restriktiv. In vielen Modellen aus der Schaltkreissimulation und der mechanischen Systeme ist diese geforderte Bedingung beispielsweise nicht erfüllt.

Auf der Basis numerischer Experimente ist festgestellt worden, dass die bekannten numerischen Verfahren Stabilitätsprobleme bei manchen Aufgaben aufweisen, [Hanke et al., 1998], [Wensch et al., 1995]. Im ersten der zitierten Artikel ist folgendes Beispiel zu finden:

Beispiel 3.0.8

$$\begin{pmatrix} x_1' \\ x_2' \\ 0 \end{pmatrix} + \begin{pmatrix} \lambda & -1 & -1 \\ \eta t(1-\eta t) - \eta & \lambda & -\eta t \\ 1-\eta t & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0,\ t \geq 0.$$

Es handelt sich um ein Index-2-Hessenberg-System wie das Beispiel 2.1.9, wobei die Regularität der Matrix

$$B_{21}B_{12} = \begin{pmatrix} 1-\eta t & 1 \end{pmatrix} \begin{pmatrix} -1 \\ -\eta t \end{pmatrix} = -1$$

für alle t garantiert ist. Die allgemeine Lösung ist durch

$$\begin{aligned} x_1(t) &= x_1(0)e^{-\lambda t}, \\ x_2(t) &= (\eta t - 1)x_1(t), \\ x_3(t) &= -(\eta t - 1)x_1(t), \end{aligned}$$

gegeben.

Die Lösungsmannigfaltigkeit für diese Gleichung ist

$$M_1(t) = \left\{ z \in I\!R^3 : z = \begin{pmatrix} z_1 \\ (\eta t - 1) z_1 \\ (1 - \eta t) z_1 \end{pmatrix} \right\},$$

und der Raum $S_1(t)$ ist hier

$$S_1(t) = \left\{ z \in I\!R^3 : (1 - \eta t) z_1 + z_2 = 0 \right\}.$$

Offensichtlich ist die Lösung exponentiell asymptotisch stabil im Lyapunovschen Sinne, wenn $\lambda > 0$ ist. Nunmehr liegt bei diesem Problem Kontraktivität auf der Lösungsmannigfaltigkeit vor. Man betrachtet das euklidische Skalarprodukt und einen Vektor (y, x, t), der zu $M_1(t)$ gehört und die Gleichungen

$$\begin{aligned} y_1 &= -\lambda x_1 + x_2 + x_3, \\ y_2 &= \left[\eta - \eta t(1 - \eta t)\right] x_1 - \lambda x_2 + \eta t x_3, \\ 0 &= (1 - \eta t) x_1 + x_2, \\ y_3 &= 0 \end{aligned}$$

erfüllt. Die linke Seite der Kontraktivitätsbedingung ist dann

$$\langle y, x \rangle_2 = -\lambda x_1^2 + x_1 x_2 + x_1 x_3 + \left[\eta - \eta t(1 - \eta t)\right] x_1 x_2 - \lambda x_2^2 + \eta t x_2 x_3.$$

Man nutzt jetzt die Bedingung $x \in M_1(t)$ und setzt

$$\begin{aligned} x_2 &= (\eta t - 1) x_1, \\ x_3 &= -(\eta t - 1) x_1, \end{aligned}$$

in den Ausdruck für das Skalarprodukt ein. Es folgt

$$\langle y, x \rangle_2 = \left[-\lambda + \eta(\eta t - 1) - \lambda(\eta t - 1)^2\right] x_1^2.$$

Aus diesem Ausdruck kann man erkennen, dass $\langle y, x \rangle_2$, für $\lambda > 0$ und t hinreichend groß, negativ ist (es liegt schwache Kontraktivität vor). Diese Tatsache reicht, um die Kontraktivität eines algebraisch stabilen IRK(DAE)-

Verfahrens zumindest in der P-Komponente für t hinreichend groß zu garantieren, wenn die Verfahrensstufen zu M_1 gehören, siehe Beweis von Theorem 2.3.5.

Weiterhin ist die Aufgabe stark kontraktiv auf $M_1(t)$, wenn $\eta = -\lambda$. In diesem Fall folgt für das Skalarprodukt

$$\begin{aligned}\langle y, x\rangle_2 &= -\lambda x_1^2 - \frac{\lambda}{2}x_2^2 - \frac{\lambda}{2}x_3^2 + \eta(\eta t - 1)x_1^2 \\ &= -\lambda x_1^2 - \frac{\lambda}{2}x_2^2 - \frac{\lambda}{2}x_3^2 + \eta x_1 x_2 \\ &= -\lambda(x_1^2 + \frac{1}{2}x_2^2 + \frac{1}{2}x_3^2 + x_1 x_2) \\ &= -\lambda \langle x, x\rangle_S = -\lambda \|x\|_S^2 ,\end{aligned}$$

wobei die symmetrische positiv definite Matrix S als

$$S = \begin{pmatrix} 1 & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}$$

definiert ist. Jetzt folgt wegen der Äquivalenz der Normen in $\mathbb{R}^3$ die Kontraktivitätsungleichung

$$\langle y, x\rangle_2 \leq -\lambda\gamma^2 \|x\|_2^2 .$$

Andererseits stellt die Anwendung des impliziten Euler-Verfahrens auf dieses Problem die Lösung folgenden linearen Systems in jedem Schritt dar:

$$\begin{pmatrix} 1+h\lambda & -h & -h \\ h(\eta t_{i+1}(1-\eta t_{i+1}) - \eta) & 1+h\lambda & -h\eta t_{i+1} \\ 1-\eta t_{i+1} & 1 & 0 \end{pmatrix} \begin{pmatrix} x_{1,i+1} \\ x_{2,i+1} \\ x_{3,i+1} \end{pmatrix} = \begin{pmatrix} x_{1,i} \\ x_{2,i} \\ 0 \end{pmatrix}.$$

Aus diesem System folgt für die erste Komponente der Lösung die Rekursion

$$x_{1,i+1} = \frac{1+h\eta}{1+h(\eta+\lambda)} x_{1,i},$$

und für die Kontraktivität des Verfahrens in dieser Komponente erhält man die Bedingung

$$\left| \frac{1+h\eta}{1+h(\eta+\lambda)} \right| < 1.$$

Diese Bedingung kann für $\eta < 0$ verletzt werden, insbesondere für $\eta = -\lambda$, obwohl in diesem Fall Kontraktivität auf $M_1(t)$ vorliegt. Aus dieser Tatsache kann man schließen, dass Kontraktivität auf der Mannigfaltigkeit $M_0(t)$ nicht vorhanden ist, Theorem 2.3.5.

Die numerischen Ergebnisse dieses Verfahrens für die Komponente x_1, $h = 0.1$ und verschiedene Werte von λ und η sind in der Abbildung 3.2 zu sehen.

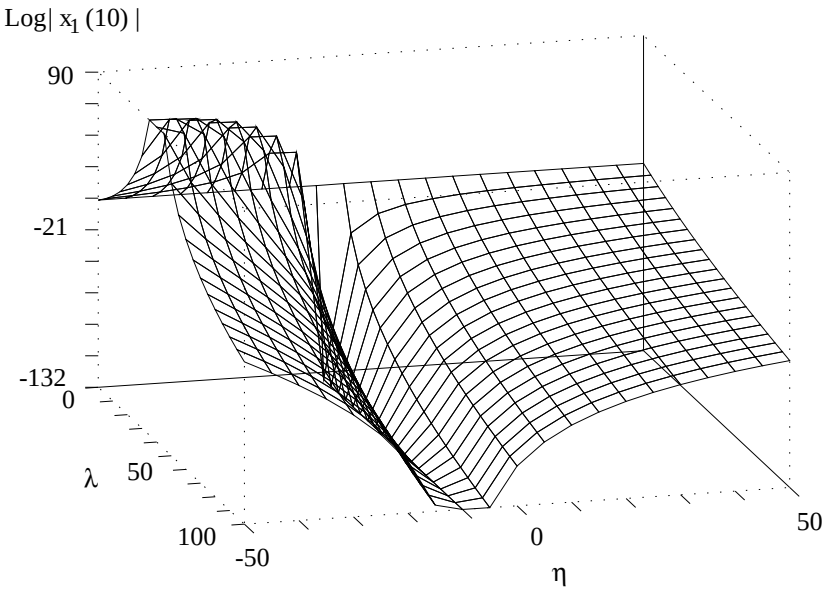

Abbildung 3.2: Approximierte Lösung durch das implizite Euler-Verfahren für das Beispiel 3.0.8. Die Grafik stellt den Betrag von $x_1(10)$ für den Diskretisierungsschritt $h = 0.1$ und verschiedene Werte von λ und η dar.

Weiterhin ist der Fakt anzumerken, dass die in [Hanke et al., 1998] gefundene Bedingung ($Q_{N_1}^{S_1,} \equiv 0$) im vorherigen Beispiel nicht erfüllt ist.

3.1 Entkopplung einer ADGl

In dieser Sektion wird ein sehr wichtiges Werkzeug für die Untersuchung Index-2-ADGln entwickelt. Die hier eingeführten Techniken sind etwas spe-

zifischer und gezielter als jene, die man normalerweise in der Fachliteratur findet, [März, 1997], [Hanke et al., 1998], [Tischendorf, 1996]. Der Vorteil dieser verfeinerten Vorgehensweise liegt darin, dass mit ihrer Hilfe Ergebnisse bezüglich der Lösungsexistenz und insbesondere der Kommutativität zwischen Entkopplung und Diskretisierung verbessert werden können.

Es wird in diesem Abschnitt eine quasilineare Index-2-ADGl der Form

$$A(t)(Px)'(t) + b(x(t), t) = 0 \quad t \in \Im, \tag{3.1}$$

betrachtet, die die Glattheitsvoraussetzungen aus Kapitel 2 erfüllt. Aus diesem Kapitel wird ebenfalls die Definition der relevanten Unterräume und Projektoren längs einer Lösung übernommen.

Für die neuen Index-2-Entkopplungsmethoden braucht man zunächst die folgende Definition:

Definition 3.1.1 *Ein Unterraum $K(t)$ von $\mathbb{R}^m$ wird ein entkoppelndes Komplement zu $N_1(t)$ genannt, wenn die Bedingungen*

1. $K(t) \supset N$,
2. $K(t) \cap N_1(t) = \{0\}$,
3. $K(t) \oplus N_1(t) = \mathbb{R}^m$,

für alle $t \in \Im$ erfüllt sind.

Bemerkung 3.1.2 *Ein entkoppelndes Komplement zu $N_1(t)$ existiert im Index-2-Fall immer, weil $N \cap N_1(t) = \{0\}$ ($N \subset S_1(t)$) ist. Nunmehr ist $S_1(t)$ selbst ein Beispiel dafür.*

Im Weiteren wird ausschließlich mit Projektoren $Q^K_{N_1}(t) := Pr(N_1(t), K(t))$ längs eines entkoppelnden Komplementraumes gearbeitet. Der in der Fachliteratur als kanonisch bezeichnete Projektor ist ein Beispiel dafür ($Q^{S_1}_{N_1}(t)$). Er ist der $\mathbb{R}^m$-Zerlegung

$$\mathbb{R}^m = N_1(t) \oplus S_1(t),$$

zugeordnet und projiziert auf $N_1(t)$ längs $S_1(t)$.

Wenn man erneut das Lemma A.14 von [Griepentrog and März, 1986] verwendet, folgt die Darstellung

$$Q^{S_1}_{N_1}(t) = Q^*_{N_1}(t) G^{-1}_{2,*}(t) B_1(t),$$

für den kanonischen Projektor.

Mit dem kanonischen Projektor läßt sich eine relativ einfache Entkopplung erreichen ([März, 1992]), jedoch ermöglichen die Entkopplungen mittels anderer Projektoren, sowohl Lösbarkeitsaussagen unter schwächeren Glattheitsvoraussetzungen zu treffen, [März and Rodríguez Santiesteban, 1999], als auch neue Einblicke in das Verhalten numerischer Verfahren für Index-2-ADGln zu gewinnen.

3.1.1 Linerare Index-2-Entkopplungen

In diesem Abschnitt wird der lineare Fall von (3.1) betrachtet:

$$A(t)(Px)'(t) + B(t)x(t) = q(t). \quad t \in \Im. \tag{3.2}$$

Außerdem wird vorausgesetzt:

Bedingung 3.1.3 *Es gibt eine stetig differenzierbare Basis von* $N_1(t)$.

Was bedeutet diese Bedingung zum Beispiel für ein Hessenberg-System? In diesem Fall, siehe Beispiel 2.1.9, ist $N_1(t)$ durch

$$N_1(t) = \left\{ \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in I\!R^m : z_1 = -B_{12}(t)z_2 \right\}$$

gegeben. Die Bedingung 3.1.3 fordert dann die Differenzierbarkeit von $B_{12}(t)$.

In der Fachliteratur wird zumeist die Entkopplung mit Hilfe des kanonischen Projektors $Q_{N_1}^{S_1}(t)$ durchgeführt, [März, 1992], [März, 1989], [Tischendorf, 1996], [Hanke et al., 1998]. Wie in diesen Arbeiten benötigt man die Differenzierbarkeit des verwendeten Projektors beziehungsweise Projektionen dessen. Solch eine Voraussetzung bedeutet, dass außerdem $S_1(t)$ glatt sein muss. Wenn man erneut den Hessenberg-Fall betrachtet, heißt es, dass $B_{21}(t)$ ebenfalls differenzierbar sein muss.

Die Entkopplungsvariante mit dem kanonischen Projektor hat den Vorteil, dass das entkoppelte System etwas einfacher ist. Ein Nachteil ist allerdings, dass man Freiheit "verschenkt". Um diese Idee etwas zu erläutern, soll die im Index-2-Fall vorhandene $I\!R^m$-Zerlegung

$$I\!R^m = N_1(t) \oplus S_1(t) \tag{3.3}$$

betrachtet werden. Der kanonische Projektor $Q_{N_1}^{S_1}(t)$ entspricht dieser Zerlegung und deswegen spiegelt er die Veränderung dieser beiden Räume wider.

Aber wie in diesem Kapitel gezeigt wird, kann man das System mit Projektoren entkoppeln, die nur an einen der Räume $N_1(t)$ oder $S_1(t)$ gebunden sind. Diese Trennung der entkoppelnden Projektoren von jeweils einem der Räume macht die Prozedur technisch aufwändiger, aber man wird mit stärkeren Aussagen belohnt. Konkret wird man in der Lage sein, die numerischen asymptotischen Stabilitätsergebnisse von [Hanke et al., 1998] zu verbessern.

Der Ansatz wird folgendermaßen realisiert: Unter der Bedingung 3.1.3 kann man einen differenzierbaren Projektor $Q_{N_1}^K(t)$ auf $N_1(t)$ längs eines entkoppelnden Komplementsraumes $K(t)$ finden. Sei $P_K^{N_1}(t) := I - Q_{N_1}^K(t)$.

Für die Entkopplung werden die folgenden Ausdrücke eingeführt:

$$\begin{aligned} G_{2,K}^{-1}A &= G_{2,K}^{-1}G_1P = G_{2,K}^{-1}G_{2,K}P_K^{N_1}P = P_K^{N_1}P, \\ G_{2,K}^{-1}Bx &= G_{2,K}^{-1}BPP_K^{N_1}x + G_{2,K}^{-1}BPQ_{N_1}^Kx + G_{2,K}^{-1}BQx, \\ &= G_{2,K}^{-1}BPP_K^{N_1}x + G_{2,K}^{-1}G_{2,K}Q_{N_1}^Kx + G_{2,K}^{-1}G_1Qx, \\ &= G_{2,K}^{-1}BPP_K^{N_1}x + Q_{N_1}^Kx + Qx. \end{aligned}$$

Durch eine Skalierung von (3.2) mit $G_{2,K}^{-1}(t)$ folgt dann

$$P_K^{N_1}P(Px)' + G_{2,K}^{-1}BPP_K^{N_1}x + Q_{N_1}^Kx + Qx = G_{2,K}^{-1}q. \tag{3.4}$$

An dieser Stelle definiert man wie in [Tischendorf, 1996] die Hilfsprojektoren $T(t)$ und $U(t) := I - T(t)$, wobei $T(t)$ auf $\operatorname{im}(Q(t)Q_{N_1}^K(t))$ projiziert. Man kann die Entkopplungen ohne diese Projektoren durchführen, jedoch müssen, wenn man diese Prozedur für nicht-lineare Probleme verwendet, restriktivere strukturelle Bedingungen vorausgesetzt werden, [März, 1995]. In [Tischendorf, 1996] wurde mit dem kanonischen Projektor gearbeitet. Es sollen hier unter den gleichen Voraussetzungen zwei Verallgemeinerungen davon vorgestellt werden.

Der Raum $\operatorname{im}(Q(t)Q_{N_1}^K(t))$ ist nichts anderes als $N \cap S(t)$, unabhängig von der Wahl von $Q_{N_1}^K(t)$, [Tischendorf, 1996]. Weiterhin sind TQ, UQ, $UQ + PQ_{N_1}^K$ und $TQP_K^{N_1}$ auch Projektoren, und man kann die Einheitsmatrix wie folgt zerlegen:

$$I = PP_K^{N_1} + PQ_{N_1}^K + UQ + TQ.$$

Durch die Multiplikation von (3.4) mit $PP_K^{N_1}$, $TQP_K^{N_1}$ und $PQ_{N_1}^K + UQ$ erhält

man

$$\begin{aligned} PP_K^{N_1}(Px)' + PP_K^{N_1}G_{2,K}^{-1}BPP_K^{N_1}x &= PP_K^{N_1}G_{2,K}^{-1}q, && (3.5)\\ -QQ_{N_1}^K(Px)' + TQP_K^{N_1}G_{2,K}^{-1}BPP_K^{N_1}x + TQx &= TQP_K^{N_1}G_{2,K}^{-1}q, && (3.6)\\ (UQ + PQ_{N_1}^K)G_{2,K}^{-1}BPP_K^{N_1}x + (UQ + PQ_{N_1}^K)x &= (UQ + PQ_{N_1}^K)G_{2,K}^{-1}q, && (3.7) \end{aligned}$$

was den Grundstein für die verschiedenen Entkopplungen darstellt. Das System (3.5)-(3.7) ist äquivalent zu (3.4), weil die Identität

$$I = PP_K^{N_1} + TQP_K^{N_1} + (UQ + Q_{N_1}^K)(UQ + PQ_{N_1}^K)$$

gilt.

Bis jetzt wurde den Ideen von [Tischendorf, 1996] bis auf die Verwendung eines allgemeinen Projektors $Q_{N_1}^K(t)$, der nur an $N_1(t)$ gebunden ist, gefolgt.

Entkopplung mit einem Projektor auf $N_1(t)$

Man berechnet zunächst

$$\begin{aligned} PP_K^{N_1}(Px)' &= PP_K^{N_1}(PP_K^{N_1}x + PQ_{N_1}^K x)' = PP_K^{N_1}(PP_K^{N_1}x)' + PP_K^{N_1}(PQ_{N_1}^K x)'\\ &= (PP_K^{N_1}x)' - (PP_K^{N_1})'PP_K^{N_1}x - (PP_K^{N_1})'PQ_{N_1}^K x, \end{aligned}$$

und definiert $u(t) := PP_K^{N_1}(t)x(t)$, $y(t) := (U(t)Q + PQ_{N_1}^K(t))x(t)$ und $w(t) := T(t)Qx(t)$. Da angenommen wurde, dass der Projektor $Q_{N_1}^K(t)$ stetig differenzierbar ist, kann (3.5)-(3.7) folgendermaßen transformiert werden:

$$\begin{aligned} u' - (PP_K^{N_1})'(u + Py) + PP_K^{N_1}G_{2,K}^{-1}Bu &= PP_K^{N_1}G_{2,K}^{-1}q, && (3.8)\\ -QQ_{N_1}^K(Py)' + (QQ_{N_1}^K)'u + TQP_K^{N_1}G_{2,K}^{-1}Bu + w &= TQP_K^{N_1}G_{2,K}^{-1}q, && (3.9)\\ y + (UQ + PQ_{N_1}^K)G_{2,K}^{-1}B_1u &= (UQ + PQ_{N_1}^K)G_{2,K}^{-1}q. && (3.10) \end{aligned}$$

Von (3.10) erhält man y, nach Einsetzen von y in (3.8) resultiert eine GDGL für u, die unter den Voraussetzungen lösbar ist. Zum Schluss, falls

$$PQ_{N_1}^K G_{2,K}^{-1}q \quad \textit{und} \quad PQ_{N_1}^{S_1}$$

stetig differenzierbar sind, bekommt man w aus (3.9). Vorausgesetzt, dass die so gerechneten Funktionen u, y und w jeweils zu den Räumen im $PP_K^{S_1}$, im $(UQ + PQ_{N_1}^K)$ und im TQ gehören, kann die Lösung $x(t)$ als

$$x(t) = u(t) + y(t) + w(t)$$

gebildet werden. Aus (3.10), (3.9) geht hervor, dass sowohl y als auch w in den laut Raumzerlegung geforderten Räumen liegen. Für u stellt sich die Frage, ob im $(PP_K^{S_1})$ ein invarianter Raum für (3.8) ist. Sei u eine Lösung von (3.8), die an einem Punkt $t_0 \in \Im$ die Bedingung $u(t_0) = PP_K^{N_1}(t_0)u(t_0)$ erfüllt. Man definiert $s(t) := (I - PP_K^{N_1}(t))u(t)$. Eine Ableitung von s liefert

$$\begin{aligned}
s' &= -(PP_K^{N_1})'u + (I - PP_K^{N_1})u' \\
&= -(PP_K^{N_1})'u + (I - PP_K^{N_1})\left[(PP_K^{N_1})'(u + Py) - PP_K^{N_1}G_{2,K}^{-1}Bu + PP_{N_1}^K G_{2,K}^{-1}q\right] \\
&= -(PP_K^{N_1})'u + (I - PP_K^{N_1})(PP_K^{N_1})'(u + Py) \\
&= -(PP_K^{N_1})'u + (PP_K^{N_1})'PP_K^{N_1}u + (I - PP_K^{N_1})(PP_K^{N_1})'Py \\
&= -(PP_K^{N_1})'(I - PP_K^{N_1})u + (I - PP_K^{N_1})(PP_K^{N_1})'PQ_{N_1}^K G_{2,K}^{-1}(q - Bu) \\
&= -(PP_K^{N_1})'s,
\end{aligned}$$

also erfüllt s die Anfangswertaufgabe

$$\begin{aligned}
s'(t) &= -(PP_K^{N_1}(t))'s(t), \quad t \in \Im, \\
s(t_0) &= 0,
\end{aligned}$$

woraus folgt, dass $s(t) \equiv 0 \Rightarrow u(t) \in \text{im}\,(PP_K^{N_1}(t))$ für alle t.

Aus der Entkopplung (3.8)-(3.10) kann man die drei wesentlichen Komponenten eines Index-2-Systems erkennen: die durch eine inhärente Differentialgleichung bestimmte, die rein algebraische und jene Komponente, die eine Differentiation benötigt. Die durch eine Differentiation bestimmte Komponente stellt einen qualitativen Unterschied zu den Index-1-Problemen dar, [Griepentrog and März, 1986], und kennzeichnet Index-2-Aufgaben als schlecht gestellt, wenn nur stetige Störungen betrachtet werden. Dieser Fakt motivierte die Anwendung von Regularisierungsansätzen auf ADGln mit höherem Index, [Hanke, 1991], [Hanke, 1992], [Hanke, 1994], [O'malley and Kalachev, 1994], [Kalachev and O'malley, 1996], [Ascher and Lin, 1993].

Bemerkung 3.1.4 *Die Entkopplung von [Tischendorf, 1996] ist ein Sonderfall von (3.8)-(3.10): Wenn $Q_{N_1}^K$ kanonisch gewählt wird, verschwindet der Term $PQ_{N_1}^{S_1}G_{2,S_1}^{-1}Bu$ in (3.10) und man erhält die Gleichungen aus [Tischendorf, 1996].*

Theorem 3.1.5 *Man betrachtet die Anfangswertaufgabe*

$$\begin{aligned} &A(t)(Px)'(t) + B(t)x(t) = q(t), \quad t \in \Im_0 := [t_0, \infty) \subset \Im, \quad (3.11)\\ &P(t_0)P_K^{N_1}(t_0)(x(t_0) - x^0) = 0, \quad x^0 \in I\!R^m, \end{aligned}$$

wobei $q \in \left\{p \in C(\Im, I\!R^m) : PQ_{N_1}^K G_{2,K}^{-1} p \in C^1(\Im, I\!R^m)\right\}$. *Seien* $N_1(t)$ *und* $S_1(t)$ *stetig differenzierbar.*

- *Dann ist (3.11) eindeutig lösbar in* $C_N^1(\Im_0, I\!R^m)$.
- *Falls die homogene Gleichung betrachtet wird, gilt für ihre Lösung*

$$x(t) = P_{sol}(t)u(t), \tag{3.12}$$

mit

$$P_{sol}(t) := \left(I - (QQ_{N_1}^{S_1})'(t) - QP_{S_1}^{N_1}(t)G_{2S_1}^{-1}(t)B(t)\right) PP_{S_1}^{N_1}(t).$$

$P_{sol}(t)$ ist ein Projektor auf den Lösungsraum der homogenen Gleichung längs $N \oplus N_1(t)$.

Beweis: D a $N_1(t)$ und $S_1(t)$ stetig differenzierbar sind, kann man C^1-Projektoren $Q_{N_1}^K(t)$ auf $N_1(t)$ längs einem entkoppelnden Komplementsraum $K(t)$ und $Q_{N_1}^{S_1}(t)$ wählen. Für die Gleichung (3.11) kann man nun die Entkopplung (3.8)-(3.10) realisieren und die AWA für die u-Komponente, mit dem Anfangswert $u(t_0) = PP_K^{N_1}(t_0)x^0$, ist eindeutig lösbar. Dies beweist die erste Aussage.

Nun betrachtet man den homogenen Fall ($q \equiv 0$) und $Q_{N_1}^K(t)$ als den kanonischen Projektor in (3.8)-(3.10). Aus der Entkopplung ergibt sich

$$\begin{aligned} y(t) &= -U(t)Q(t)G_{2,S_1}^{-1}(t)B_1(t)u(t),\\ w(t) &= -(QQ_{N_1}^{S_1})'(t)u(t) - T(t)Q(t)P_{S_1}^{N_1}(t)G_{2,S_1}^{-1}(t)B(t)u(t), \end{aligned}$$

woraus folgt

$$\begin{aligned} x(t) &= u(t) - U(t)Q(t)G_{2,S_1}^{-1}(t)B_1(t)u(t) - (QQ_{N_1}^{S_1})'(t)u(t) - T(t)Q(t)P_{S_1}^{N_1}(t)\\ &\quad \times G_{2,S_1}^{-1}(t)B(t)u(t)\\ &= \left(I - (QQ_{N_1}^{S_1})'(t) - Q(t)P_{S_1}^{N_1}(t)G_{2,S_1}^{-1}(t)B(t)\right) P(t)P_{S_1}^{N_1}(t)u(t),\\ &= P_{sol}(t)u(t). \end{aligned}$$

Weiterhin gilt $P_{sol}(t)P_{sol}(t) = P_{sol}(t)$. Außerdem kann man auch

$$P_{sol} = \left(I - (QQ_{N_1}^{S_1})'PP_{S_1}^{N_1} - QP_{S_1}^{N_1}G_{2,S_1}^{-1}BPP_{S_1}^{N_1}\right)PP_{S_1}^{N_1}$$

schreiben, wobei die Matrix $\left(I - (QQ_{N_1}^{S_1})'PP_{S_1}^{N_1} - QP_{S_1}^{N_1}G_{2,S_1}^{-1}BPP_{1,S_1}\right)$ nach dem Lemma 2.1.5 regulär ist, und dadurch

$$\ker P_{sol}(t) = \ker PP_{S_1}^{N_1}(t) = N \oplus N_1(t)$$

gilt.

Bemerkung 3.1.6

1. Für jeden Punkt $x_0 \in \operatorname{im}(P_{sol}(t_0))$ (und nur für diese Punkte) existiert genau eine Trajektorie der homogenen Gleichung, die durch x_0 geht.

2. Man nehme wieder den Hessenberg-Index-2-Fall (Beispiel 2.1.9). Der kanonische Projektor hat hier die Struktur, [Hanke et al., 1998],

$$Q_{N_1}^{S_1}(t) = \begin{pmatrix} H(t) & 0 \\ -F(t) & 0 \end{pmatrix},$$

wobei

$$F := (B_{21}B_{12})^{-1}B_{21} \quad \textit{und} \quad H := B_{12}F.$$

Der Projektor P_{sol} ist dann durch

$$P_{sol=}\begin{pmatrix} I_r - H & 0 \\ (F' - FB_{11})(I_r - H) & 0 \end{pmatrix}$$

gegeben, und die nicht-null Zeilen der inhärenten Differentialgleichung sind

$$u_1' + H'u_1 + (I-H)B_{11}u_1 = (I-H)q_1 - (H' + (I-H)B_{11})B_{12}(B_{21}B_{12})^{-1}q_2.$$

Entkopplung mit einem Projektor längs $S_1(t)$

In [März and Rodríguez Santiesteban, 1999] wurde diese Art von Entkopplung eingeführt, aber hier wird die Vorgehensweise etwas geändert. Man wählt $Q_{N_1}^K(t)$ in (3.5)-(3.7) als den kanonischen Projektor und betrachtet außerdem einen Projektor $Q_*^{S_1}(t)$, der längs $S_1(t)$ projiziert. In

[März and Rodríguez Santiesteban, 1999] wurde mit einer besonderen Wahl dieses Projektors gearbeitet, nämlich mit dem Orthoprojektor $Q_{S_1^\perp}^{S_1}(t)$ längs $S_1(t)$.

Weiterhin wird der Komplementsprojektor $P_{S_1}^*(t) := I - Q_*^{S_1}(t)$ definiert. Man sammelt einige nützliche Eigenschaften der eingeführten Projektoren.

$$\begin{aligned} Q_{N_1}^{S_1}(t)Q_*^{S_1}(t) = Q_{N_1}^{S_1}(t) \quad &, \quad Q_*^{S_1}(t)Q_{N_1}^{S_1}(t) = Q_*^{S_1}(t), \\ P_{S_1}^{N_1}(t)P_{S_1}^*(t) = P_{S_1}^*(t) \quad &, \quad P_{S_1}^*(t)P_{S_1}^{N_1}(t) = P_{S_1}^{N_1}(t), \\ Q_*^{S_1}(t)P_{S_1}^{N_1}(t) = 0 \quad &, \quad Q_{N_1}^{S_1}(t)P_{S_1}^*(t) = 0. \end{aligned}$$

Lemma 3.1.7 *Es gelten die folgenden Aussagen:*

1. $S_1(t) = \{z \in I\!R^m : W_1(t)B(t)z = 0\}$ *und folglich* $\ker W_1B = \ker Q_*^{S_1} = S_1$,

2. $\operatorname{im} W_1B = \operatorname{im} W_1$.

Beweis: P er Definition von $S_1(t)$ hat man

$$S_1(t) = \{z \in I\!R^m : B_1z \in \operatorname{im} G_1(t)\}.$$

Da $W_1(t)$ längs $\operatorname{im} G_1(t)$ projiziert, folgt daraus die erste Aussage.

Offensichtlich ist $\operatorname{im} W_1B \subset \operatorname{im} W_1$. Die andere Inklusion folgt aus der Gleichung

$$\dim(\operatorname{im} W_1) = m - \dim S_1 = \dim(\operatorname{im} W_1B).$$

In dieser Art der Entkopplung gilt die $I\!R^m$-Zerlegung

$$\begin{aligned} I &= PP_{S_1}^* + PQ_*^{S_1} + UQ + TQ, \\ x &= PP_{S_1}^*x + PQ_*^{S_1}x + UQx + TQx, \\ x &= z + (PQ_*^{S_1} + UQ)y + w, \end{aligned}$$

wobei $z := PP_{S_1}^*x$, $y := (UQ + PQ_{N_1}^{S_1})x$ und $w := TQx$. In dieser Entkopplungsart wird, wie in dem ersten Fall, von (3.5)- (3.7) ausgegangen. Wenn der Projektor $Q_*^{S_1}(t)$ stetig differenzierbar gewählt werden kann (also $S_1(t)$ besitzt eine C^1-Basis), transformiert man die Terme

$$\begin{aligned}
PP_{S_1}^{N_1}(Px)' &= PP_{S_1}^{N_1}P(Px)' = PP_{S_1}^{N_1}(PP_{S_1}^{*} + PQ_{*}^{S_1}(Px)' \\
&= PP_{S_1}^{N_1}PP_{S_1}^{*}(Px)' + PP_{S_1}^{N_1}Q_{*}^{S_1}(Px)' \\
&= PP_{S_1}^{*}(Px)' + PP_{S_1}^{N_1}Q_{*}^{S_1}(Px)' \\
&= (PP_{S_1}^{*}x)' - (PP_{S_1}^{*})'Px + PP_{S_1}^{N_1}(PQ_{*}^{S_1}x)' - PP_{S_1}^{N_1}Q_{*}^{S_{1,}}Px, \\
QQ_{N_1}^{S_1}(Px)' &= QQ_{N_1}^{S_1}Q_{*}^{S_1}(Px)' = QQ_{N_1}^{S_1}(PQ_{*}^{S_1}x)' - QQ_{N_1}^{S_1}Q_{*}^{S_{1,}}(Px),
\end{aligned}$$

und analog zu (3.8)-(3.10) erhält man

$$z' - ((PP_{S_1}^{*})' + PP_{S_1}^{N_1}Q_{*}^{S_{1'}})(z + Q_{*}^{S_1}y) \tag{3.13}$$

$$\begin{aligned}
-PP_{S_1}^{N_1}G_2^{-1}BPP_{S_1}^{N_1}(z + Q_{*}^{S_1}y) + PP_{S_1}^{N_1}(PQ_{*}^{S_1}y)' &= PP_{S_1}^{N_1}G_2^{-1}q,
\end{aligned}$$

$$(QQ_{N_1}^{S_1}Q_{*}^{S_{1'}} + TQP_{S_1}^{N_1}G_2^{-1}BPP_{S_1}^{N_1})(z + PQ_{*}^{S_1}y) \tag{3.14}$$

$$\begin{aligned}
+w - QQ_{N_1}^{S_1}(PQ_{*}^{S_1}y)' &= TQP_{S_1}^{N_1}G_2^{-1}q,
\end{aligned}$$

$$\begin{aligned}
(I + UQG_{2,S_1}^{-1}BPP_{S_1}^{N_1}Q_{*}^{S_1})y + UQG_2^{-1}BPP_{S_1}^{N_1}z &= (UQ + PQ_{N_1}^{S_1})G_2^{-1}q \qquad (3.15)
\end{aligned}$$

Die Koeffizientenmatrix von y in (3.15) ist nach dem Lemma 2.1.5 regulär und ihre Inverse ist

$$I - UQG_2^{-1}BPP_{S_1}^{N_1}Q_{*}^{S_1}.$$

Demzufolge, falls $PQ_{*}^{S_1}y = PQ_{*}^{S_1}G_2^{-1}q$ stetig differenzierbar ist, darf man diesen Ausdruck in (3.13) und (3.14) einsetzen und man wird, wenn auch die resultierenden Funktionen wieder zu den entsprechenden Räumen gehören, die Lösung von (3.2) bilden können.

Proposition 3.1.8 *$PQ_{*}^{S_1}G_2^{-1}q = PQ_{*}^{S_1}(W_1B)^{+}W_1q$.*

Beweis: F ür den Orthoprojektor $Q_{S_1^{\perp}}^{S_1}$ gilt

$$\begin{aligned}
Q_{S_1^{\perp}}^{S_1}G_2^{-1} &= (W_1B)^{+}W_1BG_2^{-1} = (W_1B)^{+}W_1BPQ_{N_1}^{S_1}G_2^{-1} \\
&= (W_1B)^{+}W_1G_2PQ_{N_1}^{S_1}G_2^{-1}BPQ_{N_1}^{S_1}G_2^{-1} \\
&= (W_1B)^{+}W_1G_2PQ_{N_1}^{S_1}G_2^{-1} = (W_1B)^{+}W_1,
\end{aligned}$$

dann hat man für einen beliebigen Projektor $Q_{*}^{S_1}$

$$PQ_{*}^{S_1}G_2^{-1}q = PQ_{*}^{S_1}Q_{S_1^{\perp}}^{S_1}G_2^{-1}q = PQ_{*}^{S_1}(W_1B)^{+}W_1q.$$

Lemma 3.1.9 *Der Unterraum $PS_1(t) = \mathrm{im}\,(PP_{S_1}^{*}(t))$ ist ein invarianter Raum für die inhärente Differentialgleichung (3.13).*

Beweis: S ei $s(t) := (I - PP^*_{S_1}(t))z(t)$ und $z(t_0) \in \mathrm{im}\,(PP^*_{S_1}(t_0))$. Wenn man (3.13) mit $I - PP^*_{S_1}$ multipliziert, bekommt man $((I - PP^*_{S_1})PP^{N_1}_{S_1} = 0)$

$$\begin{aligned}
(I - PP^*_{S_1})z' - (I - PP^*_{S_1})(PP^*_{S_1})'(z + Q^{S_1}_* y) &= 0,\\
s' + (PP^*_{S_1})'z - (I - PP^*_{S_1})(PP^*_{S_1})'z - (I - PP^*_{S_1})(PP^*_{S_1})'Q^{S_1}_* y &= 0,\\
s' + PP^*_{S_1}(PP^*_{S_1})'z &= 0,\\
s' + (PP^*_{S_1})'z - (PP^*_{S_1})'PP^*_{S_1}z &= 0,\\
s' + (PP^*_{S_1})'s &= 0,
\end{aligned}$$

und da $s(t_0) = 0$, folgt daraus $s(t) \equiv 0$.

Theorem 3.1.10 *Für die Anfangswertaufgabe*

$$\begin{aligned}
&A(t)(Px)'(t) + B(t)x(t) = q(t), \quad t \in \Im_0 := [t_0, \infty) \subset \Im, \quad (3.16)\\
&PP^*_{S_1}(t_0)(x(t_0) - x^0) = 0, \quad x^0 \in I\!R^m,
\end{aligned}$$

mit $S_1(t)$ stetig differenzierbar und $q \in \left\{p \in C(\Im, I\!R^m) : PQ^{S_1}_* G_2^{-1} p \in C^1(\Im, I\!R^m)\right\}$, existiert genau eine $C^1_N(\Im_0, I\!R^m)$-Lösung. Für die homogene AWA gilt der Ausdruck

$$x(t) = \overline{P}_{sol}(t)z(t),$$

mit

$$\overline{P}_{sol}(t) := (I - QQ^{S_1}_{N_1}(t)Q^{S_1,}_*(t) - QP^{N_1}_{S_1}(t)G_2^{-1}(t)B(t))PP^*_{S_1}(t).$$

$\overline{P}_{sol}$ *ist ein Projektor auf den Lösungsraum der homogenen Gleichung.*

Beweis: D er Beweis unterscheidet sich unwesentlich von jenem zu Theorem 3.1.5. Da $S_1(t)$ stetig differenzierbar ist, gibt es einen C^1-Projektor längs $S_1(t)$. Außerdem ist $PQ^{S_1}_* G_2^{-1} q$ stetig differenzierbar. Demzufolge können die Gleichungen (3.13)-(3.15) aufgestellt und nacheinander gelöst werden. Die Anfangsbedingung in (3.16) entspricht einer Bedingung für $z(t_0)$ in $\mathrm{im}\,(PP^*_{S_1}(t_0))$. Schließlich bildet man die Lösung von (3.2) als

$$x(t) = z(t) + (PQ^{S_1}_* + UQ)(t)y(t) + w(t).$$

Für den Fall $q(t) \equiv 0$ wird darauf aufmerksam gemacht, dass

$$y = -UQP^{N_1}_{S_1}G_2^{-1}Bz \text{ und } Q^{S_1}_* y = 0.$$

Die Lösung ist dann

$$\begin{aligned} x(t) &= z(t) + (PQ_*^{S_1}(t) + U(t)Q)y + w(t), \\ &= z(t) - U(t)QP_{S_1}^{N_1}(t)G_2^{-1}(t)B(t)z(t) - T(t)QP_{S_1}^{N_1}(t)G_2^{-1}(t)B(t)z(t) \\ &\quad -QQ_{N_1}^{S_1}(t)Q_*^{S_{1'}}(t)z(t), \\ &= \left(I - QQ_{N_1}^{S_1}(t)Q_*^{S_{1'}}(t) - QP_{S_1}^{N_1}(t)G_2^{-1}(t)B(t)\right)PP_{S_1}^*(t)z(t), \\ &= \overline{P}_{sol}(t)z(t). \end{aligned}$$

Es ist leicht zu überprüfen, dass $\overline{P}_{sol}$ ein Projektor ist. Weiterhin folgt aus der Regularität von

$$\left(I - QQ_{N_1}^{S_1}(t)Q_*^{S_{1'}}(t)PP_{S_1}^*(t) - QP_{S_1}^{N_1}(t)G_2^{-1}(t)B(t)PP_{S_1}^*(t)\right),$$

dass $\ker(\overline{P}_{sol}) = \ker(PP_{S_1}^*)$.

Bemerkung 3.1.11

1. Falls der kanonische Projektor $Q_{N_1}^{S_1}$ differenzierbar ist, gelten zwischen den Projektoren $\overline{P}_{sol}$ und P_{sol} die Beziehungen

$$P_{sol}(t) = \overline{P}_{sol}(t)PP_{S_1}^{N_1}(t) \quad , \quad \overline{P}_{sol}(t) = P_{sol}(t)PP_{S_1}^*(t).$$

2. Die Voraussetzungen dieses Theorems sind etwas schwächer als jene in Theorem 3.1.5. Die Glattheit von $N_1(t)$ war nicht nötig und es gilt

$$PQ_*^{S_1}G_2^{-1}q = PQ_*^{S_1}PQ_{N_1}^{S_1}G_2^{-1}q.$$

3. Die gleiche Aussage von Theorem 3.1.10 ist auch unter den Voraussetzungen, dass W_1B und W_1q stetig differenzierbar sind, gültig. Der Leser wird auf die Proposition 3.1.8 verwiesen. Das entspricht der analogen Aussage aus [März and Rodríguez Santiesteban, 1999].

4. Im Hessenberg-Fall kann man einen wesentlichen Unterschied zwischen den Theoremen 3.1.5 und 3.1.10 gut erkennen. Hier sind die Räume $N_1(t)$ und $S_1(t)$ durch

$$\begin{aligned} N_1(t) &= \left\{ \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} z \in I\!R^m : z_1 = -B_{12}(t)z_2 \right\}, \\ S_1(t) &= \left\{ \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} z \in I\!R^m : B_{21}(t)z_1 = 0 \right\}, \end{aligned}$$

gegeben. Das Theorem 3.1.5 verlangt, dass sowohl B_{12} als auch B_{21}

differenzierbar sein müssen, während für das Theorem 3.1.10 nur Differenzierbarkeit von B_{21} nötig ist.

3.1.2 Lokale Entkopplung quasilinearer Index-2-ADGln

Die kommende Analyse folgt den Techniken aus [Tischendorf, 1996]. Sie beruht auf dem Ansatz der Quasilinearisierung und gilt nur in einer Umgebung einer vorgegebenen Lösungstrajektorie. Hier wird diese Prozedur erweitert, indem man statt einer Entkopplung mit dem kanonischen Projektor $Q_{N_1}^{S_1}$ die zwei Varianten aus dem Abschnitt 3.1.1 einsetzt.

Man betrachtet die ADGl

$$A(t)(Px)'(t) + b(x(t), t) = 0, \quad t \in \Im, \tag{3.17}$$

wobei A, b zweimal stetig differenzierbar in $D \times \Im$ sind und D eine offene Menge in $\mathbb{R}^m$ ist. Sei außerdem (3.17) längs einer isolierten C_N^1-Lösung $x_*(t)$ tractable mit Index-2.

Alle relevanten Matrizen und Projektoren von Abschnitt 3.1.1 können für (3.17) als Funktion von (x, t) definiert werden. Sie werden in der Lösung $x_*(t)$ ausgewertet und als nur eine Funktion von t bezeichnet, zum Beispiel

$$B(t) := B(x_*(t), t) = b_x(x_*(t), t).$$

Es wird vorausgesetzt, dass es einen differenzierbaren Projektor $Q_{N_1}^K(t)$ auf $N_1(x_*(t), t)$ gibt, und sei $\hat{b}(x, t) := (U(t)Q + PQ_{N_1}^K(t))G_{2,K}^{-1}(t)b(x, t)$.

Man betrachtet für (3.17) die AWA

$$\begin{aligned} A(t)(Px)'(t) + b(x(t), t) = 0, \quad & t \in \Im_0 \\ Pr(t_0)(x(t_0) - x_0) = 0, & \end{aligned} \tag{3.18}$$

wobei $Pr(t)$ einer der Projektoren $PP_K^{N_1}(t)$ und $PP_{S_1}^*(t)$ sein kann.

Die AWA (3.18) kann wie folgt umgeschrieben werden:

$$\begin{aligned} &A(t)(Px)'(t) + B(t)x(t) + \hat{h}(x(t), t) - r_*(t) = 0, \\ &Pr(t_0)(x(t_0) - x_0) = 0, \end{aligned} \tag{3.19}$$

mit

$$\begin{aligned} &B(t) := b_x'(x_*(t), t), \\ &\hat{h}(x, t) := b(x, t) - b(x_*(t), t) - B(t)(x - x_*(t)), \\ &r_*(t) := A(t)(Px_*)'(t) + B(t)x_*(t). \end{aligned}$$

Durch die Umformulierung (3.19) sollen die linearen Techniken des Abschnitts 3.1.1 angewendet werden. Ab hier wird oft das Argument t ausgesetzt, um die Formel zu vereinfachen. Es wird unterstrichen, dass der Term $\hat{h}(x,t)$ in einer Lösungsumgebung sehr klein ist. Zusätzlich gilt

$$\begin{aligned} \hat{h}(x_*(t),t) &= 0, \\ \hat{h}'_x(x_*(t),t) &= 0, \quad \forall t \in \Im_0 \\ \hat{h}'_t(x_*(t),t) &= 0. \end{aligned}$$

Ebenso wie bei den linearen Aufgaben wird (3.19) mit $G_{2,K}^{-1}(t)$ skaliert und man erhält

$$P_K^{N_1} P(Px)' + G_{2,K}^{-1} B P P_K^{N_1} x + Q_{N_1}^K x + Qx + G_{2,K}^{-1} \hat{h}(x,\cdot) = G_{2,K}^{-1} r_*.$$

Als Nächstes wird diese Gleichung in die Komponenten $PP_K^{N_1}$, $TQP_K^{N_1}$ und $UQ + PQ_{N_1}^K$ zerlegt

$$\begin{aligned} PP_K^{N_1}(Px)' + PP_K^{N_1} G_{2,K}^{-1} B P P_K^{N_1} x & \qquad (3.20) \\ +PP_K^{N_1} G_{2,K}^{-1} \hat{h}(x,\cdot) &= PP_K^{N_1} G_{2,K}^{-1} r_*, \\ -QQ_{N_1}^K(Px)' + TQP_K^{N_1} G_{2,K}^{-1} B P P_K^{N_1} x + TQx & \qquad (3.21) \\ +TQP_K^{N_1} G_{2,K}^{-1} \hat{h}(x,\cdot) &= TQP_K^{N_1} G_{2,K}^{-1} r_*, \\ (UQ + PQ_{N_1}^K) G_{2,K}^{-1} B P P_K^{N_1} x + (UQ + PQ_{N_1}^K) x & \qquad (3.22) \\ +(UQ + PQ_{N_1}^K) G_{2,K}^{-1} \hat{h}(x,\cdot) &= (UQ + PQ_{N_1}^K) G_{2,K}^{-1} r_*. \end{aligned}$$

Dieses System ist analog zu (3.5)-(3.7) in dem quasilinearen Fall. Natürlich wird jetzt die Entkopplung technisch komplizierter und nur lokal unter strukturellen Voraussetzungen möglich sein.

Lokale Entkopplung mit einem Projektor auf $N_1(t)$

Wie im linearen Fall führt man die Variablen $u := PP_K^{N_1} x$, $y := (UQ + PQ_{N_1}^K)x$ und $w := TQx$ ein. Falls der Projektor $Q_{N_1}^K$ stetig differenzierbar ist, kann das System (3.20)-(3.22) folgendermaßen geschrieben werden

$$\begin{aligned} u' + PP_K^{N_1} G_{2,K}^{-1} B u + PP_K^{N_1} G_{2,K}^{-1} \hat{h}(u+y+w,\cdot) & \qquad (3.23) \\ -(PP_K^{N_1})'(u+Py) &= PP_K^{N_1} G_{2,K}^{-1} r_*, \\ -QQ_{N_1}^K(Py)' + (QQ_{N_1}^K)' u + TQP_K^{N_1} G_{2,K}^{-1} B u + w & \qquad (3.24) \end{aligned}$$

$$
\begin{aligned}
+TQP_K^{N_1}G_{2,K}^{-1}\hat{h}(u+y+w,\cdot) &= TQP_K^{N_1}G_{2,K}^{-1}r_*, \\
y+(UQ+PQ_{N_1}^K)G_{2,K}^{-1}\hat{h}(u+y+w,\cdot) & \qquad (3.25)\\
+(UQ+PQ_{N_1}^K)G_{2,K}^{-1}Bu &= (UQ+PQ_{N_1}^K)G_{2,K}^{-1}r_*.
\end{aligned}
$$

Dieses System stellt wegen der nicht-linearen Terme noch keine Entkopplung dar. Es wird

$$
\begin{aligned}
F(u,y,w,\cdot) &:= y+(UQ+PQ_{N_1}^K)G_{2,K}^{-1}(\hat{h}(u+y+w,\cdot)-r_*) \\
&\quad +(UQ+PQ_{N_1}^K)G_{2,K}^{-1}Bu \\
&= y-y_*+(UQ+PQ_{N_1}^K)G_{2,K}^{-1}B(u-u_*) \\
&\quad +(UQ+PQ_{N_1}^K)G_{2,K}^{-1}\hat{h}(u+y+w,\cdot)
\end{aligned}
$$

definiert und für F gilt

$$
\begin{aligned}
F(u_*(t),y_*(t),w_*(t),t) &= 0, \\
F_y'(u_*(t),y_*(t),w_*(t),t) &= I, \\
F_u'(u_*(t),y_*(t),w_*(t),t) &= (U(t)Q+PQ_{N_1}^K(t))G_{2,K}^{-1}(t)B(t), \\
F_t'(u_*(t),y_*(t),w_*(t),t) &= -y_*'(t)-(U(t)Q+PQ_{N_1}^K(t))G_{2,K}^{-1}(t)B(t)u_*'(t),
\end{aligned}
$$

für alle $t\in\Im_0$. Des Weiteren gilt:

$$
\begin{aligned}
F_y'(u,y,w,t) &= I+\hat{H}(u,y,w,t), \\
F_w'(u,y,w,t) &= \hat{H}(u,y,w,t)T(t)Q,
\end{aligned}
$$

wobei die Funktion $\hat{H}$ wie folgt definiert ist:

$$
\begin{aligned}
\hat{H}(u,y,w,t) &:= (U(t)Q+PQ_{N_1}^K(t))G_{2,K}^{-1}(t)\hat{h}_x'(u+y+w,t) \\
&= \hat{g}_x'(u+y+w,t)-\hat{g}_x'(x_*(t),t).
\end{aligned}
$$

Nun gibt es nach dem Theorem über die impliziten Funktionen für alle $t\in\Im_0$ in einer Umgebung des Punktes $(u_*(t),y_*(t),w_*(t))$ eine stetig differenzierbare Funktion

$$
f\ :\ U_{\rho_1(t)}(u_*(t),w_*(t))\ \rightarrow\ I\!R^m,
$$

$$
U_{\rho_1(t)}(u_*(t),w_*(t)) := \{(u,w)\ :\ \|(u-u_*(t)\|+\|w-w_*(t)\|<\rho_1(t)\}\,,
$$

so dass

$$\begin{aligned} F(u, f(u, w, t), w, t) &= 0, \\ f(u_*(t), w_*(t), t) &= y_*(t), \end{aligned}$$

für (u, w) in der Umgebung $U_{\rho_1(t)}(u_*(t), w_*(t))$.

Bedingung 3.1.12 *Der Radius der Umgebung $\rho_1(t)$ kann gleichmäßig in t genommen werden.*

Diese Annahme wird ebenfalls für jede Anwendung des Theorems über die impliziten Funktionen in diesem Kapitel vorausgesetzt.

Als Nächstes soll f als y in (3.24) eingesetzt und diese Gleichung nach w gelöst werden. Da in dieser Gleichung die Ableitung von y vorkommt, muss man im Allgemeinen die Differenzierbarkeit von $w(\cdot)$ verlangen. Dieser Fakt widerspricht grundsätzlich die Struktur der Gleichung (3.9), wo nur die Ableitung von Py vorkommt. Um diesen Konflikt auszuschließen, wird eine strukturelle Voraussetzung wie in [Tischendorf, 1996] getroffen. Diese Bedingung ist nichts anderes als (2.22), was in diesem Abschnitt vollständig beleuchtet wird. Wenn die Bedingung (2.22) erfüllt ist, wird man die Index-2-Entkopplung lokal in einer Umgebung der Lösung x_* durchführen können. Die Bedingung ist, dass

$$Q_{N_1}^K(t)(I + \hat{b}'_x(x, t) - \hat{b}'_x(x_*(t), t))^{-1} T(t) Q = 0, \tag{3.26}$$

für alle $t \in \Im_0$ und x in einer Umgebung von x_*. Man beachte, dass (3.26) auch als

$$Q_{N_1}^K(t)(I + \hat{H}(u, y, w, t))^{-1} T(t) Q = 0$$

geschrieben werden kann.

Vorausgesetzt, dass (3.26) gilt, erfüllt die Funktion f die folgenden Eigenschaften:

$$\begin{aligned} f(u, w, t) &= (U(t)Q + PQ_{N_1}^K(t)) f(u, w, t), \\ (Pf)'_w(u, w, t) &= 0. \end{aligned}$$

Die erste Gleichung ist aus der Definition von F einfach nachzuvollziehen. Für die zweite geht man folgendermaßen vor:

$$\begin{aligned} (Pf_w)'(u, w, t) &= (PQ_{N_1}^K f'_w)(u, w, t) \\ &= -(PQ_{N_1}^K)(t)(I + \hat{H}(u, y, w, t))^{-1} \hat{H}(u, y, w, t) T(t) Q \\ &= -(PQ_{N_1}^K)(t) \left[I - (I + \hat{H}(u, y, w, t))^{-1} \right] T(t) Q \\ &= (PQ_{N_1}^K)(t)(I + \hat{H}(u, y, w, t))^{-1} T(t) Q = 0. \end{aligned}$$

Weiterhin erfüllt f

$$QQ_{N_1}^K(t)f_t'(u_*(t),w_*(t),t) = QQ_{N_1}^K(t)y_*'(t) + QQ_{N_1}^K(t)G_{2,K}^{-1}(t)B(t)u_*'(t),$$

für alle $t \in \Im_0$.

Aus den letzten Berechnungen ergibt sich, dass in einer Umgebung der Lösung die Funktion $Py = Pf(u,w,t)$ unabhängig von w ist. Um diese Tatsache deutlich zu machen, wird $Pf(u,0,t)$ geschrieben. Man führt jetzt die Notation

$$y_p(u,t) := Py$$

ein, wobei $y_p()$ wie f in $U_{\rho_1}(u_*,w_*)$ definiert ist.

Daraufhin setzt man $y = f(u,w,t)$ und $\hat{y}_p(u,u',t) := y_p'(u,t)$ in (3.23), (3.24) ein, und man erhält

$$\begin{aligned} u' + PP_K^{N_1}G_{2,K}^{-1}Bu + PP_K^{N_1}G_{2,K}^{-1}\hat{h}(u+f(u,w,\cdot)+w,\cdot) & \qquad (3.27)\\ -(PP_K^{N_1})'(u+Pf(u,0,\cdot)) &= PP_K^{N_1}G_{2,K}^{-1}r_*,\\ (QQ_{N_1}^K)'u + w + TQP_K^{N_1}G_{2,K}^{-1}(Bu+\hat{h}(u+f(u,w,\cdot)+w,\cdot)) & \qquad (3.28)\\ -QQ_{N_1}^K\hat{y}_p &= TQP_K^{N_1}G_{2,K}^{-1}r_*. \end{aligned}$$

Es wird die Funktion

$$G(u,\hat{y}_p,w,\cdot) := (QQ_{N_1}^K)'u+w+TQP_K^{N_1}G_{2,K}^{-1}(Bu+\hat{h}(u+f(u,w,\cdot)+w,\cdot)-r_*)-QQ_{N_1}^K\hat{y}_p$$

definiert, und für diese Funktion G gilt

$$G(u_*(t),(Py_*)'(t),w_*(t),t) = 0, \quad G_w'(u_*(t),w_*(t),t) = I.$$

Nun kann man erneut das Theorem über die impliziten Funktionen anwenden, wonach eine Umgebung

$$U_{\rho_2}(u_*,(Py_*)') := \{(u,\hat{y}_p,t) \,:\, \|u-u_*(t)\| < \rho_2,\ \hat{y}_p \in I\!R^m,\ t \in \Im_0\}$$

und eine Funktion

$$g \,:\, U_{\rho_2}(u_*,(Py_*)') \,\rightarrow\, I\!R^m,$$

mit

$$G(u,\hat{y}_p,g(u,\hat{y}_p,\cdot),\cdot) = 0, \quad w_*(t) = g(u_*(t),(Py_*)'(t),t)$$

existieren. Für die Funktion g gilt außerdem

$$g(u,\hat{y}_p,t) = T(t)Qg(u,\hat{y}_p,t),$$

für alle $(u, \hat{y}_p, t) \in U_{\rho_2}(u_*, (Py_*)')$, wie aus der Definition von $G(\cdot)$ hervorgeht.

Schließlich setzt man $w = g(u, \hat{y}_p, \cdot)$ in (3.27) ein, und es ergibt sich

$$\begin{aligned} & u' + PP_K^{N_1} G_{2,K}^{-1} Bu - (PP_K^{N_1})'(u + y_p(u, \cdot)) \qquad (3.29) \\ + PP_K^{N_1} G_{2,K}^{-1} \hat{h}(u + f(u, g(u, \hat{y}_p(u, u', \cdot), \cdot), \cdot) + g(u, \hat{y}_p(u, u', \cdot), \cdot), \cdot) \;\; &= \;\; PP_K^{N_1} G_{2,K}^{-1} r_*. \end{aligned}$$

Die Gleichung (3.29) stellt lokal eine implizite Differentialgleichung für die u-Komponente der Lösung von (3.18) dar. Außerdem soll an dieser Stelle betont werden, dass $\hat{y}_p$ eine Funktion von (u, u', t) ist,

$$\hat{y}_p = (Py)' = (Pf(u, 0, t))' = P' f(u, 0, t) + Pf_u(u, 0, t)u' + Pf_t(u, 0, t).$$

Die Jacobi-Matrix von (3.29) nach u' ist

$$J := I + \left[PP_K^{N_1} G_{2,K}^{-1} \hat{h}_x (I + f_w) \right] (g_{\hat{y}_p} Pf_u),$$

und längs der Trajektorie $(u_*(t), u_*'(t), t)$ gilt

$$J(u_*(t), u_*'(t), t) = I.$$

Hier wird analog zu dem Theorem über die impliziten Funktionen angenommen, dass die Gleichung (3.29) zumindest in einer t-gleichmäßigen Umgebung

$$U_{\rho_3}(u_*) := \{u \;:\; \|u - u_*(t)\|, \; t \in \Im_0\}$$

eine implizite GDGl für u und keine ADGl darstellt. Im Endeffekt ist die durchgeführte Entkopplung mindestens in einer Umgebung der Lösung gültig, die die Erfüllung der Bedingungen für U_{ρ_1}, U_{ρ_2} und U_{ρ_3} garantiert.

Hier ist, ebenso wie in dem linearen Fall, im $(PP_K^{N_1})$ ein invarianter Unterraum für (3.29). Wenn man (3.29) mit $(I - PP_K^{N_1}(t))$ multipliziert, bekommt man

$$\begin{aligned} (I - PP_K^{N_1})u' - (I - PP_K^{N_1})(PP_K^{N_1})'(u + Pf(u, g(u, \hat{y}_p, \cdot), \cdot)) &= 0, \\ (I - PP_K^{N_1})u' - (I - PP_K^{N_1})(PP_K^{N_1})'u &= 0, \\ \left[(I - PP_K^{N_1})u\right]' + (PP_K^{N_1})'u - (PP_K^{N_1})'PP_K^{N_1}u &= 0, \\ \left[(I - PP_K^{N_1})u\right]' + (PP_K^{N_1})'(I - PP_K^{N_1})u &= 0. \end{aligned}$$

Sei $s := (I - PP_K^{N_1})u$, dann hat man für s die folgende AWA:

$$\begin{aligned} & s' + (PP_K^{N_1})'s = 0, \quad t \in \Im_0, \\ & s(t_0) = 0, \end{aligned}$$

und das bedeutet $s(t) \equiv 0$, also $u(t) \in \operatorname{im} PP_K^{N_1}(t)$ für alle $t \in \Im_0$.

So ist man zu einer nicht-linearen Entkopplung von (3.17) gelangt. Die u-Komponente ist durch (3.29) gegeben und die restlichen durch

$$\begin{aligned} Py(t) &= Pf(u(t), 0, t), && (3.30)\\ w(t) &= g(u(t), (Py)'(t), t), && (3.31)\\ Qy(t) &= Qf(u(t), w(t), t). && (3.32) \end{aligned}$$

Lokale Entkopplung mit einem Projektor längs S_1

In diesem Fall, wie bei den linearen Problemen, wählt man $Q_{N_1}^K$ in (3.20)-(3.22) als den kanonischen Projektor $Q_{N_1}^{S_1}$. Es werden die Projektoren $P_{S_1}^*$, $Q_*^{S_1}$ längs der Lösung definiert und

$$\begin{aligned} z(t) &:= PP_{S_1}^*(t)x(t),\\ y(t) &:= (UQ + PQ_{N_1}^{S_1})(t)x(t),\\ w(t) &:= T(t)Qx(t). \end{aligned}$$

Für die Lösung gilt dann der Ausdruck

$$x(t) = z(t) + (UQ + PQ_*^{S_1})(t)y(t) + w(t).$$

Es ist in diesem Fall, analog zu dem vorhergehenden Abschnitt, eine strukturelle Bedingung anzunehmen. Bei dieser Art der Entkopplung sieht die Bedingung etwas komplizierter als (3.26) aus. Nichtsdestotrotz sind beide äquivalent. (Für den Beweis und die Bedeutung der strukturellen Bedingung siehe Abschnitt 3.1.2).

Es wird vorausgesetzt, dass

$$Q_{N_1}^{S_1}(t)(I+(UQG_2^{-1}BPP_{S_1}^{N_1}Q_*^{S_1})(t)+\hat{H}(x,t)(PQ_*^{S_1}+UQ)(t))^{-1}\hat{H}(x,t)T(t)Q = 0, \quad (3.33)$$

in einer Umgebung der Lösung erfüllt ist.

Das System (3.20)-(3.22) kann nun folgendermaßen transformiert werden:

$$\begin{aligned} z' - ((PP_{S_1}^*)' + PP_{S_1}^{N_1}Q_*^{S_1\prime} - PP_{S_1}^{N_1}G_2^{-1}BPP_{S_1}^{N_1})(z + Q_*^{S_1}y) && (3.34)\\ +PP_{S_1}^{N_1}(PQ_*^{S_1}y)' + PP_{S_1}^{N_1}G_2^{-1}\hat{h}(z + (UQ + PQ_*^{S_1})y + w, \cdot) &= PP_{S_1}^{N_1}G_2^{-1}r_*, \\ w + (QQ_{N_1}^{S_1}Q_*^{S_1\prime} + TQP_{S_1}^{N_1}G_2^{-1}BPP_{S_1}^{N_1})(z + PQ_*^{S_1}y) && (3.35)\\ -QQ_{N_1}^{S_1}(PQ_*^{S_1}y)' + TQP_{S_1}^{N_1}G_2^{-1}\hat{h}(z + (UQ + PQ_*^{S_1})y + w, \cdot) &= TQP_{S_1}^{N_1}G_2^{-1}r_*, \end{aligned}$$

$$
\begin{aligned}
&(I + UQG_2^{-1}BPP_{S_1}^{N_1}Q_*^{S_1})y + UQG_2^{-1}BPP_{S_1}^{N_1}z \qquad (3.36)\\
&+(UQ + PQ_{N_1}^{S_1})G_2^{-1}\hat{h}(z + (UQ + PQ_*^{S_1})y + w, \cdot) \;=\; (UQ + PQ_{N_1}^{S_1})G_2^{-1}r_*.
\end{aligned}
$$

Als Nächstes wird die Funktion

$$
\begin{aligned}
F(z, y, w, \cdot) \;:=\; & (I + UQG_2^{-1}BPP_{S_1}^{N_1}Q_*^{S_1})y + UQG_2^{-1}BPP_{S_1}^{N_1}z\\
& +(UQ + PQ_{N_1}^{S_1})G_2^{-1}\hat{h}(z + (UQ + PQ_*^{S_1})y + TQw, \cdot)\\
& -(UQ + PQ_{N_1}^{S_1})G_2^{-1}r_*,\\
=\; & (I + UQG_2^{-1}BPP_{S_1}^{N_1}Q_*^{S_1})(y - y_*) + UQG_2^{-1}BPP_{S_1}^{N_1}(z - z_*)\\
& +(UQ + PQ_{N_1}^{S_1})G_2^{-1}\hat{h}(z + (UQ + PQ_*^{S_1})y + TQw, \cdot)
\end{aligned}
$$

definiert. Diese Funktion erfüllt die Gleichungen

$$
\begin{aligned}
F(z_*(t), y_*(t), w_*(t), t) &= 0,\\
F_y'(z_*(t), y_*(t), w_*(t), t) &= I + UQG_2^{-1}BPP_{S_1}^{N_1}Q_*^{S_1},
\end{aligned}
$$

wobei die Matrix $I + UQG_2^{-1}BPP_{S_1}^{N_1}Q_*^{S_1}$ nach dem Lemma 2.1.5 für alle $t \in \Im_0$ regulär ist.

Man bekommt durch das Theorem über die impliziten Funktionen (und unter der Bedingung 3.1.12), in einer Umgebung der Lösung, eine Applikation

$$
f \,:\, U_{\rho_1}(z_*, w_*) \,\rightarrow\, I\!R^m,
$$

mit

$$
U_{\rho_1}(z_*, w_*) := \{(z, w, t) \,:\, \|z - z_*(t)\| + \|w - w_*(t)\| < \rho_1,\; t \in \Im_0\}\,.
$$

Die implizit gegebene Funktion f erfüllt in dieser Umgebung

$$
\begin{aligned}
F(z, f(z, w, t), w, t) &= 0,\\
f(z_*(t), w_*(t), t) &= y_*(t),
\end{aligned}
$$

für alle $(z, w, t) \in U_{\rho_1}(z_*, w_*)$.

Weiterhin gilt für F

$$
\begin{aligned}
F_y(z, y, w, \cdot) \;=\; & I + UQG_2^{-1}BPP_{S_1}^{N_1}Q_*^{S_1} + \hat{H}(z + (PQ_*^{S_1} + UQ)y + TQw, \cdot)\\
& \times(PQ_*^{S_1} + UQ),\\
F_w(z, y, w, \cdot) \;=\; & \hat{H}(z + (PQ_*^{S_1} + UQ)y + TQw, \cdot)TQ,
\end{aligned}
$$

und dadurch ist $PQ_*^{S_1}y = PQ_*^{S_1}f(z,w,t)$ unter der Bedingung (3.33) unabhängig von w,

$$(PQ_*^{S_1}y)_w = -PQ_*^{S_1}Q_{N_1}^{S_1}(I+UQG_2^{-1}BPP_{S_1}^{N_1}Q_*^{S_1}+\hat{H}(PQ_*^{S_1}+UQ))^{-1}\hat{H}TQ = 0.$$

Man setzt $y = f(z,w,t)$, $y_v(z,t) := (PQ_*^{S_1}f)(z,0,t)$ und $\hat{y}_v(z,z',t) := y_v'(z,z',t)$ in das System (3.37), (3.38) ein, und es folgt

$$\begin{aligned} & z' - ((PP_{S_1}^*)' + PP_{S_1}^{N_1}Q_*^{S_1\prime} - PP_{S_1}^{N_1}G_2^{-1}BPP_{S_1}^{N_1})(z+y_v(z,\cdot)) && (3.37)\\ +PP_{S_1}^{N_1}\hat{y}_v(z,z',t) + PP_{S_1}^{N_1}G_2^{-1}\hat{h}(z+(UQ+PQ_*^{S_1})f(z,w,\cdot)+w,\cdot) &= PP_{S_1}^{N_1}G_2^{-1}r_*, \\ & w + (QQ_{N_1}^{S_1}Q_*^{S_1\prime} + TQP_{S_1}^{N_1}G_2^{-1}BPP_{S_1}^{N_1})(z+y_v(z,\cdot)) && (3.38)\\ -QQ_{N_1}^{S_1}\hat{y}_v(z,z',t) + TQP_{S_1}^{N_1}G_2^{-1}\hat{h}(z+(UQ+PQ_*^{S_1})f(z,w,\cdot)+w,\cdot) &= TQP_{S_1}^{N_1}G_2^{-1}r_*. \end{aligned}$$

Jetzt wird die Funktion

$$\begin{aligned} G(z,w,\hat{y}_v,\cdot) &:= -QQ_{N_1}^{S_1}\hat{y}_v + (QQ_{N_1}^{S_1}Q_*^{S_1\prime} + TQP_{S_1}^{N_1}G_2^{-1}BPP_{S_1}^{N_1})(z+y_v(z,\cdot)) + w \\ &\quad +TQP_{S_1}^{N_1}G_2^{-1}\hat{h}(z+(UQ+PQ_*^{S_1})f(z,w,\cdot)+w,\cdot) - TQP_{S_1}^{N_1}G_2^{-1}r_*, \\ &= w - w_* + QQ_{N_1}^{S_1}Q_*^{S_1\prime}(z-z_*+y_v(z,\cdot)-PQ_*^{S_1}y_*) \\ &\quad +TQP_{S_1}^{N_1}G_2^{-1}BPP_{S_1}^{N_1}(z-z_*+y_v(z,\cdot)-PQ_*^{S_1}y_*) \\ &\quad -QQ_{N_1}^{S_1}(\hat{y}_v-(PQ_*^{S_1}y_*)') + TQP_{S_1}^{N_1}G_2^{-1} \\ &\quad \times\hat{h}(z+(UQ+PQ_*^{S_1})f(z,w,\cdot)+w,\cdot) \end{aligned}$$

definiert und es gilt

$$\begin{aligned} G(z_*(t),w_*(t),(PQ_*^{S_1}y_*)'(t),t) &= 0, \\ G_w(z_*(t),w_*(t),(PQ_*^{S_1}y_*)'(t),t) &= I. \end{aligned}$$

Es wird erneut das Theorem über die impliziten Funktionen benutzt und es folgt, dass eine Funktion g existiert

$$g \,:\, U_{\rho_2}(z_*,(PQ_*^{S_1}y_*)') \,\to\, I\!R^m,$$

wobei

$$U_{\rho_2}(z_*,(PQ_*^{S_1}y_*)') := \{(z,\hat{y}_v,t) \,:\, \|u-u_*(t)\| < \rho_2,\ \hat{y}_v \in I\!R^m,\ t\in\Im_0\},$$

die

$$\begin{aligned} G(z,g(z,\hat{y}_v,t),\hat{y}_v,t) &= 0, \\ g(z_*(t),(PQ_*^{S_1}y_*)'(t),t) &= w_*(t), \end{aligned}$$

in $U_{\rho_2}(z_*, (PQ_*^{S_1} y_*)')$ erfüllt.

Nun setzt man $w = g(z, \hat{y}_v(z, z', t), t)$ in (3.39) ein und erhält

$$\begin{aligned} z' - ((PP_{S_1}^*)' + PP_{S_1}^{N_1} Q_*^{S_1\prime} - PP_{S_1}^{N_1} G_2^{-1} B P P_{S_1}^{N_1})(z + y_v(z,\cdot)) \\ + PP_{S_1}^{N_1} \hat{y}_v(z, z', \cdot) + PP_{S_1}^{N_1} G_2^{-1} \hat{h}(z + (UQ + PQ_*^{S_1}) \\ \times f(z, g(z, \hat{y}_v(z, z', \cdot), \cdot), \cdot) + g(z, \hat{y}_v(z, z', \cdot), \cdot), \cdot) &= PP_{S_1}^{N_1} G_2^{-1} r_*. \end{aligned} \tag{3.39}$$

Die Gleichung (3.39) ist eine implizite Differentialgleichung für z, die in einer Umgebung der Form

$$U_{\rho_3}(z_*) := \{z \in I\!R^m \,:\, \|z - z_*(t)\| < \rho_3,\, t \in \Im_0\}$$

eine GDGl ist (unter der t-Gleichmäßigkeitsvoraussetzung), weil die Jacobi-Matrix nach der Variablen z' längs der Lösung die Einheitsmatrix ist:

$$\begin{aligned} J(z,t) &= I + PP_{S_1}^{N_1} Q_*^{S_1} f_z + PP_{S_1}^{N_1} G_2^{-1} \hat{h}_x (UQ f_w + I) g_{\hat{y}_v} f_z, \\ &= I - PP_{S_1}^{N_1} (Q_*^{S_1} + G_2^{-1} \hat{h}_x (UQ f_w + I) g_{\hat{y}_v}) F_y^{-1} F_z, \\ J(z_*(t), t) &= I - PP_{S_1}^{N_1} Q_*^{S_1} (I - UQ G_2^{-1} B P P_{S_1}^{N_1} Q_*^{S_1}) UQ G_2^{-1} B P P_{S_1}^{N_1}, \\ &= I. \end{aligned}$$

Diese GDGl besitzt außerdem den invarianten Raum $\operatorname{im} PP_{S_1}^* = PS_1$ aus einem ähnlichen Argument wie in der Entkopplung längs N_1.

Um die lokale Entkopplung in diesem Fall zusammenzufassen, stellt man alle Gleichungen zusammen:

$$\begin{aligned} z' - ((PP_{S_1}^*)' + PP_{S_1}^{N_1} Q_*^{S_1\prime} - PP_{S_1}^{N_1} G_2^{-1} B P P_{S_1}^{N_1})(z + y_v(z,\cdot)) \\ + PP_{S_1}^{N_1} \hat{y}_v(z, z', \cdot) + PP_{S_1}^{N_1} G_2^{-1} \hat{h}(z + (UQ + PQ_*^{S_1}) \\ \times f(z, g(z, \hat{y}_v(z, z', \cdot), \cdot), \cdot) + g(z, \hat{y}_v(z, z', \cdot), \cdot), \cdot) &= PP_{S_1}^{N_1} G_2^{-1} r_*, \end{aligned}$$

$$\begin{aligned} y_v &= (PQ_*^{S_1} f)(z, 0, \cdot), \\ \hat{y}_v &= (PQ_*^{S_1} f)'(z, z', 0, \cdot), \\ w &= g(z, \hat{y}_v, \cdot), \\ y &= f(z, w, \cdot). \end{aligned}$$

Bedeutung der Strukturbedingungen

Bei der Entkopplung im Abschnitt 3.1.2 wurde angenommen, dass die Bedingung

$$Q^K_{N_1}(t)(I + \hat{b}'_x(x,t) - \hat{b}'_x(x_*(t),t))^{-1}T(t)Q = 0$$

in einer Umgebung der Trajektorie $(x_*(t), t)$ erfüllt ist. Diese Annahme unterscheidet sich von der in [Tischendorf, 1996] nur dadurch, dass der Projektor $Q^K_{N_1}$ nicht der kanonische sein muss. Jedoch kann man zeigen, dass die Bedingung (3.26) unabhängig von der Wahl des Projektors $Q^K_{N_1}$ ist und folglich beide Bedingungen gleichbedeutend sind. In der Tat kann die Strukturbedingung ebenso geschrieben werden:

$$Q^K_{N_1}(I + \hat{H}(x,t))^{-1}T(t)Q = 0,$$

und wenn $Q^K_{N_1}$, $Q^{K_a}_{N_1}$ zwei Projektoren auf N_1 und $\hat{H}$, $\hat{H}_a$ mit dem entsprechenden Projektor gerechnet sind, dann gilt

$$\hat{H} = (UQ + PQ^K_{N_1})G^{-1}_{2,K}\hat{h}_x = (UQ + PQ^{K_a}_{N_1})G^{-1}_{2,K}\hat{h}_x + PQ^K_{N_1}P^{N_1}_{K_a}G^{-1}_{2,K}\hat{h}_x,$$

und weil die Beziehung zwischen den Matrizen $G_{2,K}$ und G_{2,K_a} durch

$$\begin{aligned} G_{2,K} &= G_{2,K_a}(I + Q^K_{N_1}P^{N_1}_{K_a}), \\ G^{-1}_{2,K} &= (I - Q^K_{N_1}P^{N_1}_{K_a})G^{-1}_{2,K_a}, \end{aligned}$$

gegeben ist, bekommt man

$$\begin{aligned} \hat{H} &= \hat{H}_a - (UQ + PQ^{K_a}_{N_1})Q^K_{N_1}P^{N_1}_{K_a}G^{-1}_{2,K_a}\hat{h}_x + PQ^K_{N_1}P^{N_1}_{K_a}G^{-1}_{2,K_a}\hat{h}_x \\ &= \hat{H}_a - PQ^K_{N_1}P^{N_1}_{K_a}G^{-1}_{2,K_a}\hat{h}_x + PQ^K_{N_1}P^{N_1}_{K_a}G^{-1}_{2,K_a}\hat{h}_x = \hat{H}_a. \end{aligned}$$

Folglich gilt für die Strukturbedingung (wenn man die Argumente weglässt)

$$\begin{aligned} Q^K_{N_1}(I + \hat{H})^{-1}TQ &= Q^K_{N_1}(Q^{K_a}_{N_1} + P^{N_1}_{K_a})(I + \hat{H}_a)^{-1}TQ, \\ &= Q^{K_a}_{N_1}(I + \hat{H}_a)^{-1}TQ + Q^K_{N_1}P^{N_1}_{K_a}(I + \hat{H}_a)^{-1}TQ, \end{aligned}$$

und da

$$PP^{N_1}_{K_a}(I + \hat{H}_a)^{-1}TQ = 0,$$

erhält man

$$Q^K_{N_1}(I + \hat{H})^{-1}TQ = Q^{K_a}_{N_1}(I + \hat{H}_a)^{-1}TQ.$$

Bezüglich der Bedingung (3.33) für die Entkopplung mit Projektor längs S_1 gilt eine ähnliche Aussage. Diese Bedingung lautet

$$Q_{N_1}^{S_1}(t)(I+(UQG_2^{-1}BPP_{S_1}^{N_1}Q_*^{S_1})(t)+\hat{H}(x,t)(PQ_*^{S_1}+UQ)(t))^{-1}\hat{H}(x,t)T(t)Q=0,$$

in einer Umgebung der Trajektorie $(x_*(t),t)$.

Zuerst gilt, dass diese Bedingung für $Q_*^{S_1}=Q_{N_1}^{S_1}$ die Strukturbedingung von [Tischendorf, 1996] ergibt. In der Tat bekommt man in diesem Fall

$$Q_{N_1}^{S_1}(t)(I+\hat{H}(x,t)(PQ_{N_1}^{S_1}+UQ)(t))^{-1}\hat{H}(x,t)T(t)Q=0,$$

und da

$$(PQ_{N_1}^{S_1}+UQ)\hat{H}=\hat{H},$$

folgt

$$\begin{aligned}
& Q_{N_1}^{S_1}(t)(I+\hat{H}(x,t)(PQ_{N_1}^{S_1}+UQ)(t))^{-1}\hat{H}(x,t)T(t)Q\\
= \;& Q_{N_1}^{S_1}(t)\sum_{k=0}^{\infty}(-1)^k\left[\hat{H}(x,t)(PQ_{N_1}^{S_1}+UQ)(t)\right]^k\hat{H}(x,t)T(t)Q\\
= \;& Q_{N_1}^{S_1}(t)\sum_{k=0}^{\infty}(-1)^k\left[\hat{H}(x,t)\right]^k\hat{H}(x,t)T(t)Q\\
= \;& Q_{N_1}^{S_1}(t)(I+\hat{H}(x,t))^{-1}\hat{H}(x,t)T(t)Q\\
= \;& -Q_{N_1}^{S_1}(t)(I+\hat{H}(x,t))^{-1}T(t)Q.
\end{aligned}$$

Als Nächstes wird gezeigt, dass die Bedingung (3.33) auch unabhängig von der Wahl des Projektors $Q_*^{S_1}$ und deshalb die Bedingung von [Tischendorf, 1996] äquivalent zu (3.33) ist. Sei $Q_{*_a}^{S_1}$ ein anderer Projektor längs S_1 und $P_{S_1}^{*_a}$ dementsprechend definiert, dann gilt

$$\begin{aligned}
& Q_{N_1}^{S_1}(t)\left[I+(UQG_2^{-1}BPP_{S_1}^{N_1}Q_*^{S_1})(t)+\hat{H}(x,t)(PQ_*^{S_1}+UQ)(t)\right]^{-1}\\
& \times\hat{H}(x,t)T(t)Q\\
= \;& Q_{N_1}^{S_1}(t)\left[I+\hat{H}(x,t)(PQ_*^{S_1}+UQ)(t)(I-(UQG_2^{-1}BPP_{S_1}^{N_1}Q_*^{S_1})(t))\right]^{-1}\\
& \times\hat{H}(x,t)T(t)Q\\
= \;& Q_{N_1}^{S_1}(t)\left[I+\hat{H}(x,t)(PQ_*^{S_1}+UQ-UQG_2^{-1}BPP_{S_1}^{N_1}Q_*^{S_1})(t)\right]^{-1}\hat{H}(x,t)T(t)Q\\
= \;& Q_{N_1}^{S_1}(t)\left[I+\hat{H}(x,t)(PQ_{*_a}^{S_1}+UQ-UQG_2^{-1}BPP_{S_1}^{N_1}Q_{*_a}^{S_1})(t)\right.\\
& \left.+\hat{H}(x,t)(I-UQG_2^{-1}B)PP_{S_1}^{*_a}Q_*^{S_1}\right]^{-1}\hat{H}(x,t)T(t)Q
\end{aligned}$$

$$
\begin{aligned}
&= Q_{N_1}^{S_1}(t)\left\{\left[I+\hat{H}(x,t)(PQ_{*_a}^{S_1}+UQ-UQG_2^{-1}BPP_{S_1}^{N_1}Q_{*_a}^{S_1})(t)\right]\right.\\
&\quad\times\left[I+\left[I+\hat{H}(x,t)(PQ_{*_a}^{S_1}+UQ-UQG_2^{-1}BPP_{S_1}^{N_1}Q_{*_a}^{S_1})(t)\right]^{-1}\hat{H}(x,t)\right.\\
&\quad\left.\left.\times(I-UQG_2^{-1}B)PP_{S_1}^{*_a}Q_*^{S_1}\right]\hat{H}(x,t)T(t)Q\right\}^{-1}\\
&= Q_{N_1}^{S_1}(t)\left[I+\left[I+\hat{H}(x,t)(PQ_{*_a}^{S_1}+UQ-UQG_2^{-1}BPP_{S_1}^{N_1}Q_{*_a}^{S_1})(t)\right]^{-1}\hat{H}(x,t)\right.\\
&\quad\times(I-UQG_2^{-1}B)PP_{S_1}^{*_a}Q_*^{S_1}\Big]^{-1}Q_{N_1}^{S_1}(t)\left[I+\hat{H}(x,t)(PQ_{*_a}^{S_1}+UQ\right.\\
&\quad\left.-UQG_2^{-1}BPP_{S_1}^{N_1}Q_{*_a}^{S_1})(t)\right]^{-1}\hat{H}(x,t)T(t)Q\\
&= Q_{N_1}^{S_1}(t)\left[I+\left[I+\hat{H}(x,t)(PQ_{*_a}^{S_1}+UQ-UQG_2^{-1}BPP_{S_1}^{N_1}Q_{*_a}^{S_1})(t)\right]^{-1}\hat{H}(x,t)\right.\\
&\quad\times(I-UQG_2^{-1}B)PP_{S_1}^{*_a}Q_*^{S_1}\Big]^{-1}Q_{N_1}^{S_1}(t)\\
&\quad\times\left[I+(UQG_2^{-1}BPP_{S_1}^{N_1}Q_{*_a}^{S_1})(t)+\hat{H}(x,t)(PQ_{*_a}^{S_1}+UQ)(t)\right]^{-1}\hat{H}(x,t)T(t)Q,
\end{aligned}
$$

was die Unabhängigkeit vom Projektor $Q_*^{S_1}$ zeigt.

Nachdem die Äquivalenz der drei Strukturbedingungen gezeigt wurde, soll noch einmal auf [Tischendorf, 1996] verwiesen werden. An jener Stelle wurden wichtige Fälle diskutiert, die in der Strukturbedingung enthalten sind. Aus dieser Analyse geht hervor, dass zum Beispiel Hessenberg-Systeme und Modelle aus der Schaltkreissimulation die Strukturbedingung erfüllen.

3.2 Kommutativität zwischen Entkopplung und Diskretisierung

In dieser Sektion werden zwei der wichtigsten numerischen Verfahrensfamilien im Kontext der Index-2-ADGln untersucht. Konkreter gesagt wird die Gültigkeit des Diagramms anfangs des Kapitels untersucht.

Durch die verfeinerten Entkopplungsmethoden, die in der Sektion 3.1 eingeführt wurden, können die Ergebnisse von [Hanke et al., 1998] verbessert werden und zum Beispiel Modelle aus der Schaltkreissimulation in die Analyse einbezogen werden. Diese Aussagen sind schon zum Teil durch [März and Rodríguez Santiesteban, 1999] bekannt geworden.

3.2.1 BDF-Verfahren

Dieses lineare Mehrschrittverfahren stellte die ersten Ansätze dar, die für ADGl vorgeschlagen wurden, [Gear, 1971], und verfügen über eine langjährige erfolgreiche Anwendung auf dem Gebiet der Schaltkreissimulation. Ein BDF-Verfahren für die ADGl (3.1) wird folgendermaßen realisiert, [Griepentrog and März, 1986], [Brenan et al., 1989]:

$$A_i \frac{1}{h} \sum_{j=0}^{k} \alpha_j x_{i-j} + b(x_i, t_i) = 0, \quad i \geq k, \tag{3.40}$$

wobei $h > 0$ die Schrittweite, das Gitter durch $t_i = t_0 + hi$ gegeben ist und die Funktionswerte $x_0, \ldots, x_{k-1}$ bekannt sind.

Wie in dem kontinuierlichen Fall wird angenommen, dass diese AWA eine isolierte Lösung $x_*(\cdot)$ besitzt.

Die diskrete Gleichung (3.40) kann wieder folgendermaßen geschrieben werden:

$$A_i \frac{1}{h} \sum_{j=0}^{k} \alpha_j x_{i-j} + B_i x_i + \hat{h}(x_i, t_i) = r_{*,i}, \quad i \geq k, \tag{3.41}$$

wobei $\hat{h}$ und r_* wie im Abschnitt 3.1.2 definiert sind

$$\begin{aligned} \hat{h}(x,t) &:= b(x,t) - b(x_*(t),t) - B(t)(x - x_*(t)), \\ B(t) &:= b'_x(x_*(t),t), \\ r_*(t) &:= A(t)(Px_*)'(t) + B(t)x_*(t). \end{aligned}$$

Jetzt wird die Gleichung (3.41) wie in 3.1.2 entkoppelt und anschließend die diskrete und analytische Entkopplung verglichen. Es wird je nach Entkopplungsart angenommen, dass entweder N_1 oder S_1 konstant ist.

Wenn man (3.41) mit $(PP_K^{N_1} G_{2,K}^{-1})_i$, $(TQP_K^{N_1} G_{2,K}^{-1})_i$ und $\left[(UQ + PQ_{N_1}^K) G_{2,K}^{-1}\right]_i$ multipliziert, folgt

$$(PP_K^{N_1})_i \frac{1}{h} \sum_{j=0}^{k} \alpha_j P x_{i-j} + (PP_K^{N_1} G_{2,K}^{-1} B P P_K^{N_1} x)_i \tag{3.42}$$

$$+(PP_K^{N_1} G_{2,K}^{-1})_i \hat{h}(x_i, t_i) = (PP_K^{N_1} G_{2,K}^{-1} r_*)_i,$$

$$-(QQ_{N_1}^K)_i \frac{1}{h} \sum_{j=0}^{k} \alpha_j P x_{i-j} + (TQP_K^{N_1} G_{2,K}^{-1} B P P_K^{N_1} x)_i \tag{3.43}$$

$$+(TQx)_i + (TQP_K^{N_1} G_{2,K}^{-1})_i \hat{h}(x_i, t_i) = (TQP_K^{N_1} G_{2,K}^{-1} r_*)_i,$$

$$\begin{aligned}
&\left[(UQ + PQ_{N_1}^K)G_{2,K}^{-1}\right]_i (BPP_K^{N_1}x)_i + U_iQx_i \qquad (3.44)\\
&+PQ_{N_1}^K x_i + \left[(UQ + PQ_{N_1}^K)G_{2,K}^{-1}\right]_i \hat{h}(x_i,t_i) \;=\; \left[(UQ + PQ_{N_1}^K)G_{2,K}^{-1}r_*\right]_i .
\end{aligned}$$

Lokale Entkopplung mit einem Projektor auf N_1

In dieser Entkopplungsart wird angenommen, dass $N_1(x_*(t),t)$ konstant ist und es wird ein konstanter Projektor $Q_{N_1}^K$ ausgewählt. Das System (3.42)-(3.44) vereinfacht sich zu

$$\begin{aligned}
&\frac{1}{h}\sum_{j=0}^{k}\alpha_j u_{i-j} + (PP_K^{S_1}G_{2,K}^{-1}Bu)_i \qquad (3.45)\\
&+(PP_K^{S_1}G_{2,K}^{-1})_i\hat{h}(u_i+y_i+w_i,t_i) \;=\; (PP_K^{S_1}G_{2,K}^{-1}r_*)_i,\\
&-QQ_{N_1}^K\frac{1}{h}\sum_{j=0}^{k}\alpha_j Py_{i-j} + (TQP_K^{N_1}G_{2,K}^{-1}Bu)_i + w_i \qquad (3.46)\\
&+(TQP_K^{N_1}G_{2,K}^{-1})_i\hat{h}(u_i+y_i+w_i,t_i) \;=\; (TQP_K^{N_1}G_{2,K}^{-1}r_*)_i,\\
&\left[(UQ + PQ_{N_1}^K)G_{2,K}^{-1}\right]_i (Bu)_i + y_i \qquad (3.47)\\
&+\left[(UQ + PQ_{N_1}^K)G_{2,K}^{-1}\right]_i \hat{h}(u_i+y_i+w_i,t_i) \;=\; \left[(UQ + PQ_{N_1}^K)G_{2,K}^{-1}r_*\right]_i ,
\end{aligned}$$

wobei $u_i = (PP_K^{N_1}x)_i$, $y_i = (UQ + PQ_{N_1}^K)_i x_i$ und $w_i = T_iQx_i$.

Es wird darauf aufmerksam gemacht, dass (3.47) nichts anderes als (3.25) an der Stelle t_i ist. Dann ergibt sich dieselbe Gleichung

$$F(u_i, y_i, w_i, t_i) = 0$$

wie in 3.1.2 und es existiert die implizite Funktion

$$f : U_\rho(x_*) \to I\!R^m,$$

mit

$$\begin{aligned}
y_i &= f(u_i, w_i, t_i),\\
Py_i &= Pf(u_i, 0, t_i),
\end{aligned}$$

für $(u_i, w_i, t_i) \in U_\rho(x_*)$.

Nun wird $y_i = f(u_i, w_i, t_i)$ und

$$\hat{y}_{p,i}^h := \frac{1}{h}\sum_{j=0}^{k}\alpha_j Py_{i-j}.$$

in (3.46) eingesetzt

$$\frac{1}{h}\sum_{j=0}^{k}\alpha_j u_{i-j} + (PP_K^{N_1}G_{2,K}^{-1}Bu)_i \tag{3.48}$$

$$\begin{aligned} +(PP_K^{N_1}G_{2,K}^{-1})_i\hat{h}(u_i + f(u_i,w_i,t_i) + w_i, t_i) &= (PP_K^{N_1}G_{2,K}^{-1}r_*)_i, \\ -QQ_{N_1}^K\hat{y}_{p,i}^h + (TQP_K^{N_1}G_{2,K}^{-1}Bu)_i + w_i & \qquad (3.49) \\ +(TQP_K^{N_1}G_{2,K}^{-1})_i\hat{h}(u_i + f(u_i,w_i,t_i) + w_i, t_i) &= (TQP_K^{N_1}G_{2,K}^{-1}r_*)_i. \end{aligned}$$

Die Gleichung (3.49) ist nichts anderes als

$$G(u_i, \hat{y}_{p,i}^h, w_i, t_i) = 0,$$

in dem Fall N_1 konstant. Wie in 3.1.2 definiert diese Gleichung die Funktion

$$g \;:\; U_{\rho_2}(x_*) \;\rightarrow\; I\!R^m,$$

mit

$$w_i = g(u_i, \hat{y}_{p,i}^h, t_i).$$

Zuletzt kann w_i in (3.48) eingesetzt werden und man erhält

$$\frac{1}{h}\sum_{j=0}^{k}\alpha_j u_{i-j} + (PP_K^{N_1}G_{2,K}^{-1}Bu)_i \tag{3.50}$$

$$+(PP_K^{N_1}G_{2,K}^{-1})_i\hat{h}(u_i + f(u_i, g(u_i,\hat{y}_{p,i}^h,t_i),t_i) + g(u_i,\hat{y}_{p,i}^h,t_i),t_i) = (PP_K^{N_1}G_{2,K}^{-1}r_*)_i.$$

Nun kann man aus (3.50) u_i ausrechnen und für die anderen Komponenten gilt

$$\begin{aligned} Py_i &= Pf(u_i, 0, t_i), && (3.51)\\ \hat{y}_{p,i}^h &= \frac{1}{h}\sum_{j=0}^{k}\alpha_j Py_{i-j}, && (3.52)\\ w_i &= g(u_i, \hat{y}_{p,i}^h, t_i), && (3.53)\\ y_i &= f(u_i, w_i, t_i). && (3.54) \end{aligned}$$

Das System (3.50)-(3.54) stellt die BDF-Anwendung auf (3.29)-(3.32) dar, wenn $(Q_{N_1}^K)' = 0$. Mit der Anwendung des BDF-Verfahrens auf (3.29)-(3.32) ist gemeint, dass die Ableitung einer Funktion immer durch die BDF-Differenzen-Approximation zu ersetzen ist.

Die Analyse dieses Abschnittes zeigt in anderen Worten, dass, wenn der Raum N_1 längs der Lösung konstant ist, das Kommutativitätsdiagramm 3.1 für die lokale Entkopplung gilt. Falls $N_1(x_*(t),t)$ nicht konstant ist, dann gilt die Kommutativität mit dieser Entkopplungsart nicht mehr. In diesem Fall kommt der problematische Term

$$(PP_K^{N_1})'(u+y_p(u,\cdot))$$

in (3.29) vor und die Kommutativität ist verletzt.

Lokale Entkopplung mit einem Projektor längs S_1

In dieser Art der Entkopplung bezeichnet man

$$\begin{aligned} z_i &:= PP^*_{S_1}(t_i)x(t_i), \\ y_i &:= (UQ+PQ^{S_1}_{N_1})(t_i)x(t_i), \\ w_i &:= T(t_i)Qx(t_i). \end{aligned}$$

Es wird vorausgesetzt, dass $S_1(x_*(t),t)$ konstant ist und es wird der Projektor $Q_*^{S_1}$ konstant gewählt. Das System (3.20)-(3.22) kann dann folgendermaßen geschrieben werden:

$$\begin{aligned} &\frac{1}{h}\sum_{j=0}^{k}\alpha_j z_{i-j}+(PP^{N_1}_{S_1})_i\frac{1}{h}\sum_{j=0}^{k}\alpha_j(PQ^{S_1}_*y)_{i-j} && (3.55)\\ &+(PP^{N_1}_{S_1}G_2^{-1}BPP^{N_1}_{S_1})_i(z+Q^{S_1}_*y)_i+(PP^{N_1}_{S_1}G_2^{-1})_i \\ &\qquad\times\hat{h}(z_i+(UQ+PQ^{S_1}_*)_iy_i+w_i,t_i) &&= (PP^{N_1}_{S_1}G_2^{-1}r_*)_i,\\ &-(QQ^{S_1}_{N_1})_i\frac{1}{h}\sum_{j=0}^{k}\alpha_j(PQ^{S_1}_*y)_{i-j}+(TQP^{N_1}_{S_1}G_2^{-1}BPP^{N_1}_{S_1})_i && (3.56)\\ &\qquad\times(z+Q^{S_1}_*y)_i+w_i+(TQP^{N_1}_{S_1}G_2^{-1})_i \\ &\qquad\times\hat{h}(z_i+(UQ+PQ^{S_1}_*)_iy_i+w_i,t_i) &&= (TQP^{N_1}_{S_1}G_2^{-1}r_*)_i,\\ &(I+UQG_2^{-1}BPP^{N_1}_{S_1}Q^{S_1}_*)_iy_i+(UQG_2^{-1}BPP^{N_1}_{S_1})_iz_i && (3.57)\\ &\qquad+(UQ+PQ^{S_1}_{N_1})_iG^{-1}_{2,i} \\ &\qquad\times\hat{h}(z_i+(UQ+PQ^{S_1}_*)_iy_i+w_i,t_i) &&= (UQ+PQ^{S_1}_{N_1})_i(G_2^{-1}r_*)_i. \end{aligned}$$

Die dritte Gleichung ist nichts anderes als

$$F(z_i,y_i,w_i,t_i)=0.$$

Diese Gleichung kann nach y_i gelöst werden,

$$\begin{aligned} y_i &= f(z_i, w_i, t_i), \\ y_v(t_i) &= PQ_*^{S_1} y_i = PQ_*^{S_1} f(z_i, 0, t_i), \end{aligned}$$

für $(z_i, w_i, t_i) \in U_p(x_*)$.

Nun wird y_i und

$$\hat{y}_{v,i}^h := \frac{1}{h} \sum_{j=0}^{k} \alpha_j (PQ_*^{S_1} y)_{i-j}$$

in (3.56), (3.57) eingesetzt und man erhält das folgende System:

$$\begin{aligned} &\frac{1}{h} \sum_{j=0}^{k} \alpha_j z_{i-j} + (PP_{S_1}^{N_1} G_2^{-1} BPP_{S_1}^{N_1})_i (z_i + Q_*^{S_1} f(z_i, 0, t_i)) \qquad (3.58) \\ &\quad + (PP_{S_1}^{N_1})_i \hat{y}_{v,i}^h + (PP_{S_1}^{N_1} G_2^{-1})_i \hat{h} (z_i + (UQ + PQ_*^{S_1})_i \\ &\qquad \times f(z_i, w_i, t_i) + w_i, t_i) = (PP_{S_1}^{N_1} G_2^{-1} r_*)_i, \\ &\quad - (QQ_{N_1}^{S_1})_i \hat{y}_{v,i}^h + (TQP_{S_1}^{N_1} G_2^{-1} BPP_{S_1}^{N_1})_i (z_i + w_i \qquad (3.59) \\ &+ Q_*^{S_1} f(z_i, 0, t_i)) + (TQP_{S_1}^{N_1} G_2^{-1})_i \hat{h} (z_i + (UQ + PQ_*^{S_1})_i \\ &\qquad \times f(z_i, w_i, t_i) + w_i, t_i) = (TQP_{S_1}^{N_1} G_2^{-1} r_*)_i. \end{aligned}$$

Die Gleichung (3.60) ist genau

$$G(z_i, w_i, \hat{y}_{v,i}^h, t_i) = 0,$$

und für w_i gilt

$$w_i = g(z_i, \hat{y}_{v,i}^h, t_i)$$

in der Umgebung $U_{\rho_2}(x_*)$.

Beim Einsetzen von w_i in (3.59) erhält man schließlich

$$\begin{aligned} &\frac{1}{h} \sum_{j=0}^{k} \alpha_j z_{i-j} + (PP_{S_1}^{N_1})_i \hat{y}_{v,i}^h + (PP_{S_1}^{N_1} G_2^{-1} BPP_{S_1}^{N_1})_i \qquad (3.60) \\ &\qquad \times (z_i + Q_*^{S_1} f(z_i, 0, t_i)) + (PP_{S_1}^{N_1} G_2^{-1})_i \\ &\times \hat{h} (z_i + (UQ + PQ_*^{S_1})_i f(z_i, g(z_i, \hat{y}_{v,i}^h, t_i), t_i) + g(z_i, \hat{y}_{v,i}^h, t_i), t_i) = (PP_{S_1}^{N_1} G_2^{-1} r_*)_i, \end{aligned}$$

was die BDF-Anwendung auf (3.39) darstellt. Man versteht erneut darunter, dass die Ableitung von $y_v(z(t), t) = PQ_*^{S_1} f(z(t), 0, t)$ nach t durch

$$\hat{y}_{v,i}^h = \frac{1}{h} \sum_{j=0}^{k} \alpha_j PQ_*^{S_1} f(z_{i-j}, 0, t_{i-j})$$

ersetzt wird.

Analog zu dem Fall, in dem $N_1(x_*(t),t)$ konstant ist, erhält man das Ergebnis, dass das Diagramm 3.1 für die Entkopplung mit Projektor längs S_1 gilt, wenn der Raum $S_1(x_*(t),t)$ konstant ist. Wenn diese Bedingung nicht erfüllt ist, kommen die Terme

$$(PP^*_{S_1})', \quad PP^{N_1}_{S_1}Q^{S_{1'}}_*,$$

in (3.39) vor. Somit ist die Kommutativität bei dieser Entkopplungsart verletzt.

3.2.2 Runge-Kutta-Verfahren

Für die IRK-Verfahren ist ebenfalls eine analoge Analyse möglich. In der Sektion 2.3 wurden die IRK-Verfahren für ADGln eingeführt. Dort wird das Verfahren, vorausgesetzt, dass die Matrix $(\alpha_{ij})_{i,j=1}^s$ regulär ist, wie folgt umformuliert:

$$x_l = \rho x_{l-1} + \sum_{i=1}^{s}\sum_{j=1}^{s} \beta_i \hat{\alpha}_{ij} X_{lj}, \qquad (3.61)$$

$$A_{li}\frac{1}{h}\sum_{j=1}^{s}\hat{\alpha}_{ij}(X_{lj}-x_{l-1}) + b_{li} = 0,\ i=1,\ldots,s. \qquad (3.62)$$

Dabei handelt es sich bei dem Ausdruck

$$\frac{1}{h}\sum_{j=1}^{s}\hat{\alpha}_{ij}(X_{lj}-x_{l-1})$$

um X'_{li} und damit um eine RK-Approximation für die Ableitung von $x(t)$ an der Stelle t_{li}.

Die Gleichung (3.62) unterscheidet sich unwesentlich von der BDF-Diskretisierung (3.40). Deswegen ist die Entkopplung in einer analogen Weise möglich.

Wie bei den BDF-Verfahren wird angenommen, dass eine AWA für die ADGl eine isolierte Lösung $x_*(\cdot)$ besitzt.

Die diskrete Gleichung (3.62) kann folgendermaßen umgeschrieben werden:

$$A_{li}\frac{1}{h}\sum_{j=1}^{s}\hat{\alpha}_{ij}(X_{lj}-x_{l-1}) + (BX)_{li} + \hat{h}(X_{li},t_{li}) = r_{*,li}, \quad i=1,\ldots,s, \qquad (3.63)$$

wobei $\hat{h}$ und r_* wie in 3.1.2 definiert sind:

$$\begin{aligned}\hat{h}(x,t) &:= b(x,t) - b(x_*(t),t) - B(t)(x - x_*(t)),\\ B(t) &:= b'_x(x_*(t),t),\\ r_*(t) &:= A(t)x'_*(t) + B(t)x_*(t).\end{aligned}$$

In den kommenden Abschnitten wird die Gleichung (3.63) wie in 3.1.2 entkoppelt und anschließend werden die diskreten und analytischen Entkopplungen verglichen. Es wird je nach Entkopplungsart angenommen, dass entweder N_1 oder S_1 konstant ist.

Wenn man (3.63) mit $(PP_K^{N_1}G_{2,K}^{-1})_{li}$, $(TQP_K^{N_1}G_{2,K}^{-1})_{li}$ und $\left[(UQ + PQ_{N_1}^K)G_{2,K}^{-1}\right]_{li}$ multipliziert, folgt

$$(PP_K^{N_1})_{li}\frac{1}{h}\sum_{j=1}^{s}\hat{\alpha}_{ij}P(X_{lj} - x_{l-1}) \tag{3.64}$$

$$\begin{aligned}+(PP_K^{N_1}G_{2,K}^{-1}BPP_K^{N_1}X)_{li} + (PP_K^{N_1}G_{2,K}^{-1})_{li}\hat{h}(X_{li},t_{li}) &= (PP_K^{N_1}G_{2,K}^{-1}r_*)_{li},\\ -(QQ_{N_1}^K)_{li}\frac{1}{h}\sum_{j=1}^{s}\hat{\alpha}_{ij}P(X_{lj} - x_{l-1}) &\qquad (3.65)\\ +(TQP_K^{N_1}G_{2,K}^{-1}BPP_K^{N_1}X)_{li} + (TQX)_{li}&\\ +(TQP_K^{N_1}G_{2,K}^{-1})_{li}\hat{h}(X_{li},t_{li}) &= (TQP_K^{N_1}G_{2,K}^{-1}r_*)_{li},\\ \left[(UQ + PQ_{N_1}^K)G_{2,K}^{-1}\right]_{li}(BPP_K^{N_1}X)_{li} + U_iQX_{li} &\qquad (3.66)\\ +PQ_{N_1}^KX_{li} + \left[(UQ + PQ_{N_1}^K)G_{2,K}^{-1}\right]_{li}\hat{h}(X_{li},t_{li}) &= \left[(UQ + PQ_{N_1}^K)G_{2,K}^{-1}r_*\right]_{li}.\end{aligned}$$

Lokale Entkopplung mit einem Projektor auf N_1

In diesem Fall setzt man voraus, dass der Raum $N_1(x_*(t),t)$ konstant ist. Nun wird ein konstanter Projektor $Q_{N_1}^K$ auf diesen Raum längs einem entkoppelnden K gewählt.

Zunächst projiziert man (3.61) in die Komponenten $PP_K^{N_1}$, $PQ_{N_1}^K$ und Q

$$(PP_K^{N_1}x)_l = \rho(PP_K^{N_1}x)_{l-1} + \sum_{i=1}^{s}\sum_{j=1}^{s}\beta_i\hat{\alpha}_{ij}(PP_K^{N_1}X)_{lj}, \tag{3.67}$$

$$(Qx)_l = \rho(Qx)_{l-1} + \sum_{i=1}^{s}\sum_{j=1}^{s}\beta_i\hat{\alpha}_{ij}(QX)_{lj}, \tag{3.68}$$

$$(PQ_{N_1}^Kx)_l = \rho(PQ_{N_1}^Kx)_{l-1} + \sum_{i=1}^{s}\sum_{j=1}^{s}\beta_i\hat{\alpha}_{ij}(PQ_{N_1}^KX)_{lj}. \tag{3.69}$$

Die Gleichung (3.67) entspricht der Anwendung des IRK-Verfahrens auf die inhärente Differentialgleichung (3.29) für die Variable $u = PP_K^{N_1}x$. Jedoch ist dies bei den restlichen Gleichungen nicht der Fall. In den Komponenten Qx und $PQ_{N_1}^K x$ findet im Allgemeinen eine Rekursion statt, obwohl sie nicht dynamisch sind. Bemerkenswert ist, dass keine Rekursionen in den nicht-dynamischen Komponenten auftreten, wenn es sich um ein IRK(DAE)-Verfahren handelt. In der Tat ist in diesem Falle $\rho = 0$ und

$$\sum_{i=1}^{s} \beta_i \hat{\alpha}_{ij} = \begin{cases} 1 & j = s \\ 0 & j \neq s \end{cases} .$$

Auf diese Weise erhält man für die Q und $PQ_{N_1}^K$-Komponenten den richtigen Ausdruck, nämlich

$$\begin{aligned} (Qx)_l &= (QX)_{ls}, \\ (PQ_{N_1}^K x)_l &= (PQ_{N_1}^K X)_{ls}. \end{aligned}$$

Jetzt soll sich dem System (3.65)-(3.67) zugewendet werden. Für die Variablen $U_{li} = (PP_K^{N_1}X)_{li}$, $Y_{li} = (UQ + PQ_{N_1}^K)_{li}X_{li}$ und $W_{li} = T_{li}QX_{li}$ schreibt es sich als

$$\begin{aligned} &\frac{1}{h}\sum_{j=1}^{s} \hat{\alpha}_{ij}(U_{lj} - u_{l-1}) + (PP_K^{S_1}G_{2,K}^{-1}BU)_{li} && (3.70) \\ &\quad +(PP_K^{S_1}G_{2,K}^{-1})_i \hat{h}(U_{li} + Y_{li} + W_{li}, t_{li}) &&= (PP_K^{S_1}G_{2,K}^{-1}r_*)_{li}, \\ &-QQ_{N_1}^K \frac{1}{h}\sum_{j=1}^{s} \hat{\alpha}_{ij}((PY)_{lj} - (Py)_{l-1}) + W_{li} && (3.71) \\ &\quad +(TQP_K^{N_1}G_{2,K}^{-1}BU)_{li} + (TQP_K^{N_1}G_{2,K}^{-1})_{li} \\ &\qquad \times \hat{h}(U_{li} + Y_{li} + W_{li}, t_{li}) &&= (TQP_K^{N_1}G_{2,K}^{-1}r_*)_{li}, \\ &\quad \left[(UQ + PQ_{N_1}^K)G_{2,K}^{-1}\right]_{li}(BU)_{li} + Y_{li} && (3.72) \\ &+ \left[(UQ + PQ_{N_1}^K)G_{2,K}^{-1}\right]_{li} \hat{h}(U_{li} + Y_{li} + W_{li}, t_{li}) &&= \left[(UQ + PQ_{N_1}^K)G_{2,K}^{-1}r_*\right]_{li}. \end{aligned}$$

Die Gleichung (3.73) ist nichts anderes als (3.25) an der Stelle t_{li} und (3.73) entspricht

$$F(U_{li}, Y_{li}, W_{li}, t_{li}) = 0.$$

Nach der Analyse von 3.1.2 existiert die implizit definierte Funktion

$$f : U_\rho(x_*) \to I\!R^m,$$

mit

$$\begin{aligned} Y_i &= f(U_{li}, W_{li}, t_{li}), \\ PY_{li} &= Pf(U_{li}, 0, t_{li}), \end{aligned}$$

für $(U_{li}, W_{li}, t_{li}) \in U_\rho(x_*)$.

Nun setzt man $Y_{li} = f(U_{li}, W_{li}, t_{li})$ und

$$\widehat{Y}^h_{p,li} := \frac{1}{h} \sum_{j=1}^{s} \hat{\alpha}_{ij}((PY)_{lj} - (Py)_{l-1})$$

in (3.72) ein,

$$\begin{aligned} &\frac{1}{h} \sum_{j=1}^{s} \hat{\alpha}_{ij}(U_{lj} - u_{l-1}) + (PP^{N_1}_K G^{-1}_{2,K} BU)_{li} \qquad (3.73) \\ +(PP^{N_1}_K G^{-1}_{2,K})_{li} \hat{h}(U_{li} + f(U_{li}, W_{li}, t_{li}) + W_{li}, t_{li}) &= (PP^{N_1}_K G^{-1}_{2,K} r_*)_{li}, \\ -QQ^K_{N_1} \widehat{Y}^h_{p,li} + (TQP^{N_1}_K G^{-1}_{2,K} BU)_{li} + W_{li} \qquad &(3.74) \\ +(TQP^{N_1}_K G^{-1}_{2,K})_{li} \hat{h}(U_{li} + f(U_{li}, W_{li}, t_{li}) + W_{li}, t_{li}) &= (TQP^{N_1}_K G^{-1}_{2,K} r_*)_{li}. \end{aligned}$$

Die Gleichung (3.74) entspricht

$$G(U_{li}, \widehat{Y}^h_{p,li}, W_{li}, t_{li}) = 0,$$

wenn der Raum N_1 konstant ist. Wie in 3.1.2 definiert diese Gleichung die Funktion

$$g \; : \; U_{\rho_2}(x_*) \; \rightarrow \; I\!R^m$$

implizit, wobei

$$W_{li} = g(U_{li}, \widehat{Y}^h_{p,li}, t_{li}).$$

Zuletzt kann W_{li} in (3.73) eingesetzt werden und man erhält

$$\begin{aligned} &\frac{1}{h} \sum_{j=1}^{s} \hat{\alpha}_{ij}(U_{lj} - u_{l-1}) + (PP^{N_1}_K G^{-1}_{2,K} BU)_{li} + (PP^{N_1}_K G^{-1}_{2,K})_{li} \qquad (3.75) \\ \times \hat{h}(U_{li} + f(U_{li}, g(U_{li}, \widehat{Y}^h_{p,li}, t_{li}), t_{li}) + g(U_{li}, \widehat{Y}^h_{p,li}, t_{li}), t_{li}) &= (PP^{N_1}_K G^{-1}_{2,K} r_*)_{li}. \end{aligned}$$

Nun kann man aus (3.75) U_{li} errechnen und für die anderen Komponenten der RK-Stufen gilt:

$$PY_{li} = Pf(U_{li}, 0, t_{li}), \qquad (3.76)$$

$$\widehat{Y}^h_{p,li} : = \frac{1}{h}\sum_{j=1}^{s}\hat{\alpha}_{ij}((PY)_{lj} - (Py)_{l-1}), \tag{3.77}$$

$$W_{li} = g(U_{li}, \widehat{Y}^h_{p,li}, t_{li}), \tag{3.78}$$

$$Y_{li} = f(U_{li}, W_{li}, t_{li}). \tag{3.79}$$

Die Gleichung (3.75) zusammen mit (3.67) ist die Anwendung des IRK-Verfahrens auf die inhärente Differentialgleichung (3.29), falls die Bedingung $Q^{K\prime}_{N_1} = 0$ erfüllt ist. Weiterhin stellen, wenn es sich um ein IRK(DAE) handelt, die Gleichungen (3.67)-(3.69), (3.75)-(3.79) die IRK-Anwendung auf (3.29)-(3.32) dar. In anderen Worten gilt Kommutativität zwischen der Entkopplung mit Projektor auf N_1 und dem IRK(DAE)-Verfahren. Erneut ist mit der Anwendung des IRK-Verfahrens auf (3.29)-(3.32) zu verstehen, dass die Ableitung einer Funktion immer durch die IRK-Ableitungsapproximation zu ersetzen ist.

Falls $N_1(x_*(t), t)$ nicht konstant ist, dann gilt die Kommutativität mit dieser Entkopplungsart nicht mehr. In diesem Fall kommt der problematische Term

$$(PP^{N_1}_K)'(u + y_p(u, \cdot))$$

in (3.29) vor und die Kommutativität ist verletzt.

Lokale Entkopplung mit einem Projektor längs S_1

In diesem Abschnitt wird vorausgesetzt, dass der Raum $S_1(x_*(t), t)$ konstant ist und man wählt einen konstanten Projektor $Q^{S_1}_*$ längs S_1.

Zuerst projiziert man (3.61) in die Komponenten $PP^*_{S_1}$, $PQ^{S_1}_*$ und Q

$$(PP^*_{S_1}x)_l = \rho(PP^*_{S_1}x)_{l-1} + \sum_{i=1}^{s}\sum_{j=1}^{s}\beta_i\hat{\alpha}_{ij}(PP^*_{S_1}X)_{lj}, \tag{3.80}$$

$$(Qx)_l = \rho(Qx)_{l-1} + \sum_{i=1}^{s}\sum_{j=1}^{s}\beta_i\hat{\alpha}_{ij}(QX)_{lj}, \tag{3.81}$$

$$(PQ^{S_1}_*x)_l = \rho(PQ^{S_1}_*x)_{l-1} + \sum_{i=1}^{s}\sum_{j=1}^{s}\beta_i\hat{\alpha}_{ij}(PQ^{S_1}_*X)_{lj}. \tag{3.82}$$

Analog zu der Entkopplung mit einem Projektor auf N_1 entspricht die Gleichung (3.80) der Anwendung des IRK-Verfahrens auf die inhärente Differentialgleichung (3.29) für die Variable $z = PP^*_{S_1}x$. Aber bei den restlichen

Gleichungen ist dies nicht der Fall. In den Komponenten Qx und $PQ_*^{S_1}x$ findet im Allgemeinen eine Rekursion statt, obwohl sie nicht dynamisch sind. Wieder stellen sich die IRK(DAE)-Verfahren als vorteilhaft heraus. In diesem Fall verschwinden die Rekursionen in diesen Komponenten. Da $\rho = 0$ ist und

$$\sum_{i=1}^{s} \beta_i \hat{\alpha}_{ij} = \begin{cases} 1 & j = s \\ 0 & j \neq s \end{cases},$$

erhält man für die Q und $PQ_*^{S_1}$-Komponenten den richtigen Ausdruck, nämlich

$$\begin{aligned} (Qx)_l &= (QX)_{ls}, \\ (PQ_*^{S_1}x)_l &= (PQ_*^{S_1}X)_{ls}. \end{aligned}$$

Bei der Entkopplung mit einem Projektor längs S_1 wird die Notation

$$\begin{aligned} Z_{li} &:= PP_{S_1}^* X_{li}, \\ Y_{li} &:= (UQ + PQ_{N_1}^{S_1})_{li} X_{li}, \\ W_{li} &:= (TQX)_{li} \end{aligned}$$

eingeführt. Das System (3.65)-(3.67) kann folgendermaßen geschrieben werden:

$$\begin{aligned} &\frac{1}{h}\sum_{j=1}^{s} \hat{\alpha}_{ij}(Z_{lj} - z_{l-1}) + (PP_{S_1}^{N_1})_{li}\frac{1}{h}\sum_{j=1}^{s}\hat{\alpha}_{ij}((PQ_*^{S_1}Y)_{lj} \\ &-(PQ_*^{S_1}y)_{l-1}) + (PP_{S_1}^{N_1}G_2^{-1}BPP_{S_1}^{N_1})_{li}(Z + Q_*^{S_1}Y)_{li} \\ &+(PP_{S_1}^{N_1}G_2^{-1})_{li}\hat{h}(Z_{li} + (UQ + PQ_*^{S_1})_{li}Y_{li} + W_{li}, t_{li}) = (PP_{S_1}^{N_1}G_2^{-1}r_*)_{li}, \end{aligned} \tag{3.83}$$

$$\begin{aligned} &-(QQ_{N_1}^{S_1})_{li}\frac{1}{h}\sum_{j=1}^{s}\hat{\alpha}_{ij}((PQ_*^{S_1}Y)_{lj} - (PQ_*^{S_1}y)_{l-1}) + W_{li} \\ &+(TQP_{S_1}^{N_1}G_2^{-1}BPP_{S_1}^{N_1})_{li}(Z + Q_*^{S_1}Y)_{li} + (TQP_{S_1}^{N_1}G_2^{-1})_{li} \\ &\times\hat{h}(Z_{li} + (UQ + PQ_*^{S_1})_{li}Y_{li} + W_{li}, t_{li}) = (TQP_{S_1}^{N_1}G_2^{-1}r_*)_{li}, \end{aligned} \tag{3.84}$$

$$\begin{aligned} &(I + UQG_2^{-1}BPP_{S_1}^{N_1}Q_*^{S_1})_{li}Y_{li} + (UQG_2^{-1}BPP_{S_1}^{N_1})_{li}Z_{li} \\ &+(UQ + PQ_{N_1}^{S_1})_{li}G_{2,i}^{-1} \\ &\times\hat{h}(Z_{li} + (UQ + PQ_*^{S_1})_{li}Y_{li} + W_{li}, t_{li}) = (UQ + PQ_{N_1}^{S_1})_i(G_2^{-1}r_*)_{li}. \end{aligned} \tag{3.85}$$

Die dritte Gleichung ist nichts anderes als

$$F(Z_{li}, Y_{li}, W_{li}, t_{li}) = 0.$$

Dann kann diese Gleichung nach Y_{li} gelöst werden,

$$\begin{aligned} Y_{li} &= f(Z_{li}, W_{li}, t_{li}), \\ Y_{v,li} &= PQ_*^{S_1} Y_{li} = PQ_*^{S_1} f(Z_{li}, 0, t_{li}), \end{aligned}$$

für $(Z_{li}, W_{li}, t_{li}) \in U_p(x_*)$.

Nun setzt man Y_{li} und

$$\widehat{Y}^h_{v,li} := \frac{1}{h} \sum_{j=1}^{s} \hat{\alpha}_{ij} ((PQ_*^{S_1} Y)_{lj} - (PQ_*^{S_1} y)_{l-1})$$

in (3.83), (3.84) ein und erhält das folgende System:

$$\frac{1}{h} \sum_{j=1}^{s} \hat{\alpha}_{ij} (Z_{lj} - z_{l-1}) + (PP_{S_1}^{N_1} G_2^{-1} BPP_{S_1}^{N_1})_{li} \tag{3.86}$$

$$\begin{aligned} &\times (Z_{li} + Q_*^{S_1} f(Z_{li}, 0, t_{li})) + (PP_{S_1}^{N_1})_{li} \widehat{Y}^h_{v,li} + (PP_{S_1}^{N_1} G_2^{-1})_{li} \\ &\times \hat{h}(Z_{li} + (UQ + PQ_*^{S_1})_{li} f(Z_{li}, W_{li}, t_{li}) + W_{li}, t_{li}) \;=\; (PP_{S_1}^{N_1} G_2^{-1} r_*)_{li}, \\ &-(QQ_{N_1}^{S_1})_{li} \widehat{Y}^h_{v,li} + W_{li} + (TQP_{S_1}^{N_1} G_2^{-1} BPP_{S_1}^{N_1})_{li} \qquad (3.87) \\ &\times (Z_{li} + Q_*^{S_1} f(Z_{li}, 0, t_{li})) + (TQP_{S_1}^{N_1} G_2^{-1})_{li} \\ &\times \hat{h}(Z_{li} + (UQ + PQ_*^{S_1})_{li} f(Z_{li}, W_{li}, t_{li}) + W_{li}, t_{li}) \;=\; (TQP_{S_1}^{N_1} G_2^{-1} r_*)_{li}. \end{aligned}$$

Der Prozedur nach ist die Gleichung (3.88) genau

$$G(Z_{li}, W_{li}, \widehat{Y}^h_{v,li}, t_{li}) = 0.$$

Demzufolge ist W_{li} durch

$$W_{li} = g(Z_{li}, \widehat{Y}^h_{v,li}, t_{li})$$

in der Umgebung $U_{\rho_2}(x_*)$ gegeben.

Beim Einsetzen von W_{li} in (3.59) bekommt man schließlich

$$\begin{aligned} &\frac{1}{h} \sum_{j=1}^{s} \hat{\alpha}_{ij} (Z_{lj} - z_{l-1}) + (PP_{S_1}^{N_1})_{li} \widehat{Y}^h_{v,li} + (PP_{S_1}^{N_1} G_2^{-1} BPP_{S_1}^{N_1})_{li} \\ &\qquad \times (Z_{li} + Q_*^{S_1} f(Z_{li}, 0, t_{li})) + (PP_{S_1}^{N_1} G_2^{-1})_{li} \hat{h}(Z_{li} \qquad (3.88) \\ &+ (UQ + PQ_*^{S_1})_{li} f(Z_{li}, g(Z_{li}, \widehat{Y}^h_{v,li}, t_{li}), t_{li}) + g(Z_{li}, \widehat{Y}^h_{v,li}, t_{li}), t_{li}) \;=\; (PP_{S_1}^{N_1} G_2^{-1} r_*)_{li}. \end{aligned}$$

Die Gleichung (3.88) zusammen mit (3.80) stellt das entsprechende IRK-Verfahren dar, wenn wie angenommen der Raum S_1 konstant ist. Nunmehr

bestehen für ein IRK(DAE)-Verfahren, wie schon erwägt, keine Rekursionen in den restlichen Komponenten und es gilt die Kommutativität aus dem Diagramm 3.1 zwischen der Entkopplung mit einem Projektor längs S_1 in der IRK(DAE)-Diskretisierung. Wie im letzten Abschnitt versteht man darunter, dass alle Ableitungen nach t durch die RK-Approximation

$$v'(t_{li}) \approx \frac{1}{h} \sum_{j=1}^{s} \hat{\alpha}_{ij}(v_{lj} - v_{l-1})$$

ersetzt werden.

Wenn der Raum $S_1(x_*(t), t)$ nicht konstant ist, treten die Terme

$$(PP^*_{S_1})', \quad PP^{N_1}_{S_1} Q^{S_{1'}}_*,$$

in (3.39) auf. Infolgedessen ist die Kommutativiät bei dieser Entkopplungsart veletzt.

3.2.3 Ein numerisches Beispiel

Die Ergebnisse dieser Sektion verbessern, wie schon erwähnt, jene aus [Hanke et al., 1998]. Die in [Hanke et al., 1998] gefundene Bedingung, damit das Diagramm 3.1 gilt, lautet: $Q^{S_{1'}}_{N_1} = 0$. Das heißt in anderen Worten, dass sowohl $N_1(x_*(t), t)$ als auch $S_1(x_*(t), t)$ konstant sein müssen. Es wurde gezeigt, dass es genügt, wenn einer dieser Räume diese Eigenschaft aufweist.

Bei dem Beispiel 3.0.8 ist die Bedingung $(Q^{S_1}_{N_1})' = 0$ verletzt, da beide Räume $N_1(t)$ und $S_1(t)$ nicht konstant sind. Das folgende Beispiel ist insofern interessant, als der Raum S_1 konstant, aber die Bedingung $Q^{S_{1'}}_{N_1} = 0$ verletzt ist.

Beispiel 3.2.1

$$\begin{pmatrix} x_1' \\ x_2' \\ 0 \end{pmatrix} + \begin{pmatrix} \lambda & (\eta t - 1)^2 & -(\eta t - 1) \\ \eta t(\eta t - 1) & \lambda & -\eta t \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 0,\ t \geq 0.$$

Es handelt sich erneut um ein Index-2-Hessenberg-System. Die entscheidende Matrix $B_{21}B_{12}$ ist hier

$$B_{21}B_{12} = \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} -(\eta t - 1) \\ -\eta t \end{pmatrix} = 1,$$

also ist die Index-2-Bedingung für alle t garantiert. Die allgemeine Lösung ist durch

$$\begin{aligned} x_1(t) &= x_1(0)e^{-\lambda t}, \\ x_2(t) &= x_1(t), \\ x_3(t) &= (\eta t - 1)x_1(t), \end{aligned}$$

gegeben.

Die Lösungsmannigfaltigkeit ist in diesem Fall

$$M_1(t) = \left\{ z \in I\!R^3 : z = \begin{pmatrix} z_1 \\ z_1 \\ (\eta t - 1)z_1 \end{pmatrix} \right\},$$

und der kritische Raum $S_1(t)$ ist durch

$$S_1(t) = \left\{ z \in I\!R^3 : z_1 = z_2 \right\}$$

gegeben.

Andererseits ergibt, wenn man den Orthoprojektor auf N wählt, die Berechnung von $G_1(t)$

$$G_1(t) = \begin{pmatrix} 1 & 0 & 1 - \eta t \\ 0 & 1 & -\eta t \\ 0 & 0 & 0 \end{pmatrix},$$

woraus folgt, dass

$$N_1(t) = \left\{ z \in I\!R^3 : z_1 = (\eta t - 1)z_3 \wedge z_2 = \eta t z_3 \right\}.$$

Weiterhin ist die exakte Lösung dieser Gleichung offensichtlich exponentiell asymptotisch stabil im Lyapunovschen Sinne, wenn $\lambda > 0$ ist. Nunmehr liegt bei diesem Problem starke P-Kontraktivität auf der Lösungsmannigfaltigkeit vor. Man betrachtet das euklidische Skalarprodukt und einen Vektor (y, x, t), der zu $M_1(t)$gehört und die Gleichungen

$$\begin{aligned} y_1 &= -\lambda x_1 - (\eta t - 1)^2 x_2 + (\eta t - 1)x_3, \\ y_2 &= -\eta t(\eta t - 1)x_1 - \lambda x_2 + \eta t x_3, \\ 0 &= x_1 - x_2, \\ y_3 &= 0 \end{aligned}$$

erfüllt. Die linke Seite der Kontraktivitätsbedingung ist dann

$$\langle y, x\rangle_2 = -\lambda x_1^2 - (\eta t-1)^2 x_1 x_2 + (\eta t-1) x_1 x_3 - \eta t(\eta t-1) x_1 x_2 - \lambda x_2^2 + \eta t x_2 x_3,$$

und, wenn man die Bedingung $x \in M_1(t)$ nutzt, folgt

$$\langle y, x\rangle_2 = -\lambda(x_1^2 + x_2^2) = -\lambda \|Px\|_2^2 .$$

Die Anwendung des impliziten Euler-Verfahrens auf dieses Problem stellt in jedem Schritt die Lösung des linearen Systems

$$\begin{pmatrix} 1+h\lambda & h(\eta t_i - 1)^2 & -h(\eta t_i - 1) \\ h\eta t_{i+1}(\eta t_{i+1} - 1) & 1+h\lambda & -h\eta t_{i+1} \\ 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} x_{1,i+1} \\ x_{2,i+1} \\ x_{3,i+1} \end{pmatrix} = \begin{pmatrix} x_{1,i} \\ x_{2,i} \\ 0 \end{pmatrix}$$

dar. Aus diesem System erhält man für die erste Komponente beispielsweise

$$x_{1,i+1} = \frac{1}{1+h\lambda} x_{1,i},$$

was für alle $\lambda > 0$ die Kontraktionsbedingung

$$\left|\frac{1}{1+h\lambda}\right| \leq 1$$

erfüllt.

Die numerischen Ergebnisse mit diesem Verfahren für $h = 0.1$ und verschiedene Werte von λ und η sind in der Abbildung 3.3 dargestellt. Als Anfangswert wurde $x_1(0) = x_2(0) = 1$ und $x_3(0) = -1$ gesetzt.

3.3 Index-Reduktion durch Differentiation

In diesem Kapitel wurde bis jetzt eine quasilineare ADGl der Form

$$A(t)(Px)' + b(x(t), t) = 0, \quad t \in \Im$$

betrachtet. Leider sind einige Anwendungsfälle darin nicht eingeschlossen, zum Beispiel die Modelle aus der klassischen MNA (modified node analysis) in der Schaltkreissimulation, [Estévez Schwarz and Tischendorf, 1998], [Estévez Schwarz, 2000]. Hier handelt es sich um Gleichungen wie (2.2), also

$$A(x(t), t)(Px)'(t) + b(x(t), t) = 0, \quad t \in \Im. \tag{3.89}$$

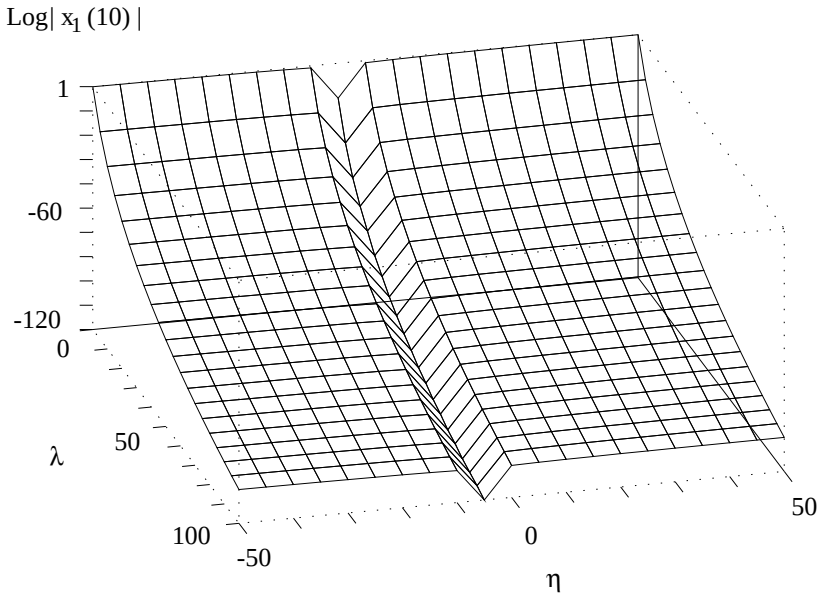

Abbildung 3.3: Approximierte Lösung für das Beispiel 3.2.1 unter Verwendung des impliziten Euler-Verfahrens bei einem Diskretisierungsschritt $h = 0.1$ und verschiedenen Werten von λ und η.

Auf die strukturellen Einzelheiten der MNA-Modelle wird am Ende des Kapitels eingegangen. Den Voraussetzungen nach sind $A(x,t)$, $b(x,t)$, $A_x(x,t)$ und $b_x(x,t)$ stetige Funktionen in D_f. Für (3.89) wird angenommen, dass der Tractability-Index-2 in einer offenen Menge $G \subset D_f$ vorliegt.

Der Index-Reduktionsprozess beruht auf dem am weitesten verbreiteten Index-Begriff, dem Differen-tiations-Index, [Gear, 1988], [Brenan et al., 1989]. Nach diesem Kriterium kann man durch sukzessive Differentiation der Gleichungen (oder einens Teils davon) den Index einer ADGl reduzieren. Es ist nicht nötig dafür zu verlangen, dass das ganze System nach der unabhängigen Variablen differenzierbar sein muss. Die Index-Reduktion, die hier vorgestellt wird, setzt nur die Differenzierbarkeit voraus, die die Lösbarkeit des Problems verlangt. Hier wird den Ideen von [März, 1998] gefolgt, für die entsprechende Index-Reduktion für lineare Systeme siehe auch [März and Rodríguez Santiesteban, 1999]. Diese Vorgehensweise erinnert auch an die Reduktionstechniken für mechanische Systeme,

[Eich-Soellner and Führer, 1998].

Die Gleichung (3.89), siehe Kapitel 2, kann auch als

$$A(x(t),t)(Px)'(t) + (I - W_1(x(t),t))b(x(t),t) + (W_1 b)(x(t),t) = 0$$

geschrieben werden. Die ersten zwei Terme in dieser Gleichung befinden sich im $\operatorname{im}(I - W_1)$, während der dritte im $\operatorname{im} W_1$ lebt. Die Index-Reduktion besteht darin, die Gleichung $(W_1 b)(x(t),t) = 0$, die die sichtbare Nebenbedingung bestimmt, durch die versteckte

$$W_1(x(t),t)\left\{(W_1 b)_x(x(t),t)(Px)'(t) + (W_1 b)_t(x(t),t)\right\} = 0$$

zu ersetzen. So erhält man eine neue Gleichung

$$(A + W_1(W_1 b)_x)(Px)' + (I - W_1)b + W_1(W_1 b)_t = 0. \tag{3.90}$$

An dieser Stelle wurden die Argumente weggelassen.

Es wird die Notation

$$\begin{aligned} \widetilde{A}(x,t) &:= A(x,t) + (W_1(W_1 b)_x)(x,t), \\ \widetilde{b}(x,t) &:= (I - W_1(x,t))b(x,t) + (W_1(W_1 b)_t)(x,t) \end{aligned}$$

eingeführt und (3.90) schreibt sich jetzt als

$$\widetilde{f}((Px)'(t), x(t), t) := \widetilde{A}(x(t),t)(Px)'(t) + \widetilde{b}(x(t),t) = 0. \tag{3.91}$$

Allerdings enthält diese Gleichung nicht die ganze Information, die in (2.2) steckt. Ausschließlich mit der zusätzlichen Nebenbedingung (2.15) sind beide Systeme äquivalent.

Die neue Aufgabe (3.91) besitzt einige gute Eigenschaften, deren erste in folgendem Lemma dargestellt wird.

Lemma 3.3.1 *Für den Nullraum von $\widetilde{A}(x,t)$ gilt*

$$\ker \widetilde{A}(x,t) = \ker A(x,t) = N.$$

Beweis: S ei $z \in \ker \widetilde{A}(x,t)$, also

$$(A + W_1(W_1 b)_x)(x,t)z = 0.$$

Diese Gleichung bedeutet

$$\begin{aligned} A(x,t)z &= 0, \\ (W_1(W_1b)_x)(x,t)z &= 0, \end{aligned}$$

das zeigt die Inklusion $\ker \widetilde{A}(x,t) \subset N$. Die andere Richtung folgt dann sofort durch Lemma 2.1.15.
Man kann die neue ADGl (3.91) genauso wie (2.2) untersuchen. Dafür ist Lemma 3.3.1 sehr wichtig, denn nach ihm sind die Lösungen von (3.91) auch in dem Funktionsraum $C_N^1(\Im, I\!R^m)$ enthalten.

Sei $\widetilde{M}_0(t)$ die sichtbare Lösungsmannigfaltigkeit von (3.91), die nach Definition durch

$$\widetilde{M}_0(t) = \left\{ x \in D : \widetilde{b}(x,t) \in \operatorname{im} \widetilde{A}(x,t) \right\}$$

gegeben ist.

Lemma 3.3.2 *$M_1(t) \subset \widetilde{M}_0(t)$ für alle $t \in \Im$.*

Beweis: S ei $x_0 \in M_1(t)$, dann gibt es $z_0 \in I\!R^m$, so dass $b(x_0,t) = A(x_0,t)z_0$. Das bedeutet

$$(W_1b)(x_0,t) = 0 \Rightarrow (I - W_1(x_0,t))b(x_0,t) = A(x_0,t)z_0,$$

und außerdem

$$Pz_0 = P(A^+b)(x_0,t).$$

Auf der anderen Seite gilt für $x_0 \in M_1(t)$ auch

$$\begin{aligned} (W_1(W_1b)_t)(x_0,t) &= (W_1(W_1b)_x)(x_0,t)P(A^+b)(x_0,t) \\ &= (W_1(W_1b)_x)(x_0,t)z_0. \end{aligned}$$

Nun erhält man

$$\begin{aligned} \widetilde{b}(x_0,t) &= (I - W_1(x_0,t))b(x_0,t) + (W_1(W_1b)_t)(x_0,t) \\ &= A(x_0,t)z_0 + (W_1(W_1b)_x)(x_0,t)z_0 \\ &= \widetilde{A}(x_0,t)z_0, \end{aligned}$$

was die Zugehörigkeit von x_0 zu $\widetilde{M}_0(t)$ impliziert.

Bemerkung 3.3.3 *Aus dem Beweis ist unschwer zu ersehen, dass*

$$M_1(t) = \widetilde{M}_0(t) \cap \left\{ x \in I\!R^m : (W_1b)(x,t) = 0 \right\}.$$

Lemma 3.3.4 *Die Mannigfaltigkeit $M_1(t)$ ist eine Invariante für (3.91).*

Beweis: S ei $x(t)$ eine Lösung von (3.91), für die $x(t_0) \in M_1(t_0)$, $t_0 \in \Im$ gilt. Sei außerdem $s(t) := (W_1 b)(x(t), t)$, dann hat man $s(t_0) = 0$. Es wird s nach t abgeleitet, und es folgt

$$s'(t) = (W_1 b)_x(x(t), t)(Px)'(t) + (W_1 b)_t(x(t), t).$$

Infolgedessen gilt

$$0 = W_1(x(t), t)s'(t) = (W_1(x(t), t)s(t))_t - W_{1t}(x(t), t)s(t) = s'(t) - W_{1t}(x(t), t)s(t),$$

woraus, zusammen mit $s(t_0) = 0$, $x(t) \in M_1(t)$, $\forall t \in \Im$ folgt.
Für die Hauptaussage dieses Abschnittes ist auch folgendes Lemma nützlich:

Lemma 3.3.5 *$G_2 Q_1 G_2^{-1}$ ist wie W_1 auch ein Projektor längs* $\operatorname{im} G_1$*. Folglich gilt*

$$W_1 = W_1 G_2 Q_1 G_2^{-1} \quad \textit{und} \quad Q_1 G_2^{-1} = Q_1 G_2^{-1} W_1.$$

Beweis: O ffensichtlich ist

$$\left(G_2 Q_1 G_2^{-1}\right)^2 = G_2 Q_1 G_2^{-1}.$$

Außerdem, da

$$Q_1 G_2^{-1} G_1 = Q_1 G_2^{-1} G_2 P_1 = 0,$$

gilt

$$\ker\left(Q_1 G_2^{-1}\right) \supset \operatorname{im} G_1.$$

Andererseits ist

$$\dim\left(\ker\left(Q_1 G_2^{-1}\right)\right) = \dim\left(\operatorname{im} G_1\right),$$

woraus

$$\ker\left(Q_1 G_2^{-1}\right) = \operatorname{im} G_1$$

folgt.
An dieser Stelle ist man in der Lage, die Hauptaussage dieses Abschnittes, nämlich dass die Gleichung (3.91) den Tractability-Index-1 besitzt, zu beweisen.

Theorem 3.3.6 *Seien die Glattheitsvoraussetzungen erfüllt, zusätzlich sei* $(W_1 b)(x,t)$ *zweimal nach beiden Argumenten stetig differenzierbar. Außerdem seien* $t_0 \in \Im$, $x_0 \in M_1(t_0)$, $x_0' := -(A^+ b)(x_0, t_0)$, *so dass* $A(x_0,t_0)x_0' + b(x_0,t_0) = 0$ *und*

$$N_1(x_0', x_0, t_0) \cap S_1(x_0', x_0, t_0) = \{0\}$$

gelten. Dann sind die folgenden Aussagen richtig:

- *Die Gleichung (3.91) ist in* (x_0', x_0, t_0) *Index-1-tractable*
- *Eine Anfangswertaufgabe für (2.2), mit* $x(t_0) = x_0 \in M_1(t_0)$, *besitzt lokal eine eindeutige Lösung* $x \in C_N^1$, *für die*

  $$x(t_0) = x_0, \quad P(x'(t_0) - x_0') = 0$$

  gilt.

Beweis: U nter den Annahmen des Theorems sind $\widetilde{A}(x,t)$, $\widetilde{b}(x,t)$, $\widetilde{A}_x(x,t)$ und $\widetilde{b}_x(x,t)$ stetige Funktionen. Nach dem Lemma 3.3.2 gilt auch $x_0 \in M_0(t_0)$. Es wird als Nächstes die erste Aussage bewiesen, wofür man die Matrix $\widetilde{G}_1$ berechnet:

$$\begin{aligned}
\widetilde{G}_1(y,x,t) &= \widetilde{A}(x,t) + \widetilde{B}(y,x,t)\widetilde{Q} \\
&= \widetilde{A}(x,t) + (\widetilde{b}_x(x,t) + (\widetilde{A}(x,t)y)_x)\widetilde{Q} \\
&= A(x,t) + (W_1(W_1b)_x)(x,t) + [b_x(x,t) - (W_1b)_x(x,t) \\
&\quad + (W_1(W_1b)_t)_x(x,t) + (A(x,t)y)_x + ((W_1(W_1b)_x)(x,t)y)_x]\,Q,
\end{aligned}$$

und man betrachtet das lineare Gleichungssystem

$$\widetilde{G}_1(y,x,t)z = 0.$$

Der Term $(W_1b)_x(x,t)Qz$ ist nach dem Lemma 2.1.15 null und für die Terme

$$(W_1(W_1b)_t)_x(x,t)Qz, \quad ((W_1(W_1b)_x)(x,t)y)_xQz,$$

erhält man

$$\begin{aligned}
(W_1(W_1b)_t)_xQz &= W_{1x}Qz(W_1b)_t + W_1(W_1b)_{tx}Qz = W_1(W_1b)_{tx}Qz, \\
(W_1(W_1b)_xy)_xQz &= W_{1x}Qz(W_1b)_xy + W_1((W_1b)_xy)_xQz = W_1(W_1b)_{xx}Qz\,y = 0.
\end{aligned}$$

So folgt $\widetilde{G}_1(y,x,t)z = 0 \Leftrightarrow$

$$A(x,t)z+(W_1(W_1b)_x)(x,t)z+b_x(x,t)Qz+(W_1(W_1b)_{tx})(x,t)Qz+(A(x,t)y)_xQz = 0,$$

$\Leftrightarrow$

$$G_1(y,x,t)z + (W_1(W_1b)_x)(x,t)z + (W_1(W_1b)_{tx})(x,t)Qz = 0,$$

$\Leftrightarrow$

$$\begin{aligned} G_1(y,x,t)z &= 0 \\ (W_1(W_1b)_x)(x,t)z + (W_1(W_1b)_{xt})(x,t)Qz &= 0, \end{aligned}$$

$\Leftrightarrow$

$$\begin{aligned} G_1(y,x,t)z &= 0 \\ (W_1(W_1b)_x)(x,t)z &= 0. \end{aligned}$$

Die erste Gleichung des letzten Systems bedeutet, dass $z \in N_1(y,x,t)$ und aus der zweiten Gleichung erhält man

$$W_1(x,t)W_{1x}(x,t)zb(x,t) + W_1(x,t)b_x(x,t)z = 0.$$

Unabhängig davon gilt für $z = Q_1(y,x,t)z$

$$\begin{aligned} Q_1(y,x,t)z &= (Q_1G_2^{-1}B)(y,x,t)z = (Q_1G_2^{-1})(y,x,t)W_1(x,t)B(y,x,t)z \\ &= (Q_1G_2^{-1})(y,x,t)W_1(x,t)(b_x(x,t) + (A(x,t)y)_x)z \\ &= -(Q_1G_2^{-1})(y,x,t)W_1(x,t)W_{1x}(x,t)z(b(x,t) + A(x,t)y). \end{aligned}$$

In dem Punkt (x_0', x_0, t_0) folgt $z = Q_1(x_0', x_0, t_0)z = 0 = z$. Das beweist die Regularität von $\widetilde{G}_1(x_0', x_0, t_0)$ und damit besitzt die Gleichung (3.91) den Tractability-Index-1 in (x_0', x_0, t_0).

Für die zweite Aussage des Theorems betrachtet man (3.91) mit dem Anfangswert $x(t_0) = x_0 \in M_1(t_0) \subset \widetilde{M}_0(t_0)$, nach dem Theorem 14 Abschnitt 1.2 von [Griepentrog and März, 1986] existiert eine eindeutige C_N^1-Lösung $x(t)$ von dieser AWA in einer Umgebung von x_0, also für $t \in \Im_0 \subset \Im$. Es reicht jetzt aus zu zeigen, dass diese Lösung ebenfalls eine von (3.89) ist. Nach dem Lemma 3.3.4 gilt $x(t) \in M_1(t)$ und damit ist $(W_1b)(x(t),t) = 0, \forall t \in \Im_0$. Außerdem ergibt $I - W_1(Px(t),t)$ multipliziert mit (3.91)

$$A(x(t),t)(Px)'(t) + (I - W_1(x(t),t))b(x(t),t) = 0,$$

was zusammen mit $(W_1 b)(x(t), t) = 0$ die ursprüngliche Gleichung

$$A(x(t), t)(Px)'(t) + b(x(t), t) = 0$$

mit sich bringt. Da jede Lösung von (2.2) auch eine Lösung von (3.91) darstellt, ist die Eindeutigkeit der Lösung gleichermaßen gesichert.

Bemerkung 3.3.7

1. Die Ergebnisse dieses Abschnittes gelten unter Annahmen, die aus praktischer Sicht akzeptabel sind. Die Voraussetzung, dass im $G_1(y, x, t)$ nur von (Px, t) abhängt, ist womöglich diejenige, die man in Zweifel ziehen könnte. Jedoch erfüllen sowohl die Hessenberg-Systeme als auch die Modelle aus der klassischen modifizierten Knoten-Analyse in der Schaltkreissimulation, [Tischendorf, 1996], [Estévez Schwarz and Tischendorf, 1998], [Estévez Schwarz, 2000], diese Bedingungen.

2. Man bemerkt, dass die Differenzierbarkeitsannahmen, hier über $(W_1 b)(x, t)$ und im Theorem 3.1.10 über $S_1(t)$, im linearen Fall im Einklang stehen.

3. Nach den Ergebnissen dieses Abschnittes bietet sich folgender Ansatz an: Um die AWA für die Index-2-Gleichung (3.89) zu lösen, führt man zunächst eine Index-Reduktion durch und diskretisiert anschließend die Index-1-Gleichung, die konstanten Nullraum besitzt. Die Idee beruht auf der Tatsache, dass die Mannigfaltigkeit $M_1(t)$ invariant für (3.89) ist. Jedoch reicht diese Eigenschaft der Gleichung (3.89) nicht aus, da die M_1-Invarianz bei der Diskretisierung nicht vorkommt. Im Allgemeinen wird mit diesem Ansatz die versteckte Nebenbedingung, aber nicht $W_1 b = 0$, durch die Näherungslösung erfüllt, was zu falschen Ergebnissen führt.

3.4 Stabilität der index-reduzierten Gleichung

Der Ansatz dieses Kapitels, die Kommutativität zwischen Diskretisierung und einer Index-Reduktions-Transformation als Werkzeug für die Stabilitätsuntersuchung zu benutzen, wird besonders unterstützt, wenn die Aufgabe im Kasten unten links des Diagramms 3.1 gewisse Stabilitätseigenschaften der

ursprünglichen Aufgabe erbt. Offensichtlich überträgt sich im Fall der Index-Reduktion durch Differentiation die Lyapunovsche Stabilität einer Lösung von (3.89), da alle Lösungen dieser Gleichung auch eine von (3.91) sind, siehe den Beweis von Theorem 3.3.6. Ziel dieses Abschnittes ist es zu zeigen, dass sich auch die Kontraktivität in gewissem Sinne überträgt. Zuerst muss geklärt werden, auf welcher Mannigfaltigkeit diese Eigenschaft betrachtet wird. Beide Aufgaben besitzen dafür verschiedene Obermengen $M_0(t)$ und $\widetilde{M}_0(t)$, jedoch kann als gemeinsamer Nenner $M_1(t)$ erkannt werden, $M_1(t) \subset M_0(t)$ und $M_1(t) \subset \widetilde{M}_0(t)$.

Theorem 3.4.1 *Man nimmt an, dass die Gleichung (3.89) den Tractability-Index-2 besitzt. Sei außerdem diese Gleichung (P-)kontraktiv auf einer Mannigfaltigkeit $\Gamma(t) \subseteq M_1(t)$. Dann ist die index-reduzierte Gleichung (3.91) auch (P-)kontraktiv auf $\Gamma(t)$.*

Beweis: M an betrachtet die Vektoren (y_1, x_1, t) und (y_2, x_2, t) in $\mathbb{R}^m \times \Gamma(t) \times \mathfrak{S}$, die die Bedingungen

$$\begin{aligned} Qy_1 = Qy_2 = 0, \\ \widetilde{A}(x_1,t)y_1 + \widetilde{b}(x_1,t) = \widetilde{A}(x_2,t)y_2 + \widetilde{b}(x_2,t) = 0, \end{aligned} \tag{3.92}$$

erfüllen. Als Projektor auf $\widetilde{N}$ kann Q gewählt werden, weil $\widetilde{N} = N$.

Da x_1, x_2 auch zu $M_0(t)$ gehören, gilt

$$W_1(t)b(x_1,t) = W_1(t)b(x_1,t) = 0.$$

Andererseits geht aus (3.92) hervor, dass

$$A(x_1,t)y_1 + (I - W_1(t))b(x_1,t) = A(x_2,t)y_2 + (I - W_1(t))b(x_2,t) = 0.$$

Aus den zwei letzten Gleichungen folgt auch für (y_1, x_1, t) und (y_2, x_2, t)

$$\begin{aligned} Qy_1 = Qy_2 = 0, \\ A(x_1,t)y_1 + b(x_1,t) = A(x_2,t)y_2 + b(x_2,t) = 0, \end{aligned}$$

und durch die Kontraktivität von (3.89) auf $\Gamma(t)$ erhält man

$$\langle y_1 - y_2, x_1 - x_2 \rangle_S \leq -c \, \|x_1 - x_2\|_S^2 .$$

Bemerkung 3.4.2 *Besteht in irgendeiner Art und Weise (P-)Kontraktivität auf $\widetilde{M}_0(t)$, so ist die Anwendung eines algebraisch stabilen IRK(DAE) auf die index-reduzierte Gleichung nach dem Kollorar 2.3.7 auch (P-)kontraktiv. Aber selbst in diesem Fall besteht noch das Problem, dass die Nebenbedingung $W_1 b = 0$ durch die diskrete Lösung nicht immer erfüllt wird, Bemerkung 3.3.7.*

3.5 Kommutativität zwischen Index-Reduktion und Diskretisierung

Hier wird eine der Hauptfragen dieses Kapitels beantwortet, und zwar unter welchen Bedingungen das Kommutativitätsdiagramm 3.1 für die Index-Reduktion durch Differentiation (IRD) und die BDF- und IRK-Verfahren zutrifft.

3.5.1 BDF-Verfahren

Ein BDF-Verfahren für die ADGl

$$A(x(t),t)(Px)'(t) + b(x(t),t) = 0, \quad t \in [t_0, \infty), \tag{3.93}$$

wie in der Sektion 3.2 schon eingeführt, lautet

$$A_i \sum_{j=0}^{k} \alpha_j (Px)_{i-j} + h b_i = 0, \quad i \geq k, \tag{3.94}$$

wobei $A_i := A(x(t_i), t_i)$ und $b_i := b(x(t_i), t_i)$.

Als Erstes stellt man fest, dass die diskrete Lösung in jedem Schritt zu $M_0(t_i)$ gehört. Es wird daran erinnert, dass dies bei den IRK-Verfahren auch zutraf. Aus (3.94) folgt

$$b(x_i, t_i) \in \operatorname{im} A(x_i, t_i),$$

und damit ist diese Aussage deutlich.

Wie in dem kontinuierlichen Fall wird (3.94) in die Komponenten auf $\operatorname{im} G_{1,i}$ und ein Komplement davon aufgesplittet:

$$A_i \sum_{j=0}^{k} \alpha_j (Px)_{i-j} + h(I - W_1)_i b_i = 0, \quad i \geq k, \tag{3.95}$$

$$(W_1 b)_i = 0. \tag{3.96}$$

Die Gleichung (3.96) ist nichts anderes als eine Folge des schon erwähnten Faktes, dass die diskrete Lösung zu $M_0(t_i)$ gehört.

Das Ziel ist nun herauszufinden, wann (3.95), (3.96) der gleichen Diskretisierung von (3.91) entspricht, also

$$(A + W_1(W_1 b)_x)_i \sum_{j=0}^{k} \alpha_j (Px)_{i-j} + h(I - W_1)_i b_i + h W_{1,i} (W_1 b)_{t,i} = 0. \quad i \geq k. \tag{3.97}$$

Aus dieser Gleichung ist ersichtlich, dass (3.95) erfüllt ist. So folgt zwangsläufig die Bedingung

$$W_{1,i}\left\{(W_1b)_{x,i}\sum_{j=0}^{k}\alpha_j(Px)_{i-j}+h(W_1b)_{t,i}\right\}=0,\quad i\geq k, \tag{3.98}$$

damit (3.97) und die Kommutativität gelten. Weiterhin folgt aus (3.94)

$$\sum_{j=0}^{k}\alpha_j(Px)_{i-j}=P\sum_{j=0}^{k}\alpha_j(Px)_{i-j}=-hP(A^+b)_i,\quad i\geq k$$

und (3.98) verwandelt sich in

$$W_{1,i}\left\{-(W_1b)_xA^+b+(W_1b)_t\right\}_i=0,\quad i\geq k. \tag{3.99}$$

Die letzte Gleichung ist die versteckte Nebenbedingung (2.17) an der Stelle t_i. Als Schlussfolgerung kann man behaupten, dass, wenn die diskrete Lösung x_i zusätzlich (3.99) erfüllt, dann das Diagramm 3.1 gilt.

In der Fachliteratur stellt man fest, dass der Zugehörigkeit der diskreten Lösung zu $M_1(t)$ besondere Aufmerksamkeit gewidmet wird. Die vorliegende Analyse unterstreicht die Bedeutsamkeit dieses Sachverhaltes.

Das System (3.95), (3.96), (3.99) ist im Gegensatz zu dem kontinuierlichen Fall überbestimmt. Deswegen findet man in der Fachliteratur eine Vielfalt von Ansätzen, die das System in irgendeinem verallgemeinerten Sinne lösen. In diesem Abschnitt ist man daran interessiert zu klären, wann dies nicht notwendig ist. Kapitel 4 soll sich mit dem anderen Fall näher auseinander setzen.

Das bishier erzielte Ergebnis kann folgendermaßen zusammengefasst werden.

Kriterium 3.5.1 *Wenn die diskrete Lösung die Bedingung (3.99) in jedem Schritt erfüllt (also wenn $x_i \in M_1(t_i)$), dann kommutieren die IRT und die BDF-Diskretisierung.*

3.5.2 IRK-Verfahren

Die Analyse der IRK-Verfahren verläuft nach dem gleichen Muster wie bei den BDF-Verfahren. Ein IRK-Verfahren für die ADGl

$$A(x(t),t)(Px)'(t)+b(x(t),t)=0,\quad t\in[t_0,\infty), \tag{3.100}$$

wie in 3.2.2 schon eingeführt, lautet

$$x_l = \rho x_{l-1} + \sum_{i=1}^{s}\sum_{j=1}^{s} \beta_i \hat{\alpha}_{ij} X_{lj}, \quad (3.101)$$

$$A_{li} \sum_{j=1}^{s} \hat{\alpha}_{ij}(X_{lj} - x_{l-1}) + h b_{li} = 0,\ i = 1, \ldots, s, \quad (3.102)$$

wobei $A_{li} := A(X_{li}, t_{li})$ und $b_{li} := b(X_{li}, t_{li})$.

Das Entscheidende hierbei ist, ob die Gleichung (3.102) der IRK-Diskretisierung der index-reduzierten Gleichung (3.91) entspricht.

Als Erstes stellt man fest, dass alle Stufen des Verfahrens zu M_0 gehören. Aus (3.102) folgt

$$b(X_{li}, t_{li}) \in \operatorname{im} A(X_{li}, t_{li})$$

und damit ist diese Aussage deutlich.

Es werden die Gleichungen (3.102) in die Komponenten auf $\operatorname{im} G_{1,li}$ und ein Komplement davon aufgesplittet:

$$A_{li} \sum_{j=1}^{s} \hat{\alpha}_{ij}(X_{lj} - x_{l-1}) + h(I - W_1)_{li} b_{li} = 0, \quad i = 1, \ldots, s, \quad (3.103)$$

$$(W_1 b)_{li} = 0, \quad i = 1, \ldots, s. \quad (3.104)$$

Andererseits erhält man, wenn das IRK-Verfahren direkt auf die index-reduzierte Gleichung (3.91) angewandt wird, für die RK-Stufen

$$(A + W_1(W_1 b)_x)_{li} \sum_{j=1}^{s} \hat{\alpha}_{ij}(X_{lj} - x_{l-1}) + h(I - W_1)_{li} b_{li} \quad (3.105)$$
$$+ h W_{1,li}(W_1 b)_{t,li} = 0,\ i = 1, \ldots, s.$$

Aus dieser Gleichung ist ersichtlich, dass die Projektion von (3.105) auf $\operatorname{im} G_{1,i}$ erfüllt ist, (3.104). So folgt zwangsläufig die Bedingung

$$W_{1,li} \left\{ (W_1 b)_{x,li} \sum_{j=1}^{s} \hat{\alpha}_{ij}(X_{lj} - x_{l-1}) + h(W_1 b)_{t,li} \right\} = 0, \quad i = 1, \ldots, s \quad (3.106)$$

damit (3.105) gilt. Weiterhin folgt aus (3.102), dass

$$P \sum_{j=1}^{s} \hat{\alpha}_{ij}(X_{lj} - x_{l-1}) = -hP(A^+ b)_{li}, \quad i = 1, \ldots, s$$

und (3.106) verwandelt sich in

$$W_{1,li}\left\{-(W_1 b)_x A^+ b + (W_1 b)_t\right\}_{li} = 0, \quad i = 1, \ldots, s. \tag{3.107}$$

Ebenso wie bei den BDF-Verfahren erhält man, dass die versteckte Nebenbedingung (2.17) für die RK-Stufen eine hinreichende Bedingung ist, damit das Kommutativitätsdiagramm 3.1 gilt.

Das Analogon des Kriteriums 3.5.1 lautet für die IRK-Verfahren:

Kriterium 3.5.2 *Wenn alle Stufen eines IRK-Verfahrens zu der Mannigfaltigkeit $M_1(t_i)$ gehören, dann kommutieren die IRT und die IRK-Diskretisierung.*

Natürlich ist dieses Kriterium, wie Kriterium 3.5.1 auch, aus praktischer Sicht unbrauchbar, aber mit ihrer Hilfe kann man für Klassen von Aufgaben anwendbare Kriterien finden. Ein Beispiel dafür sind die schon erwähnten klassischen MNA-Modelle der Schaltkreissimulation. Es werden hier Probleme betrachtet, die die topologischen Annahmen von [Estévez Schwarz and Tischendorf, 1998], [Estévez Schwarz, 2000] erfüllen. Die MNA-Gleichungen besitzen, unter anderen, folgende Eigenschaften ([Estévez Schwarz and Tischendorf, 1998], [Estévez Schwarz, 2000]):

1. der Raum $S_1(x', x, t)$ ist konstant
2. der Raum $\operatorname{im} G_1$ ist ebenfalls konstant
3. die Gleichung $(W_1 b)(x, t) = 0$ ist linear in x.

Aus der dritten Aussage folgt

$$(W_1 b)(x, t) = (W_1 b)_x(t)x + (W_1 b)(0, t). \tag{3.108}$$

Andererseits gilt, wenn 2. erfüllt ist, für $W_1 B$

$$W_1 B(x', x, t) = W_1 \left\{b_x(x, t) + [A(x, t)x']_x\right\} = (W_1 b)_x(t).$$

So reduziert sich die sichtbare Nebenbedingung $(W_1 b)(x, t) = 0$ zu

$$(W_1 B)(t)x = 0 \Leftrightarrow x \in S_1,$$

vorausgesetzt, die Gleichung $(W_1 b)(0, t) = 0$ gilt. Die letzte Annahme bedeutet bei den betrachteten Modellen, dass der Schaltkreis weder gesteuerte

Spannungs- noch gesteuerte Stromquellen enthält, [Estévez Schwarz and Tischendorf, 1998], [Estévez Schwarz, 2000].

Unter den genannten Bedingungen gelten dann $PQ_*^{S_1}x = 0$ und die Bedingung (3.98), falls die Approximation der Lösung an der Stelle t_{l-1} zu $M_0(t_{l-1})$ gehört. In der Tat

$$\begin{aligned}
& W_{1,li}\left\{(W_1b)_{x,li}\sum_{j=1}^{s}\hat{\alpha}_{ij}(X_{lj}-x_{l-1})+h(W_1b)_{t,li}\right\}\\
= \;& W_1\left\{(W_1B)_{li}\sum_{j=1}^{s}\hat{\alpha}_{ij}(X_{lj}-x_{l-1})+h(W_1B)_{t,li}X_{li}\right\}\\
= \;& W_1\left\{(W_1B)_{li}\sum_{j=1}^{s}\hat{\alpha}_{ij}PP_{S_1}^{*}(X_{lj}-x_{l-1})+h(W_1B)_{t,li}(PP_{S_1}^{*}X)_{li}\right\}\\
= \;& W_1\left\{(W_1B)_{li}PP_{S_1}^{*}\sum_{j=1}^{s}\hat{\alpha}_{ij}(X_{lj}-x_{l-1})+h(W_1B)_{t,li}PP_{S_1}^{*}X_{li}\right\}\\
= \;& 0,
\end{aligned}$$

und damit liegt für solche Aufgaben die Kommutativität zwischen der Index-Reduktion und der IRK-Diskretisierung vor. Es ist nicht schwer zu erkennen, dass für die BDF-Verfahren die gleiche Aussage bezüglich der Bedingung (3.98) gilt. Die Vorgehensweise für den Beweis ist die gleiche. Allerdings kann man bei den IRK-Verfahren mit Hilfe des Theorems 2.3.5 eine stärkere Aussage formulieren.

Theorem 3.5.3 *Seien ein klassisches MNA-Modell der Form (3.93) (P-)kontraktiv auf der Lösungsmannigfaltigkeit $M_1(t)$ und die Voraussetzungen 1-3 erfüllt. Dann ist die Anwendung eines algebraisch stabilen IRK(DAE)-Verfahrens auf die MNA-Gleichung (P-)kontraktiv.*

Kapitel 4

Stabilisierung

Im Kapitel 2 wurden Eigenschaften einer Index-2-Gleichung gefunden, die garantieren, dass ein BDF- oder IRK-Verfahren seine GDGl-Stabilitätseigenschaften beibehält. Das Beispiel 3.0.8 zeigt allerdings, dass schwer wiegende Stabilitätsprobleme auftreten können, wenn eine Aufgabe die genannten Eigenschaften nicht besitzt.

In 3.5 wurde die entscheidende Rolle der Gleichungen (3.95), (3.96), (3.99) (für das BDF-Verfahren beispielsweise) für eine stabile diskrete Lösung festgestellt. Insbesondere wurde aufgedeckt, dass die bloße Anwendung des Verfahrens im Allgemeinen nur die Erfüllung von (3.95) und (3.96) garantiert. Alle drei Gleichungen stellen ohne Weiteres ein überbestimmtes Gleichungssystem auf, allerdings liegt mit den klassischen MNA-Gleichungen ein Fall vor, bei dem die Gleichung (3.99) automatisch erfüllt wird.

Wie in 3.5 erwägt, sind in der Fachliteratur verschiedene Stabilisierungsansätze zu finden, die das System (3.95), (3.96), (3.99) in irgendeiner Art und Weise lösen. Allerdings kann man zwei Arten von Ansätzen erkennen, [Eich et al., 1990], [Eich-Soellner and Führer, 1998]:

1. Ansätze, die auf einer Projektion der approximierten Lösung basieren (Koordenatenprojektion)

2. Ansätze, die auf einer Projektion des Residuums basieren (Ableitungsprojektion).

Beide Ansätze sind nicht an ein Diskretisierungsverfahren gebunden und können deswegen beliebig mit numerischen Methoden konbiniert werden.

Diese Ansätze kommen im Wesentlichen aus der Mechanik der Mehrkörpersysteme, [Eich et al., 1990], wo Gleichungen in Hessenberg-Form auftreten. Die Index-2-Formulierung solcher Aufgaben hat die Form

$$\begin{aligned} x_1'(t) + B_{11}(t)x_1(t) + B_{12}(t)x_2(t) &= q_1(t), && (4.1)\\ B_{21}(t)x_1(t) &= q_2(t), && (4.2) \end{aligned}$$

siehe Beispiel 2.1.9. Sei Q der Orthoprojektor auf N und W_1 als Projektor längs im G_1

$$W_1 = Q = \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix}$$

gewählt.

Die Projektionen auf $\operatorname{im}(I - W_1)$ und $\operatorname{im} W_1$ für (4.1), (4.2) ergeben, ungeachtet der durch die Projektionen entstehenden Nullen, erneut diese Gleichungen. Um die versteckte Nebenbedingung zu berechnen, muss die Gleichung (4.2) nach t differenziert werden. Das ergibt

$$\begin{aligned} B_{21}'(t)x_1(t) + B_{21}(t)x_1' &= q_2'(t), \\ (B_{21}'(t) - B_{21}(t)B_{11}(t))x_1(t) - B_{21}(t)B_{12}(t)x_2(t) &= q_2'(t) - B_{21}(t)q_1(t). && (4.3) \end{aligned}$$

Der Ansatz vom Typ 1 wurde zuerst für GDGln mit Invarianten in [Shampine, 1986] eingeführt und schlägt vor, das Index-1-System (4.1), (4.3) (siehe 3.3) mit dem ausgewählten Verfahren zu lösen und nach jedem Schritt die erhaltene Näherungslösung orthogonal auf die Mannigfaltigkeit

$$M_0(t) = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in I\!R^m : B_{21}(t)x_1(t) = q_2(t) \right\}$$

zu projizieren.

Die Anwendung dieses Ansatzes auf das Beispiel 3.0.8 ist, wie die Abbildung 4.1 zeigt, erfolglos. Hier wurde wie bei den Ergebnissen in der Abbildung 3.2 das Euler-Verfahren für die Index-1-Formulierung verwendet und in jedem Schritt die genannte Projektion durchgeführt.

Die Linie $\eta = 0$ repräsentiert den Fall, in dem die Räume N_1 und S_1 konstant und damit die Ergebnisse aus Kapitel 3 anwendbar sind. Wenn das Verfahren das asymptotische Verhalten der exakten Lösung widerspiegeln würde, dann müsste die Fläche für alle Werte von η wie bei $\eta = 0$ aussehen.

Die vorhandene Situation in dem Koordinatenprojektionsansatz kann wie in der Abbildung 4.2 für die differenzierbaren Komponenten geometrisch dargestellt werden.

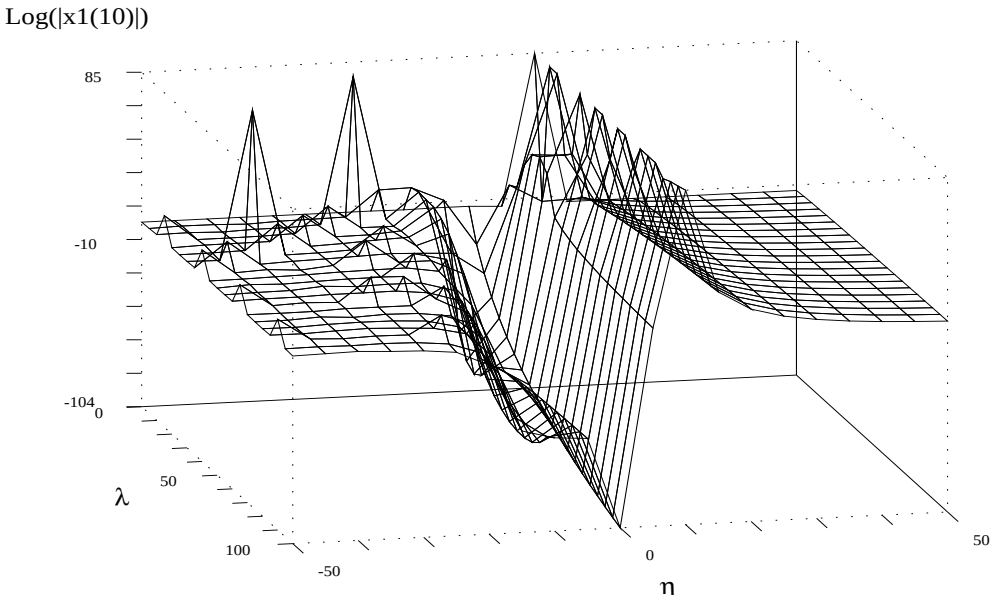

Abbildung 4.1: Ergebnisse für das Beispiel 3.0.8 unter Verwendung des impliziten Euler-Verfahrens für den Koordenatenprojektionsansatz mit einer Schrittweite $h = 0.1$ und verschiedenen Werten von λ und η.

Ein Schritt des Koordenatenprojektionsansatzes kann wie folgt beschrieben werden: Mit dem impliziten Euler-Verfahren wird die Ableitung x_1' durch den Differenzenquotienten approximiert. Man kennt jedoch nur die Größenordnung des dadurch verursachten Diskretisierungsfehlers. Insbesondere hat man keine Information über die Projektion von $\widetilde{x}_1'$ auf dem Tangentialraum an der Stelle x_1. Außerdem liegt die mit dieser möglicherweise falschen Richtung berechnete Lösung ($\widetilde{x}_1$) auf der Mannigfaltigkeit H (siehe (2.18)), jedoch in der Regel außerhalb der Mannigfaltigkeit M_0, da die sichtbare Nebenbedingung (4.2) in der Index-1-Aufgabe (4.1), (4.3) nicht berücksichtigt wird. Anschließend wird $\widetilde{x}_1$ orthogonal auf M_0 projiziert ($\bar{x}_1$) und dieser Wert als neue Approximation übernommen. Eine andere Variante dieses Verfahrens besteht darin, statt auf M_0 auf M_1 zu projizieren. Das ändert aber nichts an der Tatsache, dass die Dynamik der Aufgabe nicht berücksichtigt wird.

Die Verfahrensweise von Typ 2 betrachtet auch das Index-1-System (4.1), (4.3) mit der zusätzlichen Bedingung (4.2). Er fordert in jedem Integrationsschritt die Erfüllung der Gleichungen (4.2), (4.3), also die Zugehörigkeit der

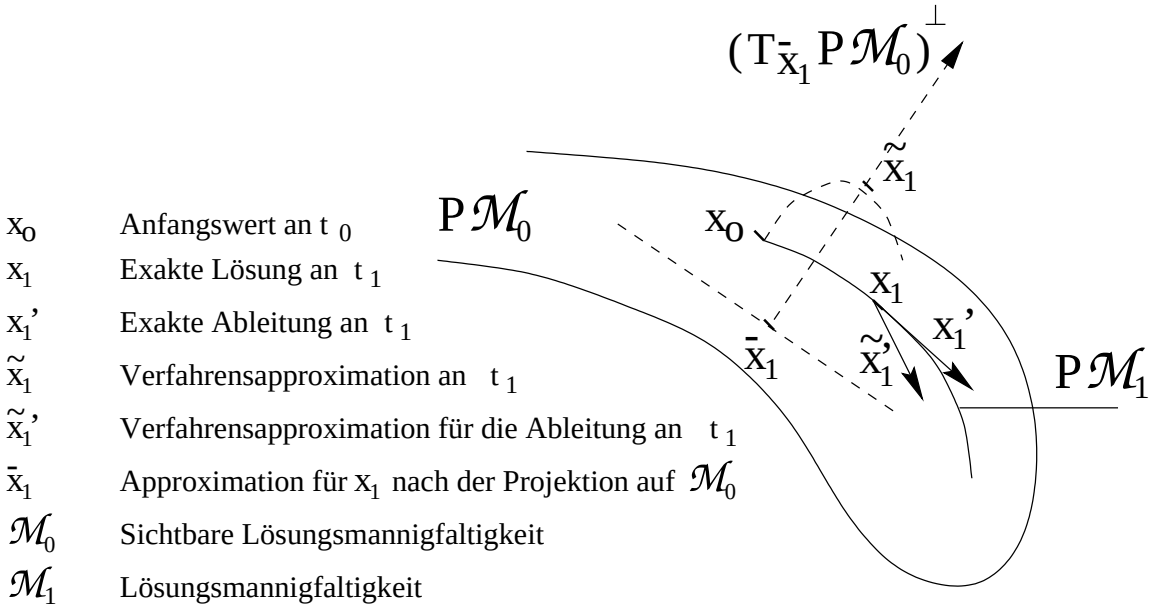

Abbildung 4.2: Geometrische Interpretation des Koordenatenprojektionsansatzes.

Approximation zu M_1, und lässt für die Diskretisierung von (4.1) einen Fehler im im $B_{21}^T(t)$ zu. Bemerkenswert an dieser Stelle ist es, dass im $B_{21}^T(t) \times \{0\}$ der Orthogonalraum nach dem euklidischen Skalarprodukt von S_1 ist und S_1, für die homogene Aufgabe, M_1 enthält. Insbesondere gilt in diesem Fall ([März and Rodríguez Santiesteban, 1999])

$$PM_1(t) = PS_1(t) = \left\{ \begin{pmatrix} x_1 \\ 0 \end{pmatrix} : B_{21}(t)x_1 = 0 \right\}.$$

Die beschriebene Methode entspricht der Diskretisierung der Gear-Gupta-Leimkuhler (GGL)-Formulierung für (4.1), (4.2), [Hairer and Wanner, 1991]. Sie lautet

$$\begin{aligned} x_1'(t) + B_{11}(t)x_1(t) + B_{12}(t)x_2(t) + B_{21}^T(t)\mu(t) &= q_1(t), && (4.4)\\ (B_{21}'(t) - B_{21}(t)B_{11}(t))x_1(t) - B_{21}(t)B_{12}(t)x_2(t) &= q_2'(t) && (4.5)\\ &\quad -B_{21}(t)q_1(t), \\ B_{21}(t)x_1(t) &= q_2(t), && (4.6) \end{aligned}$$

wobei der eingeführte Vektor μ die Dimension $m - r$ hat, wenn $x_1 \in I\!R^r$.

Die Situation ist für die differenzierbaren Komponenten in der Abbildung 4.3 geometrisch dargestellt:

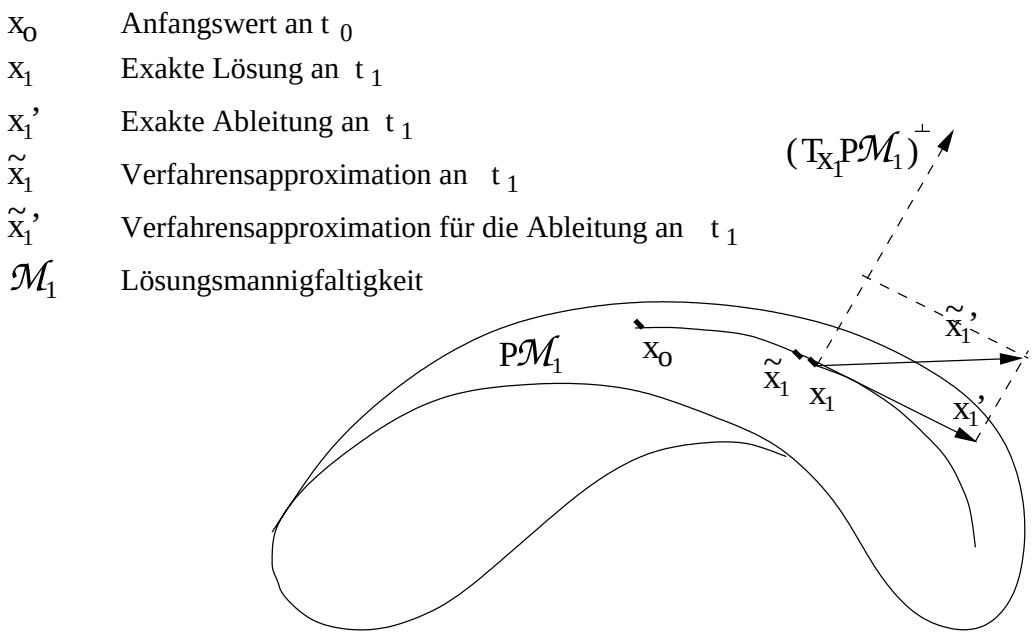

Abbildung 4.3: Geometrische Interpretation des Ableitungsprojektionsansatzes.

Nach einem Integrationsschritt mit dem impliziten Euler-Verfahren liegt erstens, im Gegensatz zu dem Koordenatenprojektionsansatz, die Lösungsapproximation $\widetilde{x}_1$ auf M_1, weil die Aufgabe (4.4)-(4.6) sowohl die sichtbare als auch die versteckte Nebenbedingung enthält. Zweitens ist die Orthoprojektion von $\widetilde{x}_1'$ auf den Tangentialraum in x_1 identisch mit x_1'. Die Erklärung dafür besteht darin, dass das Residuum bei der Diskretisierung der Gleichung (4.1) im $\operatorname{im} B_{21}^T(t) \simeq S_1^{\perp}(t)$ liegt. Diese zwei Fakten wecken die Vermutung, dass dieser Ansatz die Stabilitätsprobleme lösen kann, da zusäzlich zu der M_1-Zugehörigkeit die Dynamik des Systems respektiert wird.

In der Abbildung 4.4 sind die Ergebnisse der Anwendung des impliziten Euler-Verfahrens auf die GGL-Formulierung (4.4)-(4.6) für das Beispiel 3.0.8 dargestellt. Der Stabilisierungseffekt ist deutlich zu erkennen, die Näherungslösung besitzt das gleiche asymptotische Verhalten wie die exakte Lösung.

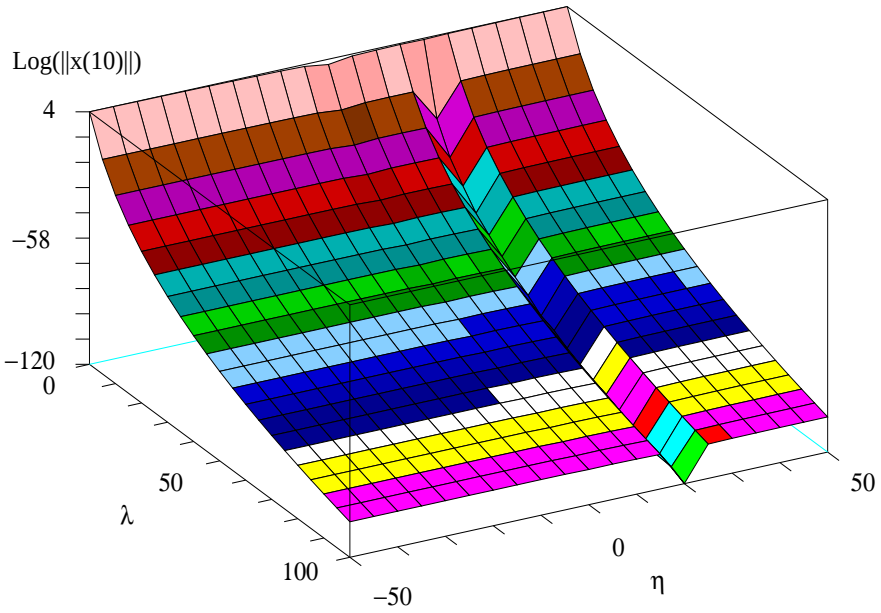

Abbildung 4.4: Ergebnisse für das Beispiel 3.0.8 unter Verwendung des impliziten Euler-Verfahrens für den GGL-Ansatz mit Schrittweite $h = 0.1$ und verschiedenen Werten von λ und η.

Die hier erwähnten Argumente und numerischen Ergebnisse sind natürlich noch kein Beweis dafür, dass die GGL-Formulierung die Stabilitätsprobleme behebt. Das ist allerdings das Ziel dieses Kapitels. Es werden zunächst einige Eigenschaften der GGL-Formulierung im Hessenberg-Fall gezeigt, dann eine Verallgemeinerung dieser Formulierung für Nicht-Hessenberg-Systeme eingeführt und anschließend analoge Eigenschaften dieser Verallgemeinerung bewiesen. Die wichtigste Aussage für die GGL-Formulierung belegt die Kontraktivität von IRK(DAE)-Verfahren für die GGL-Formulierung und erklärt zugleich die vorgestellten numerischen Ergebnisse.

4.1 Eigenschaften der GGL-Formulierung im Hessenberg-Fall

Das System (4.4)-(4.6) stellt eine ADGl dar, die mit Hilfe der Notation

$$\hat{x} := \begin{pmatrix} x_1 \\ x_2 \\ \mu \end{pmatrix}, \quad \widehat{A} := \begin{pmatrix} I & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \widehat{B} := \begin{pmatrix} B_{11} & B_{12} & B_{21}^T \\ B_{21}' - B_{21}B_{11} & B_{21}B_{12} & 0 \\ B_{21} & 0 & 0 \end{pmatrix},$$

und

$$\widehat{P} := \begin{pmatrix} P & 0 \\ 0 & 0 \end{pmatrix}, \quad \hat{q} := \begin{pmatrix} q_1 \\ q_2' - B_{21}q_1 \\ q_2 \end{pmatrix},$$

als

$$\widehat{A}(\widehat{P}\hat{x})'(t) + \widehat{B}(t)\hat{x}(t) = \hat{q}(t), \tag{4.7}$$

geschrieben werden kann.

In dieser Sektion werden alle zu (4.7) gehörenden Objekte (Matrizen, Projektoren und Räume) mit $\widehat{\ }$ gekennzeichnet.

Lemma 4.1.1 *Für die Aufgabe (4.7) ist die sichtbare Lösungsmannigfaltigkeit* $\widehat{M}_0(t)$ *durch*

$$\widehat{M}_0(t) = M_1(t) \times I\!R^{m-r}$$

gegeben.

Beweis: D ie sichtbare Lösungsmannigfaltigkeit $\widehat{M}_0(t)$ ist nach der Definition durch

$$\widehat{M}_0(t) = \left\{ \hat{z} = \begin{pmatrix} z_1 \\ z_2 \\ z_\mu \end{pmatrix} \in I\!R^r \times I\!R^{m-r} \times I\!R^{m-r} : \left[\widehat{B}(t)\hat{z} - \hat{q}(t)\right] \in \operatorname{im} \widehat{A} \right\}$$

bestimmt, das bedeutet konkret

$$\begin{aligned} (B_{21}' - B_{21}B_{11})z_1 - B_{21}B_{12}z_2 &= q_2' - B_{21}q_1, \\ B_{21}z_1 &= q_2, \end{aligned}$$

was die Zugehörigkeit von $\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$ zu M_1 ergibt.

Das Lemma bringt eine sehr wichtige Aussage aus der Stabilitätssicht, nämlich dass die Mannigfaltigkeit $M_1(t)$ durch den GGL-Ansatz zu $\widehat{M}_0(t)$ "wandert".

Lemma 4.1.2 *Falls die ursprüngliche Aufgabe (4.1), (4.2) Index-2 ist, dann ist der Tractability-Index von (4.14) größer als 1, und*

$$\widehat{G}_1(t) = \begin{pmatrix} I_r & B_{12}(t) & B_{21}^T(t) \\ 0 & -(B_{21}B_{12})(t) & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \widehat{W}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & I_{m-r} \end{pmatrix}.$$

Beweis:

$$\begin{aligned} \widehat{G}_1(t) &= \begin{pmatrix} I_r & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} B_{11}(t) & B_{12}(t) & B_{21}^T(t) \\ (B_{21}' - B_{21}B_{11})(t) & -(B_{21}B_{12})(t) & 0 \\ B_{21}(t) & 0 & 0 \end{pmatrix} \\ &\quad \times \begin{pmatrix} 0 & 0 & 0 \\ 0 & I_{m-r} & 0 \\ 0 & 0 & I_{m-r} \end{pmatrix} \\ &= \begin{pmatrix} I_r & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & B_{12}(t) & B_{21}^T(t) \\ 0 & -(B_{21}B_{12})(t) & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} I_r & B_{12}(t) & B_{21}^T(t) \\ 0 & -(B_{21}B_{12})(t) & 0 \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Offensichtlich ist $\widehat{G}_1(t)$ singulär für alle t. Auf Grund der Regularität von $(B_{21}B_{12})(t)$ besitzt diese Matrix konstanten Rank m für alle t und eine mögliche Wahl für $\widehat{W}_1$ ist durch den angegebenen Ausdruck bestimmt.
Jetzt, da der Ausdruck für $\widehat{G}_1(t)$ bekannt ist, können für (4.7) die relevanten Räume $\widehat{N}_1(t)$ und $\widehat{S}_1(t)$ berechnet werden.

Für $\hat{z} \in \widehat{N}_1(t)$ gilt

$$\begin{pmatrix} I_r & B_{12}(t) & B_{21}^T(t) \\ 0 & -(B_{21}B_{12})(t) & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{z}_1 \\ \hat{z}_2 \\ \hat{z}_\mu \end{pmatrix} = 0.$$

Diese Gleichung ist äquivalent zu

$$\begin{pmatrix} I_r & B_{12}(t) \\ 0 & -(B_{21}B_{12})(t) \end{pmatrix} \begin{pmatrix} \hat{z}_1 \\ \hat{z}_2 \end{pmatrix} = - \begin{pmatrix} B_{21}^T(t)\hat{z}_\mu \\ 0 \end{pmatrix}.$$

Aus der letzten Gleichung geht hervor, dass

$$\begin{pmatrix} \hat{z}_1 \\ \hat{z}_2 \end{pmatrix} = -\begin{pmatrix} I_r & B_{12}(t)(B_{21}B_{12})^{-1}(t) \\ 0 & -(B_{21}B_{12})^{-1}(t) \end{pmatrix} \begin{pmatrix} B_{21}^T(t)\hat{z}_\mu \\ 0 \end{pmatrix} = -\begin{pmatrix} B_{21}^T(t) \\ 0 \end{pmatrix} \hat{z}_\mu,$$

und damit folgt für $\widehat{N}_1(t)$

$$\widehat{N}_1(t) = \left\{ \hat{z} \in I\!R^r \times I\!R^{m-r} \times I\!R^{m-r} : \begin{pmatrix} \hat{z}_1 \\ \hat{z}_2 \end{pmatrix} = -\begin{pmatrix} B_{21}^T(t) \\ 0 \end{pmatrix} \hat{z}_\mu \right\}.$$

Als möglicher Projektor auf $\widehat{N}_1(t)$ kann

$$\widehat{Q}^*_{\widehat{N}_1}(t) = \begin{pmatrix} 0 & 0 & -B_{21}^T(t) \\ 0 & 0 & 0 \\ 0 & 0 & I_{m-r} \end{pmatrix}$$

gewählt werden.

Bezüglich $\widehat{S}_1(t)$ hat man nach der Definition

$$\widehat{S}_1(t) = \left\{ \hat{z} \in I\!R^r \times I\!R^{m-r} \times I\!R^{m-r} : \widehat{B}(t)\widehat{P}\hat{z} \in \operatorname{im} \widehat{G}_1(t) \right\}.$$

Das bedeutet

$$\begin{aligned} \widehat{B}\widehat{P}\hat{z} &= \begin{pmatrix} B_{11} & B_{12} & B_{21}^T \\ B_{21}' - B_{21}B_{11} & -B_{21}B_{12} & 0 \\ B_{21} & 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{z}_1 \\ 0 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} B_{11}\hat{z}_1 \\ (B_{21}' - B_{21}B_{11})\hat{z}_1 \\ B_{21}\hat{z}_1 \end{pmatrix} \in I\!R^r \times I\!R^{m-r} \times \{0\}, \end{aligned}$$

woraus folgt (siehe 2.1.9)

$$\widehat{S}_1(t) = S_1(t) \times I\!R^{m-r}.$$

An dieser Stelle ist darauf zu hinweisen, dass bezüglich der Komponenten $\hat{z}_1$, $\hat{z}_2$ die Räume $\widehat{S}_1(t)$ und $\widehat{N}_1(t)$ orthogonal in dem euklidischen Skalarprodukt sind.

Lemma 4.1.3 *Wenn die ursprüngliche Aufgabe (4.1)-(4.2) index-2-tractable ist, dann ist auch die Aufgabe (4.7) index-2-tractable.*

Beweis: E s wird gezeigt, dass

$$\widehat{S}_1(t) \cap \widehat{N}_1(t) = \{0\}\,,\ \forall t.$$

Sei $\hat{z} \in \widehat{S}_1(t) \cap \widehat{N}_1(t)$, dann gelten die Gleichungen

$$\begin{aligned} \begin{pmatrix} \hat{z}_1 \\ \hat{z}_2 \end{pmatrix} &\in S_1(t), \\ \begin{pmatrix} \hat{z}_1 \\ \hat{z}_2 \end{pmatrix} &= -\begin{pmatrix} B_{21}^T(t) \\ 0 \end{pmatrix} \hat{z}_\mu, \end{aligned}$$

und folglich $\begin{pmatrix} B_{21}^T(t) \\ 0 \end{pmatrix} \hat{z}_\mu \in S_1(t)$ ebenfalls. Die letzte Aussage bedeutet (durch die Definition von $S_1(t)$, siehe auch das Beispiel 2.1.9):

$$B_{21}(t)B_{21}^T(t)\hat{z}_\mu = 0.$$

Diese Gleichung impliziert $\hat{z}_\mu = 0$, weil nach der Index-2-Annahme $B_{21}(t)$ vollen Rank besitzt. Schließlich erhält man $\hat{z}_1 = \hat{z}_2 = \hat{z}_\mu = 0$.
Ein weiterer Unterraum von Bedeutung bei den Index-2-Aufgaben ist $\widehat{N} \cap \widehat{S}(t)$. In 3.1.1 wurde festgestellt, dass in diesem Unterraum jene Komponenten leben, die in der Differentiationsaufgabe einbezogen sind.

Lemma 4.1.4 *Für die Aufgabe (4.7) gilt*

$$\begin{aligned} \widehat{N} \cap \widehat{S}(t) &= \{0\}_r \times \{0\}_{m-r} \times I\!R^{m-r}, \\ \widehat{T} &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & I_{m-r} \end{pmatrix}. \end{aligned}$$

Beweis: A us 3.1 weiß man, dass die Gleichung

$$\widehat{N} \cap \widehat{S}(t) = \operatorname{im} \widehat{Q}\widehat{Q}^*_{\widehat{N}_1}(t)$$

unabhängig von den Projektoren $\widehat{Q}$ und $\widehat{Q}^*_{\widehat{N}_1}(t)$ Gültigkeit besitzt. Die einfache Berechnung

$$\widehat{Q}\widehat{Q}^*_{\widehat{N}_1}(t) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & I_{m-r} & 0 \\ 0 & 0 & I_{m-r} \end{pmatrix} \begin{pmatrix} 0 & 0 & -B_{21}^T(t) \\ 0 & 0 & 0 \\ 0 & 0 & I_{m-r} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & I_{m-r} \end{pmatrix},$$

ergibt dann die Aussage des Lemmas.

Lemma 4.1.5 *Die Lösungsmannigfaltigkeit $\widehat{M}_1(t)$ von (4.7) ist durch*

$$\widehat{M}_1(t) = \left\{ \hat{z} = \begin{pmatrix} \hat{z}_1 \\ \hat{z}_2 \\ \hat{z}_\mu \end{pmatrix} \in I\!R^r \times I\!R^{m-r} \times I\!R^{m-r} : \hat{z} \in \widehat{M}_0(t) \wedge \hat{z}_\mu = 0 \right\} = M_1(t) \times \{0\}$$

gegeben.

Beweis: D ie Mannigfaltigkeit $\widehat{M}_1(t)$ ist nach der Definition durch

$$\widehat{M}_1(t) = \left\{ \begin{array}{c} \hat{z} \in I\!R^r \times I\!R^{m-r} \times I\!R^{m-r} : \hat{z} \in \widehat{M}_0(t), \\ \widehat{W}_1(t) \left[(\widehat{W}_1 \widehat{B})' - \widehat{W}_1 \widehat{B} \widehat{A}^+ \widehat{B} \right](t) \hat{z} = \widehat{W}_1(t) \left[(\widehat{W}_1 \hat{q})' - \widehat{W}_1 \widehat{B} \widehat{A}^+ \hat{q} \right](t) \end{array} \right\}$$

bestimmt. Jetzt wird gezeigt, dass die versteckte Nebenbedingung nichts anderes als $\hat{z}_\mu = 0$ ist.

Der Term $\widehat{W}_1 \widehat{B} \widehat{A}^+$ ist

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ B_{21} & 0 & 0 \end{pmatrix} \begin{pmatrix} I_r & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ B_{21} & 0 & 0 \end{pmatrix},$$

dann folgt

$$\begin{aligned} \widehat{W}_1 \widehat{B} \widehat{A}^+ \widehat{B} &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ B_{21} & 0 & 0 \end{pmatrix} \begin{pmatrix} B_{11}(t) & B_{12}(t) & B_{21}^T(t) \\ (B_{21}' - B_{21} B_{11})(t) & -(B_{21} B_{12})(t) & 0 \\ B_{21}(t) & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ B_{21} B_{11} & B_{21} B_{12} & B_{21} B_{21}^T \end{pmatrix}. \end{aligned}$$

Andererseits ist

$$\widehat{W}_1 \widehat{B} \widehat{A}^+ \hat{q} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ B_{21} & 0 & 0 \end{pmatrix} \begin{pmatrix} q_1 \\ q_2' - B_{21} q_1 \\ q_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ B_{21} q_1 \end{pmatrix},$$

und die versteckte Nebenbedingung kann folgendermaßen geschrieben werden:

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & I_{m-r} \end{pmatrix} \left[\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ B'_{21} - B_{21}B_{11} & -B_{21}B_{12} & -B_{21}B_{21}^T \end{pmatrix} \begin{pmatrix} \hat{z}_1 \\ \hat{z}_2 \\ \hat{z}_\mu \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ q'_2 - B_{21}q_1 \end{pmatrix} \right] = 0,$$

$$\begin{pmatrix} 0 \\ 0 \\ (B'_{21} - B_{21}B_{11})\hat{z}_1 - B_{21}B_{12}\hat{z}_2 - B_{21}B_{21}^T\hat{z}_\mu \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ q'_2 - B_{21}q_1 \end{pmatrix} = 0.$$

Wenn $\hat{z} \in \widehat{M_0}$, dann gilt die Gleichung (also die versteckte Nebenbedingung der ursprünglichen Gleichung)

$$(B'_{21} - B_{21}B_{11})\hat{z}_1 - B_{21}B_{12}\hat{z}_2 = q'_2 - B_{21}q_1.$$

Jetzt reduziert sich die versteckte Nebenbedingung für (4.7) zu

$$B_{21}B_{21}^T\hat{z}_\mu = 0,$$

woraus $\hat{z}_\mu = 0$ folgt.

Wenn $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}(\cdot)$ eine Lösung von (4.1), (4.2) ist, dann kann man trivialerweise eine Lösung für (4.7) als $\begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix}(\cdot)$ angeben. Der folgende Korollar stellt die Umkehrung dieser Aussage dar.

Korollar 4.1.6 *Sei Θ die Applikation*

$$\Theta : \mathbb{R}^r \times \mathbb{R}^{m-r} \times \mathbb{R}^{m-r} \to \mathbb{R}^r \times \mathbb{R}^{m-r},$$
$$\Theta \begin{pmatrix} x_1 \\ x_2 \\ x_\mu \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Dann ist die Einschränkung von Θ auf $\widehat{M_1}(t)$ eine Bijektion zwischen $\widehat{M_1}(t)$ und $M_1(t)$. Das bedeutet insbesondere, dass, wenn $\hat{x}(\cdot) = \begin{pmatrix} \hat{x}_1(\cdot) \\ \hat{x}_2(\cdot) \\ \hat{x}_\mu(\cdot) \end{pmatrix}$ eine $\hat{C}^1_{\widehat{N}}$-Lösung von (4.7) ist, $\begin{pmatrix} \hat{x}_1(\cdot) \\ \hat{x}_2(\cdot) \end{pmatrix}$ eine C^1_N-Lösung von (4.1), (4.2) darstellt.

4.2 Die allgemeine lineare GGL-Formulierung

Der am Anfang des Kapitels vorgestellte Ableitungsprojektionsansatz kann für ein vollimplizites Index-2-Problem

$$A(t)(Px)'(t) + B(t)x(t) = q(t), \quad t \in \Im, \tag{4.8}$$

wobei die Gleichungsanteile W_1B, W_1q stetig differenzierbar sind, verallgemeinert werden.

Wie aus Abschnitt 2.1.2 ersichtlich, ist die Gleichung (4.8) äquivalent zu dem System

$$\begin{aligned}
A(t)(Px)'(t) + (I - W_1(t))B(t)x(t) &= (I - W_1(t))q(t), && (4.9)\\
W_1(t)B(t)x(t) &= (W_1q)(t), && (4.10)\\
W_1(t)\left[(W_1B)'(t) - (W_1BA^+B)(t)\right]x(t) &= W_1(t) && (4.11)\\
&\quad \times \left[(W_1q)'(t) - (W_1BA^+q)(t)\right].
\end{aligned}$$

Seien die Spalten von $\Phi(t)$, $\Gamma(t)$ jeweils eine Basis von $\operatorname{im} W_1(t)$, $\operatorname{im} A_1(t)^\perp$ und $\kappa = \dim N_1(t)$ die Spaltenanzahl dieser Matrizen. So folgt für $W_1(t)$ der Ausdruck

$$W_1(t) = \Phi(t)\left[\Gamma(t)^T\Phi(t)\right]^{-1}\Gamma(t)^T.$$

Mit Hilfe der letzten Gleichung kann man aus (4.10) die linear unabhängigen Bedingungen herausfiltern:

$$\begin{aligned}
&(W_1B)(t)x(t) = W_1(t)q(t),\\
&\left[\Phi(\Gamma^T\Phi)^{-1}\Gamma^TB\right](t)x(t) = \left[\Phi(\Gamma^T\Phi)^{-1}\Gamma^T\right](t)q(t),\\
&(\Gamma^TB)(t)x(t) = \Gamma^T(t)q(t).
\end{aligned}$$

Der Vorschlag für einen verallgemeinerten GGL-Ansatz lautet dann

$$\begin{aligned}
&A(t)(Px)'(t) + (I - W_1)(t)B(t)x(t) && (4.12)\\
&+W_1(t)\left[(W_1B)' - W_1BA^+B\right](t)x(t)\\
&\qquad -(A\Psi)(t)\mu(t) &&= (I - W_1)(t)q(t)\\
&\qquad\qquad &&\quad +W_1(t)\left[(W_1q)' - W_1BA^+q\right](t),\\
&\qquad (\Gamma^TB)(t)x(t) &&= \Gamma^T(t)q(t), && (4.13)
\end{aligned}$$

wobei die $m \times \kappa$-Matrix $\Psi(t)$ eine Basis eines transversalen Raumes zu $S_1(t)$ bildet. Hier gilt für die neue Variable $\mu(t) \in I\!R^\kappa$. Die Berechnung der Matrix $\Psi(t)$ ist im Allgemeinen nicht trivial und benötigt die Verwendung eines Algorithmus aus der Linearalgebra. Allerdings ist diese Aufgabe im Hessenberg-Fall viel einfacher. Man beachte, dass (4.12), (4.13) für ein Hessenberg-System, mit

$$\Gamma = \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix}$$

und $\Psi(t) = -(B^T\Gamma)(t)$, den GGL-Ansatz (4.4)-(4.6) ergibt.

Wie im Hessenberg-Fall wird das System (4.12), (4.13) als eine ADGl der Form

$$\widehat{A}(t)(\widehat{P}\hat{x})'(t) + \widehat{B}(t)\hat{x}(t) = \hat{q}(t) \tag{4.14}$$

ausgedrückt, wobei jetzt

$$\hat{x} := \begin{pmatrix} x \\ \mu \end{pmatrix}, \quad \hat{A} := \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix},$$
$$\widehat{B} := \begin{pmatrix} (I - W_1)B + W_1\left[(W_1B)' - W_1BA^+B\right] & -A\Psi \\ \Gamma^T B & 0 \end{pmatrix},$$

und

$$\widehat{P} := \begin{pmatrix} P & 0 \\ 0 & 0 \end{pmatrix}, \quad \hat{q} := \begin{pmatrix} (I - W_1)q + W_1\left[(W_1q)' - W_1BA^+q\right] \\ \Gamma^T q \end{pmatrix}.$$

Wieder werden alle auf (4.14) bezogenen Objekte (wie Matrizen, Projektoren und Räume) mit $\widehat{\ }$ gekennzeichnet.

4.2.1 Eigenschaften des GGL-Ansatzes im linearen Fall

In dieser Sektion wird gezeigt, dass die verallgemeinerte GGL-Formulierung (4.12), (4.13) die gleichen Eigenschaften besitzt, die in 4.1 für die Hessenberg-Formulierung (4.4)-(4.6) gezeigt wurden.

Lemma 4.2.1 *Für die Aufgabe (4.14) ist die sichtbare Lösungsmannigfaltigkeit* $\widehat{M}_0(t)$ *durch*

$$\widehat{M}_0(t) = M_1(t) \times I\!R^\kappa$$

gegeben.

Beweis: D ie sichtbare Lösungsmannigfaltigkeit $\widehat{M}_0(t)$ ist nach der Definition durch

$$\widehat{M}_0(t) = \left\{ \hat{z} = \begin{pmatrix} z \\ \mu \end{pmatrix} \in I\!R^m \times I\!R^\kappa : \left[\widehat{B}(t)\hat{z} - \hat{q}(t) \right] \in \operatorname{im} \widehat{A}(t) \right\}$$

bestimmt. Das bedeutet

$$\begin{aligned} (I - W_1)Bz + W_1 \left[(W_1B)' - W_1BA^+B \right] z - (I - W_1)q & \\ -W_1 \left[(W_1q)' - W_1BA^+q \right] &\in \operatorname{im} A, \\ \Gamma^T(Bz - q) &= 0. \end{aligned}$$

Die zweite Gleichung ist äquivalent zu

$$W_1Bz = W_1q,$$

was dann für die erste Gleichung

$$Bz + W_1 \left[(W_1B)' - W_1BA^+B \right] z - q - W_1 \left[(W_1q)' - W_1BA^+q \right] \in \operatorname{im} A,$$

impliziert. Da W_1 ein Projektor längs $\operatorname{im} A_1(t) \supset \operatorname{im} A(t)$ ist, folgen die zwei Bedingungen

$$\begin{aligned} B(t)z - q(t) &\in \operatorname{im} A(t), \\ W_1(t) \left[(W_1B)' - W_1BA^+B \right](t)z &= W_1(t) \left[(W_1q)' - W_1BA^+q \right](t), \end{aligned}$$

damit ein Vektor $\hat{z}$ zu $\widehat{M}_0(t)$ gehört.

Lemma 4.2.2 *Falls die ursprüngliche Aufgabe (4.8) Index-2 ist, dann ist der Tractability-Index von (4.14) größer als 1, und*

$$\widehat{G}_1(t) = \begin{pmatrix} G_1 - W_1BA^+BQ & -A\Psi \\ 0 & 0 \end{pmatrix}(t), \quad \widehat{W}_1 = \begin{pmatrix} 0 & 0 \\ 0 & I_\kappa \end{pmatrix}.$$

Beweis: Z uerst wird die Matrix $\widehat{G}_1(t)$ berechnet:

$$\begin{aligned} \widehat{G}_1 &= \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} (I - W_1)B + W_1 \left[(W_1B)' - W_1BA^+B \right] & -A\Psi \\ \Gamma^T B & 0 \end{pmatrix} \begin{pmatrix} Q & 0 \\ 0 & I_\kappa \end{pmatrix} \\ &= \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} BQ - W_1BA^+BQ & -AB\Psi \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} G_1 - W_1BA^+BQ & -A\Psi \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

Offensichtlich ist diese Matrix singulär, und wenn

$$G_1 - W_1 B A^+ B Q$$

regulär ist, dann ist $\operatorname{rank} \widehat{G}_1(t)$ konstant und

$$\widehat{W}_1 = \begin{pmatrix} 0 & 0 \\ 0 & I_\kappa \end{pmatrix}$$

eine mögliche Wahl für $\widehat{W}_1$.

In der Tat ist die Matrix $G_1 - W_1 B A^+ B Q$ regulär. Um dies zu zeigen, wird das Gleichungssystem

$$\left(G_1 - W_1 B A^+ B Q\right) z = 0$$

betrachtet.

Diese Gleichung ist äquivalent zu

$$\begin{aligned} &G_1 z = 0, \\ &W_1 B A^+ B Q z = 0. \end{aligned}$$

Das bedeutet in erster Linie, dass $z \in N_1(t)$. Außerdem folgt, wenn die erste Gleichung mit $W_1 B A^+$ multipliziert wird,

$$\begin{aligned} &W_1 B A^+ (A + BQ) z = 0, \\ &W_1 B A^+ A z = 0, \\ &W_1 B z = 0, \end{aligned}$$

was die Zugehörigkeit von z zu $S_1(t)$ mit sich bringt. Folglich ist $z \in S_1(t) \cap N_1(t)$, und da die Aufgabe (4.8) index-2-tractable ist, folgt $z = 0$ und die Regularität der Matrix.

Nachdem man den Ausdruck von $\widehat{G}_1(t)$ kennt, gilt für $\hat{z} \in \widehat{N}_1(t)$

$$\begin{aligned} z &= \left(G_1 - W_1 B A^+ B Q\right)^{-1} A \Psi \mu, \\ &= \left(G_1 - W_1 B A^+ B Q\right)^{-1} \left(G_1 - W_1 B A^+ B Q\right) P \Psi \mu, \\ &= P \Psi \mu. \end{aligned}$$

Infolgedessen ist $\widehat{N}_1(t)$ durch

$$\widehat{N}_1(t) = \left\{ \hat{z} = \begin{pmatrix} z \\ \mu \end{pmatrix} \in I\!R^m \times I\!R^\kappa : z = P \Psi(t) \mu \right\}$$

gegeben, und

$$\widehat{Q}^*_{\widehat{N}_1}(t) = \begin{pmatrix} 0 & P\Psi(t)) \\ 0 & I_\kappa \end{pmatrix}$$

ist ein Projektor auf $\widehat{N}_1(t)$.

Auf der anderen Seite ist

$$\widehat{S}_1(t) = \left\{ \hat{z} = \begin{pmatrix} z \\ \mu \end{pmatrix} \in I\!R^m \times I\!R^\kappa : \widehat{B}(t)\widehat{P}\hat{z} \in \operatorname{im} \widehat{G}_1(t) \right\},$$

das heißt

$$\widehat{B}(t)\widehat{P}\hat{z} = \begin{pmatrix} \{(I - W_1)B + W_1\left[(W_1B)' - W_1BA^+B\right]\}(t)Pz \\ \Gamma(t)^T B(t)z \end{pmatrix} \in I\!R^m \times \{0\},$$

und daraus folgt, dass $z \in S_1(t)$ und

$$\widehat{S}_1(t) = S_1(t) \times I\!R^\kappa.$$

Lemma 4.2.3 *Wenn die ursprüngliche Aufgabe (4.8) index-2-tractable ist, dann ist auch die Aufgabe (4.14) index-2-tractable.*

Beweis: E s wird gezeigt, dass

$$\widehat{S}_1(t) \cap \widehat{N}_1(t) = \{0\}, \ \forall t.$$

Sei $\hat{z} \in \widehat{S}_1(t) \cap \widehat{N}_1(t)$, dann gelten die Gleichungen

$$\begin{aligned} z &\in S_1(t), \\ z &= P\Psi(t)\mu, \end{aligned}$$

und folglich $\Psi(t)\mu \in S_1(t)$ ebenfalls. Aber da $\operatorname{im} \Psi(t)$ transversal zu $S_1(t)$ ist, muss $\Psi(t)\mu = 0$ gelten, und schließlich erhält man $z = \mu = 0$.
In 3.1.1 spielte die Schnittmenge $N \cap S(t)$ eine wichtige Rolle. Durch den Projektor $T(t)$ auf diesen Raum (und sein Komplement $U(t)$) konnte die Q-Komponente in die Teile TQx und UQx zerlegt werden, wobei nur in dem TQ-Teil eine Differentiationsaufgabe vorkommt. Das folgende Lemma bringt Aufschluss über diesen kritischen Schnittraum.

Lemma 4.2.4 *Für die Aufgabe (4.14) gilt*

$$\begin{aligned} &\widehat{N} \cap \widehat{S}(t) = \{0\} \times I\!R^\kappa, \\ &\widehat{T} = \begin{pmatrix} 0 & 0 \\ 0 & I_\kappa \end{pmatrix}. \end{aligned}$$

Beweis: I n 3.1 ist ersichtlich, dass die Gleichung

$$\widehat{N} \cap \widehat{S}(t) = \operatorname{im} \widehat{Q}\widehat{Q}^*_{\widehat{N}_1}(t)$$

unabhängig von den Projektoren $\widehat{Q}$ und $\widehat{Q}^*_{\widehat{N}_1}(t)$ gilt. Die einfache Berechnung

$$\widehat{Q}\widehat{Q}^*_{\widehat{N}_1}(t) = \begin{pmatrix} Q & 0 \\ 0 & I_\kappa \end{pmatrix} \begin{pmatrix} 0 & P\Psi(t) \\ 0 & I_\kappa \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & I_\kappa \end{pmatrix},$$

ergibt dann die Aussage des Lemmas.

Lemma 4.2.5 *Die Lösungsmannigfaltigkeit $\widehat{M}_1(t)$ von (4.14) ist durch*

$$\widehat{M}_1(t) = \left\{ \hat{z} = \begin{pmatrix} z \\ \mu \end{pmatrix} \in I\!R^m \times I\!R^\kappa : \hat{z} \in \widehat{M}_0(t) \wedge \mu = 0 \right\} = M_1(t) \times \{0\}$$

gegeben.

Beweis: D ie Mannigfaltigkeit $\widehat{M}_1(t)$ ist nach der Definition durch

$$\widehat{M}_1(t) = \left\{ \begin{array}{c} \hat{z} \in I\!R^m \times I\!R^\kappa : \hat{z} \in \widehat{M}_0(t), \\ \widehat{W}_1(t) \left[(\widehat{W}_1\widehat{B})' - \widehat{W}_1\widehat{B}\widehat{A}^+\widehat{B} \right](t)\hat{z} = \widehat{W}_1(t) \left[(\widehat{W}_1\hat{q})' - \widehat{W}_1\widehat{B}\widehat{A}^+\hat{q} \right](t) \end{array} \right\}$$

bestimmt. Es wird gezeigt, dass die versteckte Nebenbedingung nichts anderes als $\mu = 0$ ist.

Der Term $\widehat{W}_1\widehat{B}\widehat{A}^+$ ist

$$\begin{pmatrix} 0 & 0 \\ \Gamma^T B & 0 \end{pmatrix} \begin{pmatrix} A^+ & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ \Gamma^T B A^+ & 0 \end{pmatrix},$$

dann folgt

$$\begin{aligned} \widehat{W}_1\widehat{B}\widehat{A}^+\widehat{B} &= \begin{pmatrix} 0 & 0 \\ \Gamma^T B A^+ & 0 \end{pmatrix} \begin{pmatrix} (I - W_1)B + W_1 \left[(W_1 B)' - W_1 B A^+ B \right] & -A\Psi \\ \Gamma^T B & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ \Gamma^T B A^+ (I - W_1)B & -\Gamma^T B A^+ A\Psi \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ \Gamma^T B A^+ B & -\Gamma^T B \Psi \end{pmatrix}. \end{aligned}$$

Andererseits ist

$$\widehat{W}_1\widehat{B}\widehat{A}^+\hat{q} = \begin{pmatrix} 0 \\ \Gamma^T B A^+ q \end{pmatrix},$$

und die versteckte Nebenbedingung kann folgendermaßen geschrieben werden:

$$\begin{pmatrix} 0 & 0 \\ 0 & I_k \end{pmatrix} \begin{pmatrix} 0 & 0 \\ (\Gamma^T B)' - \Gamma^T B A^+ B & \Gamma^T B \Psi \end{pmatrix} \begin{pmatrix} z \\ \mu \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & I_k \end{pmatrix} \times \begin{pmatrix} 0 \\ (\Gamma^T q)' - \Gamma^T B A^+ q \end{pmatrix},$$

$$\begin{pmatrix} 0 \\ \left[(\Gamma^T B)' - \Gamma^T B A^+ B\right] z + \Gamma^T B \Psi \mu \end{pmatrix} = \begin{pmatrix} 0 \\ (\Gamma^T q)' - \Gamma^T B A^+ q \end{pmatrix}.$$

Wenn $\hat{z} \in \widehat{M_0}$, dann gelten die Gleichungen

$$\begin{aligned} W_1 B z &= W_1 q, \\ W_1 \left[(W_1 B)' - W_1 B A^+ B\right] z &= W_1 \left[(W_1 q)' - W_1 B A^+ q\right], \end{aligned}$$

und daraus folgt

$$\begin{aligned} \left[(W_1 B)' - W_1 B A^+ B\right] z &= \left[(W_1 q)' - W_1 B A^+ q\right], \\ \left\{\left[\Phi(\Gamma^T \Phi)^{-1} \Gamma^T B\right]' - \Phi(\Gamma^T \Phi)^{-1} \Gamma^T B A^+ B\right\} z &= \left\{\left[\Phi(\Gamma^T \Phi)^{-1} \Gamma^T q\right]' \right. \\ &\quad \left. - \Phi(\Gamma^T \Phi)^{-1} \Gamma^T B A^+ q\right\}, \\ \left[(\Gamma^T B)' - \Gamma^T B A^+ B\right] z &= (\Gamma^T q)' - \Gamma^T B A^+ q. \end{aligned}$$

Jetzt reduziert sich die versteckte Nebenbedingung zu

$$\Gamma^T B \Psi \mu = 0,$$

woraus, wegen der Transversalität zwischen $S_1(t)$ und $\operatorname{im} \Psi(t)$, $\mu = 0$ folgt. Wenn $x(\cdot)$ eine Lösung von (4.8) ist, dann kann trivialerweise eine Lösung für (4.14) als $\hat{x}(\cdot) = \begin{pmatrix} x(\cdot) \\ 0 \end{pmatrix}$ angegeben werden. Der folgende Korollar stellt die Umkehrung dieser Aussage dar:

Korollar 4.2.6 *Sei Θ die Applikation*

$$\begin{aligned} &\Theta : \mathbb{R}^m \times \mathbb{R}^\kappa \to \mathbb{R}^m, \\ &\Theta \begin{pmatrix} x \\ \mu \end{pmatrix} = x. \end{aligned}$$

Dann ist die Einschränkung von Θ auf $\widehat{M}_1(t)$ eine Bijektion zwischen $\widehat{M}_1(t)$ und $M_1(t)$. Das bedeutet insbesondere, dass wenn $\hat{x}(\cdot) = \begin{pmatrix} x(\cdot) \\ \mu(\cdot) \end{pmatrix}$ eine $\hat{C}^1_{\widehat{N}}$-Lösung von (4.14) ist, $x(\cdot)$ eine C^1_N-Lösung von (4.8) darstellt.

4.2.2 Analytische und diskrete Stabilität der GGL-Formulierung

Bis hier wurden Eigenschaften des GGL-Ansatzes gezeigt, die gewisse Hinweise auf eine mögliche Verbesserung der asymptotischen Stabilität für die Diskretisierungsverfahren geben. In diesem Abschnitt werden wichtige Aussagen bezüglich der Kontraktivität des GGL-Ansatzes bewiesen, die positive Auswirkungen aus numerischer Sicht haben.

Jetzt stellt sich die Frage, ob sich die Kontraktivität der Gleichung (4.8) auf die entsprechende GGL-Formulierung (4.14) überträgt.

Die Kontraktivitätsdefinition 2.2.10 nimmt für die lineare Gleichung (4.8) die folgende Form an:

Definition 4.2.7 *Die lineare ADGl*

$$A(t)x'(t) + B(t)x(t) = q(t), \quad t \in \mathfrak{I}, \tag{4.15}$$

mit konstantem Nullraum $N := \ker A(t)$ heißt kontraktiv auf $M_1(t) \subset \mathbb{R}^m$, wenn ein Skalarprodukt $\langle \cdot, \cdot \rangle_S$ und eine Konstante $c > 0$ existieren, so dass für alle Vektoren $(y, x, t) \in \mathbb{R}^m \times M_1(t) \times [t_0, \infty)$ mit

$$\begin{aligned} Qy &= 0, \\ A(t)y + B(t)x &= 0, \end{aligned}$$

die Ungleichung

$$\langle y, x \rangle_S \leq -c \, \|x\|_S^2$$

gilt, wobei der Projektor Q orthogonal nach dem Skalarprodukt $\langle \cdot, \cdot \rangle_S$ ist.

Theorem 4.2.8 *Falls die lineare ADGl (4.15) (P-)kontraktiv auf $M_1(t)$ mit dem Skalarprodukt $\langle \cdot, \cdot \rangle_S$ ist, dann ist die GGL-Gleichung (4.14) (P-)kontraktiv auf $\widehat{M}_1(t)$ mit dem Skalarprodukt $\langle \cdot, \cdot \rangle_{\widehat{S}}$, wobei*

$$\widehat{S} := \begin{pmatrix} S & 0 \\ 0 & I_\kappa \end{pmatrix}.$$

Beweis: M an betrachtet einen beliebigen Vektor $(\hat{y}, \hat{x}, t)$ mit

$$\hat{y} := \begin{pmatrix} y \\ y_\mu \end{pmatrix} \in I\!R^m \times I\!R^\kappa,\ \hat{x} := \begin{pmatrix} x \\ x_\mu \end{pmatrix} \in \widehat{M}_1(t) = M_1(t) \times \{0\}\,.$$

Außerdem werden in den euklidischen Räumen $\{I\!R^m, \langle\cdot,\cdot\rangle_S\}$ und $\left\{I\!R^m \times I\!R^\kappa, \langle\cdot,\cdot\rangle_{\widehat{S}}\right\}$ die Orthoprojektoren auf jeweils N und $\widehat{N}$, Q

$$\widehat{Q} = \begin{pmatrix} Q & 0 \\ 0 & I_\kappa \end{pmatrix},$$

gewählt.

Es wird angenommen, dass die Bedingungen

$$\begin{aligned} \widehat{Q}\hat{y} &= 0, \\ \widehat{A}(t)\hat{y} + \widehat{B}(t)\hat{x} &= 0, \end{aligned}$$

erfüllt sind. Wie man aus 4.2.1 weiß, bedeuten diese Gleichungen für den Vektor (y, x, t)

$$\begin{aligned} Qy &= 0, \\ A(t)y + B(t)x &= 0, \\ W_1(t)\left[(W_1B)' - W_1BA^+B\right](t)x &= 0, \end{aligned}$$

und zusätzlich

$$y_\mu = x_{\mu=}0.$$

Da nach der Annahme die Gleichung (4.15) kontraktiv bzw. (P-)kontraktiv auf $M_1(t)$ ist, folgt:

$$\langle y, x\rangle_S \leq -c\,\|x\|_S^2\,,$$

bzw.

$$\langle y, Px\rangle_S \leq -c\,\|Px\|_S^2\,,$$

für eine Konstante $c > 0$.

Schließlich gilt für den Vektor $(\hat{y}, \hat{x}, t)$

$$\langle \hat{y}, \hat{x}\rangle_{\widehat{S}} = \langle y, x\rangle_S + \langle 0, 0\rangle_{I_\kappa} \leq -c\,\|x\|_S^2 = -c\,\|\hat{x}\|_{\widehat{S}}^2\,,$$

bzw.

$$\left\langle \hat{y}, \widehat{P}\hat{x}\right\rangle_{\widehat{S}} = \langle y, Px\rangle_S + \langle 0, 0\rangle_{I_\kappa} \leq -c\,\|Px\|_S^2 = -c\left\|\widehat{P}\hat{x}\right\|_{\widehat{S}}^2,$$

und damit ist die (P-)Kontraktivität von (4.14) auf $\widehat{M}_1(t)$ bewiesen.
Dieses Ergebnis schafft eine wichtige Voraussetzung für einen Erfolg des GGL-Ansatzes aus numerischer Sicht, es ist aber nach den im Kapitel 3 gewonnenen Erkenntnissen nicht ausreichend. Das Hauptproblem besteht darin, dass die numerische Lösung nicht immer auf der Lösungsmannigfaltigkeit liegt, was dann die Kontraktivität ausschließlich auf der Lösungsmannigfaltigkeit nutzlos macht. Deswegen ist es viel wichtiger sicher zu stellen, ob Kontraktivität auf $\widehat{M}_0(t)$ vorliegt.

Man kann folgende Aussage beweisen.

Theorem 4.2.9 *Sei die lineare ADGl (4.15) P-kontraktiv (kontraktiv) auf $M_1(t)$ mit dem Skalarprodukt $\langle\cdot,\cdot\rangle_S$. Es bezeichne $S_1^{\perp}$ das Orthokomplement von S_1 bezüglich $\langle\cdot,\cdot\rangle_S$. Wenn $\Psi(t)$ in der GGL-Formulierung (4.14) eine Basis von $S_1^{\perp}$ darstellt, dann ist diese GGL-Gleichung P-kontraktiv (schwach kontraktiv) auf $\widehat{M}_0(t)$ mit dem Skalarprodukt $\langle\cdot,\cdot\rangle_{\widehat{S}}$, wobei*

$$\widehat{S} := \begin{pmatrix} S & 0 \\ 0 & I_\kappa \end{pmatrix}.$$

Beweis: M an betrachtet einen beliebigen Vektor $(\hat{y}, \hat{x}, t)$ mit

$$\hat{y} := \begin{pmatrix} y \\ y_\mu \end{pmatrix} \in I\!R^m \times I\!R^\kappa,\ \hat{x} := \begin{pmatrix} x \\ x_\mu \end{pmatrix} \in \widehat{M}_0(t) = M_1(t) \times I\!R^\kappa.$$

In den euklidischen Räumen $\{I\!R^m, \langle\cdot,\cdot\rangle_S\}$ und $\left\{I\!R^m \times I\!R^\kappa, \langle\cdot,\cdot\rangle_{\widehat{S}}\right\}$ werden die Orthoprojektoren Q und

$$\widehat{Q} = \begin{pmatrix} Q & 0 \\ 0 & I_\kappa \end{pmatrix}$$

auf jeweils N und $\widehat{N}$ gewählt und seien die Bedingungen

$$\begin{aligned} \widehat{Q}\hat{y} &= 0, \\ \widehat{A}(t)\hat{y} + \widehat{B}(t)\hat{x} &= 0, \end{aligned}$$

erfüllt. Wie man aus 4.2.1 weiß, bedeuten diese Gleichungen für den Vektor (y, x, t)

$$\begin{aligned} Qy &= 0, \\ A(t)(y - \Psi(t)x_\mu) + B(t)x &= 0, \\ W_1(t)\left[(W_1B)' - W_1BA^+B\right](t)x &= 0, \\ (\Gamma^T B)(t)x &= 0, \end{aligned}$$

und zusätzlich

$$y_\mu = 0.$$

Der Vektor $\Psi(t)x_\mu$ gehört zu im P, weil

$$\Psi(t)x_\mu \in S_1^\perp(t) \subset N^\perp,$$

und damit folgt $Q(y - \Psi(t)x_\mu) = 0$. Falls P-Kontraktivität für die Gleichung (4.15) auf $M_1(t)$ vorliegt, gilt

$$\langle y, Px\rangle_S = \langle y - \Psi(t)x_\mu, Px\rangle_S \leq -c\,\|Px\|_S^2$$

für eine Konstante $c \geq 0$, weil $Px \in S_1(t)$ und $\Psi(t)x_\mu \in S_1^\perp(t)$.

Schließlich gilt für den Vektor $(\hat{y}, \hat{x}, t)$

$$\left\langle \hat{y}, \widehat{P}\hat{x}\right\rangle_{\widehat{S}} = \langle y, Px\rangle_S + \langle 0, 0\rangle_{I_\kappa} \leq -c\,\|Px\|_S^2 = -c\left\|\widehat{P}\hat{x}\right\|_{\widehat{S}}^2,$$

und damit ist die P-Kontraktivität von (4.14) auf $\widehat{M}_0(t)$ gezeigt.

Wenn Kontraktivität auf $M_1(t)$ vorliegt, dann gilt

$$\langle y, x\rangle_S = \langle y - \Psi(t)x_\mu, x\rangle_S \leq -c\,\|x\|_S^2,$$

weil $\Psi(t)x_\mu \in N^\perp$. Letztendlich erhält man für $(\hat{y}, \hat{x}, t)$

$$\langle \hat{y}, \hat{x}\rangle_{\widehat{S}} = \langle y, x\rangle_S + \langle 0, x_\mu\rangle_{I_\kappa} \leq -c\,\|x\|_S^2 \leq 0,$$

was die schwache Kontraktivität von (4.14) auf $\widehat{M}_0(t)$ bedeutet.

Bemerkung 4.2.10 *Die Voraussetzungen des Theorems 4.2.9 werden in dem Hessenberg-Fall klarer. Wie schon gezeigt wurde, ist der Raum $S_1(t)$ in diesem Fall durch*

$$S_1(t) = \left\{ \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} : B_{21}(t)z_1 = 0 \right\}$$

gegeben. Jetzt entnimmt man aus (4.4)

$$\Psi(t) = \begin{pmatrix} B_{21}^T(t) \\ 0 \end{pmatrix},$$

dessen Bild offensichtlich der Orthogonalraum zu $S_1(t)$ nach dem euklidischen Skalarprodukt ist.

Das ist natürlich eine für die numerische asymptotische Stabilität bedeutende Aussage. Nach dem Theorem von Kapitel 2 gilt Folgendes:

Theorem 4.2.11 *Seien die Voraussetzungen von Theorem 4.2.9 erfüllt. Ein algebraisch stabiles IRK(DAE) angewandt auf den GGL-Ansatz (4.14) ist P-kontraktiv.*

Beweis: L aut Theorem 4.2.9 ist die GGL-Formulierung (4.14) P- bzw. schwach kontraktiv auf $\widehat{M}_0(t)$. Folglich ist dieser Beweis nach Theorem 2.3.5 erbracht.
Ein wichtiger Punkt sowohl der letzten Aussage als auch des Theorems 4.2.9 ist, dass sie nur für eine der vielen möglichen GGL-Formulierungen gelten. In der Tat ist vorausgesetzt worden, dass $\operatorname{im}\Psi(t) = S_1^{\perp}(t)$ in dem Raum $\{I\!R^m, \langle\cdot,\cdot\rangle_S\}$. Jedoch gibt es Indizien, dass diese Orthogonalitätsbedingung entbehrlich sein könnte. Dafür sprechen numerische Experimente wie das folgende:

Beispiel 4.2.12 *Man betrachtet erneut das Beispiel 3.0.8 und dafür die folgende GGL-Formulierung:*

$$\begin{pmatrix} x_1' \\ x_2' \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \lambda & -1 & -1 & \alpha_1 \\ \eta t(1-\eta t)-\eta & \lambda & -\eta t & \alpha_2 \\ \eta t(\eta t-1) & 1-\eta t & 1 & 0 \\ 1-\eta t & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \mu \end{pmatrix} = 0,$$

wobei

$$\begin{aligned} \alpha_1 &:= \cos(|\cos(t)|\,\theta)(1-\eta t) - \sin(|\cos(t)|\,\theta), \\ \alpha_2 &:= \sin(|\cos(t)|\,\theta)(1-\eta t) - \cos(|\cos(t)|\,\theta). \end{aligned}$$

Anstatt $B_{21}^T(t)$ als Koeffizienten von μ in (4.4) zu nehmen, wurde eine Drehung davon um einen variablen Winkel, der zwischen 0 und θ liegt, gewählt.

Es wurde eine analoge Berechnung, wie am Anfang dieses Kapitels für den Standard-GGL-Ansatz vorgestellt, durchgeführt und man gelangt zu ähnlichen Ergebnissen wie die Abbildung 4.5 zeigt.

4.3 GGL-Stabilitätserhaltung

In 4.2.2 erzielte man ein Stabilitätsergebnis für die Anwendung eines IRK(DAE)-Verfahrens auf die GGL-Formulierung unter der Annahme, dass die ursprüngliche

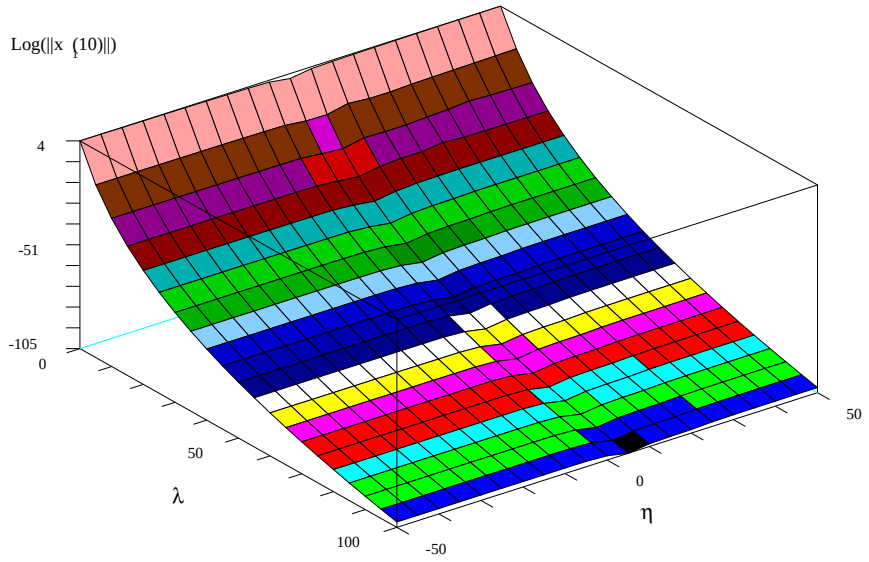

Abbildung 4.5: Numerische Ergebnisse für das Beispiel 3.0.8 unter Verwendung eines nicht-orthogonalen GGL-Ansatzes und des impliziten Euler-Verfahrens. Das Bild zeigt den Logarithmus des Betrages von $x_1(10)$ für verschiedene Werte der Parameter λ und η. Die Schrittweite betrug $h = 0.1$ und $\theta = \pi/3$.

Aufgabe P-kontraktiv ist. Jetzt wird eine Analyse unter schwächeren Voraussetzungen durchgeführt, die zusätzlich nicht nur für IRK-Verfahren anwendbar ist. Es wird der gleichen Idee wie in Kapitel 3 gefolgt. Man wird wieder ein Kommutativitätsdiagramm für die GGL-Formulierung betrachten, aber jetzt mit einer etwas anderen Bedeutung. Man kann leider nicht zeigen, dass beispielsweise Kommutativität zwischen Entkopplung und Diskretisierung gilt. Stattdessen wird bewiesen, dass bezüglich des Verfahrensstabilitätsverhaltens Kommutativität vorliegt, wenn man wie im Theorem 4.2.9 die GGL-Formulierung adäquat wählt. Das Diagramm 4.6 veranschaulicht diese Situation. Mit der GGL-Formulierung als Ausgangspunkt (links oben im dem Diagramm) kann man zwei Wege beschreiten, die sich nur

durch die Reihenfolge, in der die Schritte Entkopplung und Diskretisierung durchgeführt werden, unterscheiden. Die Aussage ist, dass, wenn der Weg über die GGL-IGDGL (inhärente gewöhnliche Differentialgleichung) zu einem kontraktiven Verfahren führt, der Weg über die ΔGGL (diskretisierte GGL-Formulierung) auch ein kontraktives Verfahren ergibt.

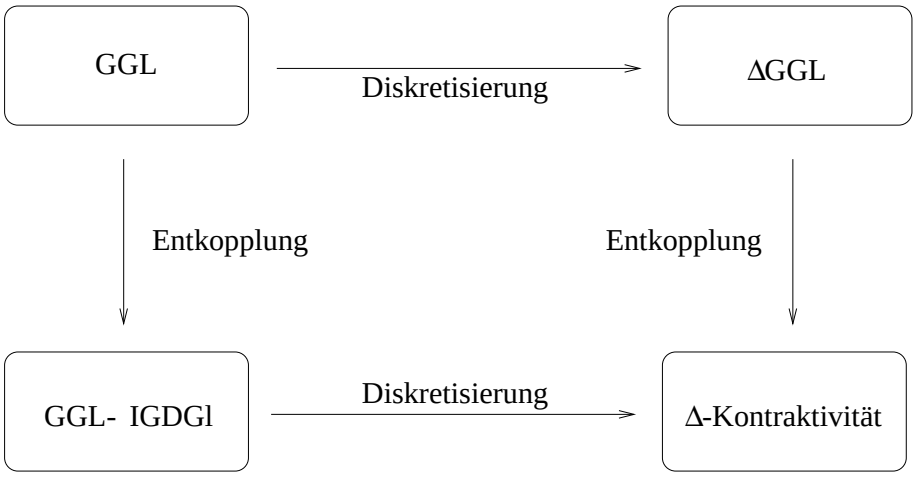

Abbildung 4.6: Kommutativität bei der GGL-Formulierung zwischen Entkopplung und Diskretisierung bezüglich der Kontraktivität des Verfahrens.

4.3.1 Entkopplung des GGL-Ansatzes

Nun wird die kanonische Entkopplung (also die mit dem Projektor auf $\widehat{N}_1(t)$ $Q_{\widehat{N}_1}^{\widehat{S}_1}(t)$) für die GGL-Formulierung (4.14) berechnet. Als Projektor auf $\widehat{N}_1(t)$ kann man

$$\widehat{Q}_{\widehat{N}_1}^*(t) = \begin{pmatrix} 0 & P\Psi(t) \\ 0 & I_\kappa \end{pmatrix}$$

wählen, dann ist

$$\begin{aligned} \widehat{B}\widehat{P}\widehat{Q}_{\widehat{N}_1}^* &= \begin{pmatrix} (I - W_1)B + W_1\left[(W_1B)' - W_1BA^+B\right] & -A\Psi \\ \Gamma^T B & 0 \end{pmatrix} \\ &\quad \times \begin{pmatrix} P & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & P\Psi \\ 0 & I_\kappa \end{pmatrix} \end{aligned}$$

$$
\begin{aligned}
&= \begin{pmatrix} (I-W_1)BP + W_1\left[(W_1B)' - W_1BA^+B\right]P & 0 \\ \Gamma^T BP & 0 \end{pmatrix}
\begin{pmatrix} 0 & P\Psi \\ 0 & I_\kappa \end{pmatrix} \\
&= \begin{pmatrix} 0 & \left\{(I-W_1)B + W_1\left[(W_1B)' - W_1BA^+B\right]\right\}P\Psi \\ 0 & \Gamma^T B\Psi \end{pmatrix},
\end{aligned}
$$

und

$$
\begin{aligned}
\widehat{G}_{2,*} &= \widehat{G}_1 + \widehat{B}\widehat{P}\widehat{Q}^*_{\widehat{N}_1} \\
&= \begin{pmatrix} G_1 - W_1BA^+BQ & -A\Psi \\ 0 & 0 \end{pmatrix} \\
&\quad + \begin{pmatrix} 0 & \left\{(I-W_1)B + W_1\left[(W_1B)' - W_1BA^+B\right]\right\}P\Psi \\ 0 & \Gamma^T B\Psi \end{pmatrix} \\
&= \begin{pmatrix} G_1 - W_1BA^+BQ & \left\{(I-W_1)B + W_1\left[(W_1B)' - W_1BA^+B\right]\right\}P\Psi - A\Psi \\ 0 & \Gamma^T B\Psi \end{pmatrix}.
\end{aligned}
$$

Die Inverse von $\widehat{G}_{2,*}$ ist durch

$$
\widehat{G}^{-1}_{2,*} = \begin{pmatrix} (G_1 - W_1BA^+BQ)^{-1} & -(G_1 - W_1BA^+BQ)^{-1}(\widehat{B}_{11} - A)P\Psi(\Gamma^T B\Psi)^{-1} \\ 0 & (\Gamma^T B\Psi)^{-1} \end{pmatrix}
$$

gegeben, wobei $\widehat{B}_{11}$ als $\widehat{B}_{11} := (I-W_1)B + W_1\left[(W_1B)' - W_1BA^+B\right]$ definiert ist. Der kanonische Projektor $\widehat{Q}^{\widehat{S}_1}_{\widehat{N}_1}$ ist dann

$$
\begin{aligned}
\widehat{Q}^{\widehat{S}_1}_{\widehat{N}_1} &= \widehat{Q}^*_{\widehat{N}_1}\widehat{G}^{-1}_{2,*}\widehat{B}\widehat{P} \\
&= \begin{pmatrix} 0 & P\Psi \\ 0 & I_\kappa \end{pmatrix} \\
&\quad \times \begin{pmatrix} (G_1 - W_1BA^+BQ)^{-1} & -(G_1 - W_1BA^+BQ)^{-1}(\widehat{B}_{11} - A)P\Psi(\Gamma^T B\Psi)^{-1} \\ 0 & (\Gamma^T B\Psi)^{-1} \end{pmatrix} \\
&\quad \times \begin{pmatrix} \widehat{B}_{11}P & 0 \\ \Gamma^T B & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & P\Psi(\Gamma^T B\Psi)^{-1} \\ 0 & (\Gamma^T B\Psi)^{-1} \end{pmatrix}
\begin{pmatrix} \widehat{B}_{11}P & 0 \\ \Gamma^T B & 0 \end{pmatrix} \\
&= \begin{pmatrix} P\Psi(\Gamma^T B\Psi)^{-1}\Gamma^T B & 0 \\ (\Gamma^T B\Psi)^{-1}\Gamma^T B & 0 \end{pmatrix} = \begin{pmatrix} PQ^{S_1}_{S_1^t}(t) & 0 \\ (\Gamma^T B\Psi)^{-1}\Gamma^T B & 0 \end{pmatrix},
\end{aligned}
$$

wobei $S_1^t(t) := \operatorname{im}\Psi(t)$, siehe (4.12), (4.13).

Man beachte, dass

$$\widehat{P}\widehat{Q}^{\widehat{S}_1}_{\widehat{N}_1}(t) = \begin{pmatrix} PQ^{S_1}_{S_1^t}(t) & 0 \\ 0 & 0 \end{pmatrix}, \quad \widehat{P}\widehat{P}^{\widehat{N}_1}_{\widehat{S}_1}(t) = \begin{pmatrix} PP^{S_1^t}_{S_1}(t) & 0 \\ 0 & 0 \end{pmatrix}$$

gilt.

Es soll als Nächstes die kanonische Entkopplung für (4.14) berechnet werden, jedoch ist dafür die Matrix $\widehat{G}_2^{-1}$ notwendig:

$$\begin{aligned}
\widehat{G}_2 &= \widehat{G}_1 + \widehat{B}\widehat{P}\widehat{Q}^{\widehat{S}_1}_{\widehat{N}_1} \\
&= \begin{pmatrix} G_1 - W_1BA^+BQ & -A\Psi \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} \widehat{B}_{11}P & 0 \\ \Gamma^TB & 0 \end{pmatrix} \begin{pmatrix} PQ^{S_1}_{S_1^t} & 0 \\ (\Gamma^TB\Psi)^{-1}\Gamma^TB & 0 \end{pmatrix} \\
&= \begin{pmatrix} G_1 - W_1BA^+BQ & -A\Psi \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} \widehat{B}_{11}PQ^{S_1}_{S_1^t} & 0 \\ \Gamma^TB & 0 \end{pmatrix} \\
&= \begin{pmatrix} G_1 - W_1BA^+BQ + \widehat{B}_{11}PQ^{S_1}_{S_1^t} & -A\Psi \\ \Gamma^TB & 0 \end{pmatrix}.
\end{aligned}$$

Proposition 4.3.1 *Die Inverse von $\widehat{G}_2$ ist durch*

$$\begin{pmatrix} (I - PQ^{S_1}_{S_1^t})\Upsilon^{-1} & \left[I - (I - PQ^{S_1}_{S_1^t})\Upsilon^{-1}\widehat{B}_{11}\right] P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1} \\ -(\widehat{B}_{21}\widehat{B}_{12})^{-1}\widehat{B}_{21}\Upsilon^{-1} & (\widehat{B}_{21}\widehat{B}_{12})^{-1}\left[I + \widehat{B}_{21}\Upsilon^{-1}\widehat{B}_{11}P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}\right] \end{pmatrix}$$

gegeben, wobei $\Upsilon := G_1 - W_1BA^+BQ$, $\widehat{B}_{21} := \Gamma^TB$ und $\widehat{B}_{12} := \Psi$.

Beweis: M an betrachtet das Gleichungssystem

$$\widehat{G}_2\hat{x} = \hat{b}, \quad \hat{x} := \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \hat{b} := \begin{pmatrix} b_1 \\ b_2 \end{pmatrix},$$

also

$$\begin{aligned}
(G_1 - W_1BA^+BQ + \widehat{B}_{11}PQ^{S_1}_{S_1^t})x_1 - A\widehat{B}_{12}x_2 &= b_1 \\
\widehat{B}_{21}x_1 &= b_2.
\end{aligned}$$

Wenn man die zweite Gleichung mit $\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}$ multipliziert, folgt

$$Q^{S_1}_{S_1^t}x_1 = \widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}b_2.$$

So kann die erste Gleichung des Systems als

$$\Upsilon x_1 = b_1 + A\widehat{B}_{12}x_2 - \widehat{B}_{11}P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}b_2$$

geschrieben werden, und daraus folgt, dass

$$x_1 = \Upsilon^{-1}\left[b_1 - \widehat{B}_{11}P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}b_2\right] + P\widehat{B}_{12}x_2.$$

Jetzt können die anderen Komponenten von x_1 berechnet werden

$$\begin{aligned} PP_{S_1}^{S_1^t}x_1 &= PP_{S_1}^{S_1^t}\Upsilon^{-1}\left[b_1 - \widehat{B}_{11}P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}b_2\right], \\ Qx_1 &= Q\Upsilon^{-1}\left[b_1 - \widehat{B}_{11}P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}b_2\right], \end{aligned}$$

und für x_1 ergibt sich

$$\begin{aligned} x_1 &= P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}b_2 + (I - PQ_{S_1^t}^{S_1})\Upsilon^{-1}\left[b_1 - \widehat{B}_{11}P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}b_2\right] \\ &= (I - PQ_{S_1^t}^{S_1})\Upsilon^{-1}b_1 + \left[I - (I - PQ_{S_1^t}^{S_1})\Upsilon^{-1}\widehat{B}_{11}\right]P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}b_2. \end{aligned}$$

Nun gilt für x_2 die Gleichung

$$A\widehat{B}_{12}x_2 = (G_1 - W_1BA^+BQ + \widehat{B}_{11}PQ_{S_1^t}^{S_1})x_1 - b_1.$$

Man multipliziert die letzte Gleichung mit $\widehat{B}_{21}(G_1 - W_1BA^+BQ)^{-1}$ und bekommt

$$\begin{aligned} \widehat{B}_{21}\widehat{B}_{12}x_2 &= \widehat{B}_{21}x_1 + \widehat{B}_{21}\Upsilon^{-1}(\widehat{B}_{11}PQ_{S_1^t}^{S_1}x_1 - b_1) \\ &= \left[I + \widehat{B}_{21}\Upsilon^{-1}\widehat{B}_{11}P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}\right]\widehat{B}_{21}x_1 - \widehat{B}_{21}\Upsilon^{-1}b_1. \end{aligned}$$

Nachdem man den Ausdruck von x_1 eingesetzt hat, folgt

$$\begin{aligned} \widehat{B}_{21}\widehat{B}_{12}x_2 &= \left[I + \widehat{B}_{21}\Upsilon^{-1}\widehat{B}_{11}P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}\right]b_2 - \widehat{B}_{21}\Upsilon^{-1}b_1, \\ x_2 &= (\widehat{B}_{21}\widehat{B}_{12})^{-1}\left[I + \widehat{B}_{21}\Upsilon^{-1}\widehat{B}_{11}P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}\right]b_2 - (\widehat{B}_{21}\widehat{B}_{12})^{-1}\widehat{B}_{21}\Upsilon^{-1}b_1, \end{aligned}$$

und schließlich erhält man für $\widehat{G}_2^{-1}$ den Ausdruck

$$\begin{pmatrix} (I - PQ_{S_1^t}^{S_1})\Upsilon^{-1} & \left[I - (I - PQ_{S_1^t}^{S_1})\Upsilon^{-1}\widehat{B}_{11}\right]P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1} \\ -(\widehat{B}_{21}\widehat{B}_{12})^{-1}\widehat{B}_{21}\Upsilon^{-1} & (\widehat{B}_{21}\widehat{B}_{12})^{-1}\left[I + \widehat{B}_{21}\Upsilon^{-1}\widehat{B}_{11}P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}\right] \end{pmatrix}.$$

Für die Berechnung der Entkopplung muss man auch den Ausdruck von $\Upsilon^{-1} = (G_1 - W_1BA^+BQ)^{-1}$ genauer kennen.

Proposition 4.3.2 *Die Inverse von* $(G_1 - W_1BA^+BQ)$ *ist durch*

$$(G_2^{-1} + Q_{N_1}^{S_1}A^+)(I - W_1) + Q_{N_1}^{S_1}(W_1B)^+W_1$$

gegeben.

Beweis: W ie in dem letzten Beweis betrachtet man ein lineares Gleichungssystem für diese Matrix

$$(G_1 - W_1BA^+BQ)x = b.$$

Das System wird in die Komponenten $I - W_1$ und W_1 gesplittet,

$$\begin{aligned} G_1x &= (I - W_1)b, \\ -W_1BA^+BQx &= W_1b, \end{aligned}$$

und aus der vorletzten Gleichung folgt

$$G_1x = G_1P_{S_1}^{N_1}x = G_2P_{S_1}^{N_1}x = (I - W_1)b,$$

$$P_{1,c}x = G_{2,c}^{-1}(I - W_1)b.$$

Um die $Q_{N_1}^{S_1}$-Komponente zu berechnen, geht man folgendermaßen vor: Man multipliziert die Gleichung

$$G_1x = (I - W_1)b$$

mit W_1BA^+ und erhält

$$\begin{aligned} W_1BA^+G_1x &= W_1BA^+(I - W_1)b, \\ W_1BA^+Ax + W_1BA^+BQx &= W_1BA^+(I - W_1)b, \\ W_1Bx &= W_1b + W_1BA^+(I - W_1)b. \end{aligned}$$

Weil $Q_{N_1}^{S_1}$ und $Q_{S_1^\perp}^{S_1} = (W_1B)^+(W_1B)$ zwei Projektoren längs S_1 sind, folgt

$$\begin{aligned} Q_{N_1}^{S_1}x &= Q_{N_1}^{S_1}(W_1B)^+W_1b + Q_{1,c}A^+(I - W_1)b, \\ &= Q_{N_1}^{S_1}\left[(W_1B)^+W_1 + A^+(I - W_1)\right]b, \end{aligned}$$

und schließlich

$$\begin{aligned} x &= G_2^{-1}(I-W_1)b + Q_{N_1}^{S_1}\left[(W_1B)^+W_1 + A^+(I-W_1)\right]b \\ &= (G_2^{-1} + Q_{N_1}^{S_1}A^+)(I-W_1)b + Q_{N_1}^{S_1}(W_1B)^+W_1b. \end{aligned}$$

Jetzt ist man in der Lage, die kanonische Entkopplung für (4.14) zu berechnen. Im Einklang mit 3.1.1 ist diese Entkopplung

$$\hat{u}' - (\widehat{P}\widehat{P}^{\widehat{N}_1}_{\widehat{S}_1})'(\hat{u}+\widehat{P}\hat{y}) + \widehat{P}\widehat{P}^{\widehat{N}_1}_{\widehat{S}_1}\widehat{G}^{-1}_{2,c}\widehat{B}\hat{u} = \widehat{P}\widehat{P}^{\widehat{N}_1}_{\widehat{S}_1}\widehat{G}_2^{-1}\hat{q}, \quad (4.16)$$

$$-\widehat{Q}\widehat{Q}^{\widehat{S}_1}_{\widehat{N}_1}(\widehat{P}\hat{y})' + (\widehat{Q}\widehat{Q}^{\widehat{S}_1}_{\widehat{N}_1})'\hat{u} + \widehat{T}\widehat{Q}\widehat{P}^{\widehat{N}_1}_{\widehat{S}_1}\widehat{G}_2^{-1}\widehat{B}\hat{u} + \hat{w} = \widehat{T}\widehat{Q}\widehat{P}^{\widehat{N}_1}_{\widehat{S}_1}\widehat{G}_2^{-1}\hat{q}, \quad (4.17)$$

$$\hat{y} + \widehat{U}\widehat{Q}\widehat{G}_2^{-1}\widehat{B}\hat{u} = (\widehat{U}\widehat{Q} + \widehat{P}\widehat{Q}^{\widehat{S}_1}_{\widehat{N}_1})\widehat{G}_2^{-1}\hat{q}. \quad (4.18)$$

Es werden zuerst die einzelnen Elemente berechnet. Die Variablen $\hat{u}$, $\hat{y}$ und $\hat{w}$ sind hier

$$\begin{aligned} \hat{u} &= \widehat{P}\widehat{P}^{\widehat{N}_1}_{\widehat{S}_1}\hat{x} = \begin{pmatrix} PP^{S_1^t}_{S_1} & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ \mu \end{pmatrix} = \begin{pmatrix} PP^{S_1^t}_{S_1}x \\ 0 \end{pmatrix} = \begin{pmatrix} z \\ 0 \end{pmatrix}, \\ \hat{y} &= (\widehat{U}\widehat{Q} + \widehat{P}\widehat{Q}^{\widehat{S}_1}_{\widehat{N}_1})\hat{x} = \left[\begin{pmatrix} I_m & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} Q & 0 \\ 0 & I_\kappa \end{pmatrix} + \begin{pmatrix} PQ^{S_1}_{S_1^t} & 0 \\ 0 & 0 \end{pmatrix}\right]\begin{pmatrix} x \\ \mu \end{pmatrix} \\ &= \begin{pmatrix} Q + PQ^{S_1}_{S_1^t} & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ \mu \end{pmatrix} = \begin{pmatrix} (Q+PQ^{S_1}_{S_1^t})x \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} (Q+PQ^{S_1}_{S_1^t})(Q+PQ^{S_1}_{N_1})x \\ 0 \end{pmatrix} = \begin{pmatrix} (Q+PQ^{S_1}_{S_1^t})(y+w) \\ 0 \end{pmatrix} \\ \hat{w} &= \widehat{T}\widehat{Q}\hat{x} = \begin{pmatrix} 0 & 0 \\ 0 & I_\kappa \end{pmatrix}\begin{pmatrix} Q & 0 \\ 0 & I_\kappa \end{pmatrix}\begin{pmatrix} x \\ \mu \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & I_\kappa \end{pmatrix}\begin{pmatrix} x \\ \mu \end{pmatrix} = \begin{pmatrix} 0 \\ \mu \end{pmatrix}. \end{aligned}$$

Es wird mit der Gleichung (4.18) angefangen. Dafür benötigt man

$$\widehat{U}\widehat{Q}\widehat{G}_2^{-1} = \begin{pmatrix} Q & 0 \\ 0 & 0 \end{pmatrix}\widehat{G}_2^{-1}$$

$$\begin{aligned} &= \begin{pmatrix} Q\Upsilon^{-1} & Q\left[I - (I-PQ^{S_1}_{S_1^t})\Upsilon^{-1}\widehat{B}_{11}\right]P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1} \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} Q\Upsilon^{-1} & -Q\Upsilon^{-1}\widehat{B}_{11}P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1} \\ 0 & 0 \end{pmatrix}, \end{aligned}$$

dann folgt

$$\begin{aligned}
\widehat{U}\widehat{Q}\widehat{G}_2^{-1}\widehat{B}\hat{u} &= \begin{pmatrix} Q\Upsilon^{-1} & -Q\Upsilon^{-1}\widehat{B}_{11}P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \widehat{B}_{11} & -A\widehat{B}_{12} \\ \widehat{B}_{21} & 0 \end{pmatrix} \begin{pmatrix} z \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} Q\Upsilon & -Q\Upsilon^{-1}\widehat{B}_{11}P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \widehat{B}_{11}z \\ \widehat{B}_{21}z \end{pmatrix} \\
&= \begin{pmatrix} Q\Upsilon^{-1}\widehat{B}_{11}(I - PQ_{S_1^t}^{S_1})z \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} Q\Upsilon^{-1}\widehat{B}_{11}z \\ 0 \end{pmatrix}.
\end{aligned}$$

Die erste Komponente dieses Vektors ist

$$\begin{aligned}
& Q\Upsilon^{-1}\widehat{B}_{11}z \\
= \; & Q\left\{(G_2^{-1} + Q_{N_1}^{S_1}A^+)(I - W_1)B + Q_{N_1}^{S_1}(W_1B)^+W_1\left[(W_1B)' - W_1BA^+B\right]\right\}z \\
= \; & Q\left[G_2^{-1}B + Q_{N_1}^{S_1}A^+B + Q_{N_1}^{S_1}(W_1B)^+W_1(W_1B)' - Q_{N_1}^{S_1}A^+B\right]z \\
= \; & Q\left[G_2^{-1}B + Q_{N_1}^{S_1}Q_{S_1^t}^{S_{1'}} - Q_{N_1}^{S_1}\left[Q_{S_1^t}^{S_1}(W_1B)^+W_1\right]'W_1B\right]z \\
= \; & Q(G_2^{-1}B + Q_{N_1}^{S_1}Q_{S_1^t}^{S_{1'}})z,
\end{aligned}$$

und man erhält

$$\widehat{U}\widehat{Q}\widehat{G}_2^{-1}\widehat{B}\hat{u} = \begin{pmatrix} Q(G_2^{-1}B + Q_{N_1}^{S_1}Q_{S_1^t}^{S_{1'}})z \\ 0 \end{pmatrix}.$$

Für die rechte Seite von (4.18) müssen $\widehat{U}\widehat{Q}\widehat{G}_2^{-1}\hat{q}$ und $\widehat{P}Q_{N_1}^{S_1}\widehat{G}_2^{-1}\hat{q}$ berechnet werden. Für den ersten Ausdruck folgt

$$\begin{aligned}
& \widehat{U}\widehat{Q}\widehat{G}_2^{-1}\hat{q} = \begin{pmatrix} Q\Upsilon^{-1} & -Q\Upsilon^{-1}\widehat{B}_{11}P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1} \\ 0 & 0 \end{pmatrix} \\
& \times \begin{pmatrix} (I - W_1)q + W_1\left[(W_1q)' - W_1BA^+q\right] \\ \Gamma^Tq \end{pmatrix} \\
& = \begin{pmatrix} Q\Upsilon^{-1}\left\{(I - W_1)q + W_1\left[(W_1q)' - W_1BA^+q\right] - \widehat{B}_{11}P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}\Gamma^Tq\right\} \\ 0 \end{pmatrix},
\end{aligned}$$

wobei für den ersten Term gilt

$$Q\Upsilon^{-1}\left\{(I - W_1)q + W_1\left[(W_1q)' - W_1BA^+q\right]\right\}$$

$$
\begin{aligned}
&= Q\left\{(G_2^{-1}+Q_{N_1}^{S_1}A^+)(I-W_1)q+Q_{N_1}^{S_1}(W_1B)^+W_1\left[(W_1q)'-W_1BA^+q\right]\right\}\\
&= Q\left\{G_2^{-1}(I-W_1)q+Q_{N_1}^{S_1}(W_1B)^+W_1(W_1q)'\right\}\\
&= Q\left\{G_2^{-1}(I-W_1)q+Q_{N_1}^{S_1}\left[Q_{S_1^t}^{S_1}(W_1B)^+W_1q\right]'-Q_{N_1}^{S_1}\left[Q_{S_1^t}^{S_1}(W_1B)^+W_1\right]'W_1q\right\},
\end{aligned}
$$

und für den zweiten

$$
\begin{aligned}
&Q\Upsilon^{-1}\widehat{B}_{11}P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}\Gamma^Tq\\
&= Q\left\{(G_2^{-1}+Q_{N_1}^{S_1}A^+)(I-W_1)B+Q_{N_1}^{S_1}(W_1B)^+W_1\left[(W_1B)'-W_1BA^+B\right]\right\}\\
&\quad\times P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}\Gamma^Tq\\
&= Q\left\{G_2^{-1}(I-W_1)B+Q_{N_1}^{S_1}Q_{S_1^t}^{S_{1'}}-Q_{N_1}^{S_1}\left[Q_{S_1^t}^{S_1}(W_1B)^+W_1\right]'W_1B\right\}\\
&\quad\times P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}\Gamma^Tq\\
&= QG_2^{-1}BP\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}\Gamma^Tq-QG_2^{-1}W_1q+QQ_{N_1}^{S_1}Q_{S_1^t}^{S_{1'}}P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}\Gamma^Tq\\
&\quad-QQ_{N_1}^{S_1}\left[Q_{S_1^t}^{S_1}(W_1B)^+W_1\right]'W_1q.
\end{aligned}
$$

So ergibt sich insgesamt

$$
\begin{aligned}
&\widehat{U}\widehat{Q}\widehat{G}_2^{-1}\hat{q}\\
&= \begin{pmatrix} QG_2^{-1}q+QQ_{N_1}^{S_1}\left[Q_{S_1^t}^{S_1}(W_1B)^+W_1q\right]'-(QG_2^{-1}B+QQ_{N_1}^{S_1}Q_{S_1^t}^{S_{1'}})P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}\Gamma^Tq\\ 0\end{pmatrix}.
\end{aligned}
$$

Für den zweiten Term der rechten Seite $\widehat{P}\widehat{Q}_{\widehat{N}_1}^{\widehat{S}_1}\widehat{G}_2^{-1}\hat{q}$ gilt

$$
\begin{aligned}
&\widehat{P}\widehat{Q}_{\widehat{N}_1}^{\widehat{S}_1}\widehat{G}_2^{-1}\hat{q}\\
&= \begin{pmatrix} PQ_{S_1^t}^{S_1} & 0\\ 0 & 0\end{pmatrix}\\
&\quad\times\begin{pmatrix} (I-PQ_{S_1^t}^{S_1})\Upsilon^{-1} & \left[I-(I-PQ_{S_1^t}^{S_1})\Upsilon^{-1}\widehat{B}_{11}\right]P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}\\ -(\widehat{B}_{21}\widehat{B}_{12})^{-1}\widehat{B}_{21}\Upsilon^{-1} & (\widehat{B}_{21}\widehat{B}_{12})^{-1}\left[I+\widehat{B}_{21}\Upsilon^{-1}\widehat{B}_{11}P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}\right]\end{pmatrix}\\
&\quad\times\begin{pmatrix}(I-W_1)q+W_1\left[(W_1q)'-W_1BA^+q\right]\\ \Gamma^Tq\end{pmatrix}\\
&= \begin{pmatrix} 0 & PQ_{S_1^t}^{S_1}\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}\\ 0 & 0\end{pmatrix}\begin{pmatrix}(I-W_1)q+W_1\left[(W_1q)'-W_1BA^+q\right]\\ \Gamma^Tq\end{pmatrix}\\
&= \begin{pmatrix} P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}\Gamma^Tq\\ 0\end{pmatrix},
\end{aligned}
$$

und nun kann man die Gleichung (4.18) zusammensetzen. Man erhält

$$\begin{aligned}
&\begin{pmatrix} (Q+PQ_{S_1^t}^{S_1})(y+w) \\ 0 \end{pmatrix} + \begin{pmatrix} Q(G_2^{-1}B + Q_{N_1}^{S_1}Q_{S_1^t}^{S_{1'}})z \\ 0 \end{pmatrix} \\
= &\begin{pmatrix} QG_2^{-1}q + QQ_{N_1}^{S_1}\left[Q_{S_1^t}^{S_1}(W_1B)^+W_1q\right]' - (QG_2^{-1}B + QQ_{N_1}^{S_1}Q_{S_1^t}^{S_{1'}})P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}\Gamma^Tq \\ 0 \end{pmatrix} \\
&+ \begin{pmatrix} P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}\Gamma^Tq \\ 0 \end{pmatrix},
\end{aligned}$$

was die Gleichung

$$\begin{aligned}
&(Q+PQ_{S_1^t}^{S_1})(y+w) + Q(G_2^{-1}B + Q_{N_1}^{S_1}Q_{S_1^t}^{S_{1'}})z \qquad (4.19)\\
= \;& QG_2^{-1}q + QQ_{N_1}^{S_1}\left[Q_{S_1^t}^{S_1}(W_1B)^+W_1q\right]' - (QG_2^{-1}B + QQ_{N_1}^{S_1}Q_{S_1^t}^{S_{1'}} - I)\\
&\times P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}\Gamma^Tq
\end{aligned}$$

ergibt.

Wenn diese Gleichung mit $Q_{S_1^t}^{S_1}$ multipliziert wird, erhält man

$$Q_{S_1^t}^{S_1}y = Q_{S_1^t}^{S_1}x = \widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}\Gamma^Tq = Q_{S_1^t}^{S_1}(W_1B)^+W_1q,$$

was äquivalent zu

$$Q_{N_1}^{S_1}y = Q_{N_1}^{S_1}x = Q_{N_1}^{S_1}G_2^{-1}q$$

ist. Diesen Ausdruck für $Q_{S_1^t}^{S_1}x$, $Q_{N_1}^{S_1}x$ bekam man schon in 3.1.1, Proposition 3.1.8.

Wie unten gezeigt wird, enthält (4.19) die Gleichungen (3.14) und (3.15). Dies war zu erwarten, weil $\widehat{M}_0 = M_1 \times I\!R^k$. Mit anderen Worten ist die sichtbare Bedingung von (4.14) die Zusammensetzung der sichtbaren und versteckten Nebenbedingungen von (4.8).

Zuerst ist darauf aufmerksam zu machen, dass (4.19) im Raum im $(Q + PQ_{S_1^t}^{S_1})$ lebt und die Projektoren $Q + PQ_{S_1^t}^{S_1}$ und $Q + PQ_{N_1}^{S_1}$ den gleichen Nullraum besitzen. Dann ist das Produkt $(Q+PQ_{N_1}^{S_1})\times$(4.19) eine äquivalente Gleichung, die in die Teile $UQ + PQ_{N_1}^{S_1}$ und TQ zerlegt werden kann. Für $(Q + PQ_{N_1}^{S_1})\times$(4.19) erhält man

$$\begin{aligned}
&(Q+PQ_{N_1}^{S_1})(y+w) + Q(G_2^{-1}B + Q_{N_1}^{S_1}Q_{S_1^t}^{S_{1'}})z\\
= \;& QG_2^{-1}q + QQ_{N_1}^{S_1}\left[Q_{S_1^t}^{S_1}(W_1B)^+W_1q\right]' - (QG_2^{-1}B + QQ_{N_1}^{S_1}Q_{S_1^t}^{S_{1'}} - PQ_{N_1}^{S_1})P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}\Gamma^Tq,
\end{aligned}$$

und für den $(UQ + PQ_{N_1}^{S_1})$-Teil folgt

$$\begin{aligned}
y + UQG_2^{-1}Bz + UQG_2^{-1}BP\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}\Gamma^T q &= UQG_2^{-1}q + PQ_{N_1}^{S_1} \\
&\quad \times \widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}\Gamma^T q, \\
y + UQG_2^{-1}Bz + UQG_2^{-1}BPQ_{S_1^t}^{S_1} y &= UQG_2^{-1}q + PQ_{N_1}^{S_1} x, \\
(I + UQG_2^{-1}BPP_{S_1}^{N_1}Q_{S_1^t}^{S_1})y + UQG_2^{-1}BPP_{S_1}^{N_1} z &= (UQ + PQ_{N_1}^{S_1})G_2^{-1}q,
\end{aligned}$$

was genau die Gleichung für die $(UQ+PQ_{N_1}^{S_1})$-Komponente der Lösung (3.15) darstellt. Andererseits ist der TQ-Teil

$$\begin{aligned}
w + TQG_2^{-1}Bz + QQ_{N_1}^{S_1}Q_{S_1^t}^{S_{1'}} z - QQ_{N_1}^{S_1}(PQ_{S_1^t}^{S_1} y)' & \\
+QQ_{N_1}^{S_1}Q_{S_1^t}^{S_{1'}} PQ_{S_1^t}^{S_1} y &= TQG_2^{-1}q - TQG_2^{-1} \\
&\quad \times BPQ_{S_1^t}^{S_1} y, \\
w + TQG_{2,c}^{-1}BPP_{S_1}^{N_1} z + QQ_{N_1}^{S_1}Q_{S_1^t}^{S_{1'}} z - QQ_{N_1}^{S_1}(PQ_{S_1^t}^{S_1} y)' & \\
+QQ_{N_1}^{S_1}Q_{S_1^t}^{S_{1'}} PQ_{S_1^t}^{S_1} y &= TQP_{S_1}^{N_1}G_2^{-1}q + QQ_{N_1}^{S_1}G_2^{-1}q \\
&\quad -TQP_{S_1}^{N_1}G_2^{-1}BPQ_{S_1^t}^{S_1} y \\
&\quad -QQ_{N_1}^{S_1}Q_{S_1^t}^{S_1} y, \\
-QQ_{N_1}^{S_1}(PQ_{S_1^t}^{S_1} y)' + (QQ_{N_1}^{S_1}Q_{S_1^t}^{S_{1'}} + TQG_2^{-1}BPP_{S_1}^{N_1}) & \\
(z + PQ_{S_1^t}^{S_1} y) + w &= TQP_{S_1}^{N_1}G_2^{-1}q,
\end{aligned}$$

und das ist wiederum die Gleichung für die TQ-Komponente der Lösung (3.14).

Als Nächstes berechnet man die inhärente Differentialgleichung (4.16). Man benötigt hier

$$\begin{aligned}
&\widehat{P}\widehat{P}_{\widehat{S}_1}^{\widehat{N}_1}\widehat{G}_2^{-1} \\
&= \begin{pmatrix} PP_{S_1}^{S_1^t} & 0 \\ 0 & 0 \end{pmatrix} \widehat{G}_2^{-1} \\
&= \begin{pmatrix} PP_{S_1}^{S_1^t}(I - PQ_{S_1^t}^{S_1} y,)\Upsilon^{-1} & PP_{S_1}^{S_1^t}\left[I - (I - PQ_{S_1^t}^{S_1})\Upsilon^{-1}\widehat{B}_{11}\right] P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1} \\ 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} PP_{S_1}^{S_1^t}\Upsilon^{-1} & -PP_{S_1}^{S_1^t}\Upsilon^{-1}\widehat{B}_{11}P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1} \\ 0 & 0 \end{pmatrix}.
\end{aligned}$$

Infolgedessen erhält man

$$\begin{aligned}
\widehat{P}\widehat{P}^{\widehat{N}_1}_{\widehat{S}_1}\widehat{G}_2^{-1}\widehat{B}\hat{u} &= \begin{pmatrix} PP^{S_1^t}_{S_1}\Upsilon^{-1} & -PP^{S_1^t}_{S_1}\Upsilon^{-1}\widehat{B}_{11}P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1} \\ 0 & 0 \end{pmatrix} \\
&\quad \times \begin{pmatrix} (I-W_1)B + W_1\left[(W_1B)' - W_1BA^+B\right] & -A\widehat{B}_{12} \\ \widehat{B}_{21} & 0 \end{pmatrix} \begin{pmatrix} PP^{S_1^t}_{S_1}x \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} PP^{S_1^t}_{S_1}\Upsilon^{-1} & -\Upsilon^{-1}\widehat{B}_{11}P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1} \\ 0 & 0 \end{pmatrix} \\
&\quad \times \begin{pmatrix} BPP^{S_1^t}_{S_1}x + W_1\left[(W_1B)' - W_1BA^+B\right]PP^{S_1^t}_{S_1}x \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} PP^{S_1^t}_{S_1}\Upsilon^{-1}\left\{B + W_1\left[(W_1B)' - W_1BA^+B\right]\right\}PP^{S_1^t}_{S_1}x \\ 0 \end{pmatrix},
\end{aligned}$$

und die erste Komponente dieses Vektors ist

$$\begin{aligned}
& \Upsilon^{-1}\left\{B + W_1\left[(W_1B)' - W_1BA^+B\right]\right\}PP^{S_1^t}_{S_1}x \\
= \;& PP^{S_1^t}_{S_1}\left[(G_2^{-1} + Q^{S_1}_{N_1}A^+)(I-W_1) + Q^{S_1}_{N_1}(W_1B)^+W_1\right] \\
& \times\left\{B + W_1\left[(W_1B)' - W_1BA^+B\right]\right\}PP^{S_1^t}_{S_1}x \\
= \;& PP^{S_1^t}_{S_1}\left[(G_2^{-1} + Q^{S_1}_{N_1}A^+)(I-W_1) + Q^{S_1}_{N_1}(W_1B)^+W_1\right]BPP^{S_1^t}_{S_1}x \\
& + PP^{S_1^t}_{S_1}Q^{S_1}_{N_1}(W_1B)^+W_1\left[(W_1B)' - W_1BA^+B\right]PP^{S_1^t}_{S_1}x \\
= \;& PP^{S_1^t}_{S_1}(G_2^{-1} + Q^{S_1}_{N_1}A^+)BPP^{S_1^t}_{S_1}x + PP^{S_1^t}_{S_1}Q^{S_1}_{N_1}Q^{S_{1'}}_{S_1^t}PP^{S_1^t}_{S_1}x - PP^{S_1^t}_{S_1}Q^{S_1}_{N_1}A^+BPP^{S_1^t}_{S_1}x \\
= \;& PP^{S_1^t}_{S_1}G_2^{-1}BPP^{S_1^t}_{S_1}x + PP^{S_1^t}_{S_1}Q^{S_1}_{N_1}Q^{S_{1'}}_{S_1^t}PP^{S_1^t}_{S_1}x \\
= \;& PP^{N_1}_{S_1}G_2^{-1}BPP^{S_1^t}_{S_1}x - PP^{S_1^t}_{S_1}P^{N_1}_{S_1}Q^{S_{1'}}_{S_1^t}PP^{S_1^t}_{S_1}x \\
= \;& PP^{N_1}_{S_1}G_2^{-1}BPP^{S_1^t}_{S_1}x - PP^{N_1}_{S_1}Q^{S_{1'}}_{S_1^t}PP^{S_1^t}_{S_1}x,
\end{aligned}$$

also

$$\widehat{P}\widehat{P}^{\widehat{N}_1}_{\widehat{S}_1}\widehat{G}_2^{-1}\widehat{B}\hat{u} = \begin{pmatrix} (PP^{N_1}_{S_1}G_2^{-1}B - PP^{N_1}_{S_1}Q^{S_{1'}}_{S_1^t})z \\ 0 \end{pmatrix}.$$

Der andere Term ist

$$(\widehat{P}\widehat{P}^{\widehat{N}_1}_{\widehat{S}_1})'(\hat{u} + \widehat{P}\hat{y}) = \begin{pmatrix} (PP^{S_1^t}_{S_1})' & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} PP^{S_1^t}_{S_1}x + PQ^{S_1}_{S_1^t}x \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} (PP_{S_1}^{S_1^t})'(PP_{S_1}^{S_1^t}x + PQ_{S_1^t}^{S_1}x) \\ 0 \end{pmatrix} = \begin{pmatrix} (PP_{S_1}^{S_1^t})'(z + Q_{S_1^t}^{S_1}y) \\ 0 \end{pmatrix},$$

und die rechte Seite

$$\begin{aligned}
&\widehat{P}\widehat{P}_{\widehat{S}_1}^{\widehat{N}_1}\widehat{G}_2^{-1}\hat{q} \\
&= \begin{pmatrix} PP_{S_1}^{S_1^t}\Upsilon^{-1} & -PP_{S_1}^{S_1^t}\Upsilon^{-1}\widehat{B}_{11}P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1} \\ 0 & 0 \end{pmatrix} \\
&\times \begin{pmatrix} (I-W_1)q + W_1\left[(W_1q)' - W_1BA^+q\right] \\ \Gamma^Tq \end{pmatrix} \\
&= \begin{pmatrix} PP_{S_1}^{S_1^t}\left\{(G_2^{-1} + Q_{N_1}^{S_1}A^+)(I-W_1)q + Q_{N_1}^{S_1}(W_1B)^+W_1\left[(W_1q)' - W_1BA^+q\right]\right\} \\ 0 \end{pmatrix} \\
&- \begin{pmatrix} PP_{S_1}^{S_1^t}\Upsilon^{-1}\widehat{B}_{11}P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}\Gamma^Tq \\ 0 \end{pmatrix}.
\end{aligned}$$

Hier rechnet man auch mit der ersten Komponente dieses Vektors weiter:

$$\begin{aligned}
&PP_{S_1}^{S_1^t}\left[G_2^{-1}(I-W_1)q + Q_{N_1}^{S_1}(W_1B)^+W_1(W_1q)'\right] \\
&-PP_{S_1}^{S_1^t}\left[G_2^{-1}(I-W_1)B + Q_{N_1}^{S_1}(W_1B)^+W_1(W_1B)'\right]P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}\Gamma^Tq \\
= \;& PP_{S_1}^{N_1}G_2^{-1}(I-W_1)q + PP_{S_1}^{S_1^t}Q_{N_1}^{S_1}(W_1B)^+W_1(W_1q)' - PP_{S_1}^{N_1}G_2^{-1}(I-W_1) \\
&\times BP\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}\Gamma^Tq - PP_{S_1}^{S_1^t}Q_{N_1}^{S_1}(W_1B)^+W_1(W_1B)'P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1}\Gamma^Tq \\
= \;& PP_{S_1}^{N_1}G_2^{-1}(I-W_1)q + PP_{S_1}^{S_1^t}Q_{N_1}^{S_1}\left[Q_{S_1^t}^{S_1}(W_1B)^+W_1q\right]' - PP_{S_1}^{N_1}G_2^{-1}BPQ_{S_1^t}^{S_1}y \\
&+PP_{S_1}^{N_1}G_2^{-1}W_1q - PP_{S_1}^{S_1^t}Q_{N_1}^{S_1}Q_{S_1^t}^{S_1'}PQ_{S_1^t}^{S_1}y.
\end{aligned}$$

$PQ_{S_1^t}^{S_1}y$ wird durch den schon bekannten Ausdruck

$$PQ_{S_1^t}^{S_1}y = PQ_{S_1^t}^{S_1}(W_1B)^+W_1q$$

ersetzt.

Weiterhin folgt für die erste Komponente der rechten Seite

$$\begin{aligned}
= \;& PP_{S_1}^{N_1}G_2^{-1}q + PP_{S_1}^{S_1^t}Q_{N_1}^{S_1}(Q_{S_1^t}^{S_1}y)' - PP_{S_1}^{N_1}G_2^{-1}BPQ_{S_1^t}^{S_1}y - PP_{S_1}^{S_1^t}Q_{N_1}^{S_1}Q_{S_1^t}^{S_1'}PQ_{S_1^t}^{S_1}y \\
= \;& PP_{S_1}^{N_1}G_2^{-1}q + PP_{S_1}^{S_1^t}(Q_{S_1^t}^{S_1}y)' - PP_{S_1}^{N_1}(PQ_{S_1^t}^{S_1}y)' - PP_{S_1}^{N_1}G_2^{-1}BPQ_{S_1^t}^{S_1}y \\
= \;& PP_{S_1}^{N_1}G_2^{-1}q + PP_{S_1}^{S_1^t}Q_{S_1^t}^{S_1'}Q_{S_1^t}^{S_1}y - PP_{S_1}^{N_1}(PQ_{S_1^t}^{S_1}y)' - PP_{S_1}^{N_1}G_2^{-1}BPP_{S_1}^{N_1}Q_{S_1^t}^{S_1}y \\
= \;& PP_{S_1}^{N_1}G_2^{-1}q + PP_{S_1}^{N_1}Q_{S_1^t}^{S_1'}Q_{S_1^t}^{S_1}y - PP_{S_1}^{N_1}(PQ_{S_1^t}^{S_1}y)' - PP_{S_1}^{N_1}G_2^{-1}BPP_{S_1}^{N_1}Q_{S_1^t}^{S_1}y,
\end{aligned}$$

und schließlich erhält man für die nicht-triviale Komponente von (4.16)

$$z' - ((PP_{S_1}^{S_1^t})' + PP_{S_1}^{N_1}Q_{S_1^t}^{S_1\prime} - PP_{S_1}^{N_1}G_2^{-1}BPP_{S_1}^{N_1})(z + Q_{S_1^t}^{S_1}y) \qquad (4.20)$$
$$+PP_{S_1}^{N_1}(PQ_{S_1^t}^{S_1}y)' \;=\; PP_{S_1}^{N_1}G_2^{-1}q,$$

was die inhärente Differentialgleichung (3.13) ist.

Um das System (4.16)-(4.18) vollständig berechnen zu können, fehlt noch die Gleichung (4.17), die nur die neue Variable μ bestimmt. Für den Projektor $\widehat{T}\widehat{Q}\widehat{P}_{\widehat{S}_1}^{\widehat{N}_1}$ gilt

$$\begin{aligned}\widehat{T}\widehat{Q}\widehat{P}_{\widehat{S}_1}^{\widehat{N}_1} &= \begin{pmatrix} 0 & 0 \\ 0 & I_\kappa \end{pmatrix}\begin{pmatrix} Q & 0 \\ 0 & I_\kappa \end{pmatrix}\begin{pmatrix} I_m - PQ_{S_1^t}^{S_1} & 0 \\ -(\widehat{B}_{21}\widehat{B}_{12})^{-1}\widehat{B}_{21} & I_\kappa \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ -(\widehat{B}_{21}\widehat{B}_{12})^{-1}\widehat{B}_{21} & I_\kappa \end{pmatrix},\end{aligned}$$

dann ist

$$\begin{aligned}&\widehat{T}\widehat{Q}\widehat{P}_{\widehat{S}_1}^{\widehat{N}_1}\widehat{G}_2^{-1} \\ &= \begin{pmatrix} 0 & 0 \\ -(\widehat{B}_{21}\widehat{B}_{12})^{-1}\widehat{B}_{21}\Upsilon^{-1} & (\widehat{B}_{21}\widehat{B}_{12})^{-1}\widehat{B}_{21}\Upsilon^{-1}\widehat{B}_{11}P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1} \end{pmatrix},\end{aligned}$$

und der Term $\widehat{T}\widehat{Q}\widehat{P}_{\widehat{S}_1}^{\widehat{N}_1}\widehat{G}_2^{-1}\widehat{B}\hat{u}$ ist durch

$$\begin{aligned}&\begin{pmatrix} 0 & 0 \\ -(\widehat{B}_{21}\widehat{B}_{12})^{-1}\widehat{B}_{21}\Upsilon^{-1}\widehat{B}_{11}(I - PQ_{S_1^t}^{S_1}) & (\widehat{B}_{21}\widehat{B}_{12})^{-1}\widehat{B}_{21}\Upsilon^{-1}A\widehat{B}_{12} \end{pmatrix}\begin{pmatrix} z \\ 0 \end{pmatrix} \\ =\;& \begin{pmatrix} 0 \\ -(\widehat{B}_{21}\widehat{B}_{12})^{-1}\widehat{B}_{21}\Upsilon^{-1}\widehat{B}_{11}z \end{pmatrix} \\ =\;& \begin{pmatrix} 0 \\ -(\widehat{B}_{21}\widehat{B}_{122})^{-1}\widehat{B}_{21}\left[G_2^{-1}(I - W_1)B + Q_{N_1}^{S_1}(W_1B)^+W_1(W_1B)'\right]z \end{pmatrix} \\ =\;& \begin{pmatrix} 0 \\ -(\widehat{B}_{21}\widehat{B}_{12})^{-1}\widehat{B}_{21}Q_{S_1^t}^{S_1\prime}z \end{pmatrix} \\ =\;& \begin{pmatrix} 0 \\ -(\widehat{B}_{21}\widehat{B}_{12})^{-1}\widehat{B}_{21}'z \end{pmatrix}\end{aligned}$$

gegeben.

Die rechte Seite von (4.17) ist jetzt einfach zu berechnen

$$
\begin{aligned}
& \widehat{T}\widehat{Q}\widehat{P}_{\widehat{S}_1}^{\widehat{N}_1}\widehat{G}_2^{-1}\hat{q} \\
= & \begin{pmatrix} 0 & 0 \\ -(\widehat{B}_{21}\widehat{B}_{12})^{-1}\widehat{B}_{21}\Upsilon^{-1} & (\widehat{B}_{21}\widehat{B}_{12})^{-1}\widehat{B}_{21}\Upsilon^{-1}\widehat{B}_{11}P\widehat{B}_{12}(\widehat{B}_{21}\widehat{B}_{12})^{-1} \end{pmatrix} \\
\times & \begin{pmatrix} (I-W_1)q + W_1\left[(W_1q)' - W_1BA^+q\right] \\ \Gamma^T q \end{pmatrix} \\
= & \begin{pmatrix} 0 \\ -(\widehat{B}_{21}\widehat{B}_{12})^{-1}\widehat{B}_{21}\Upsilon^{-1}\left\{(I-W_1)q + W_1\left[(W_1q)' - W_1BA^+q\right] - \widehat{B}_{11}PQ_{S_1^t}^{S_1}y\right\} \end{pmatrix} \\
= & \begin{pmatrix} 0 \\ -(\widehat{B}_{21}\widehat{B}_{21}^T)^{-1}\widehat{B}_{21}\left\{(Q_{S_1^t}^{S_1}y)' - (Q_{S_1^t}^{S_1}y)' + Q_{S_1^t}^{S_1}(Q_{S_1^t}^{S_1}y)'\right\} \end{pmatrix} \\
= & \begin{pmatrix} 0 \\ -(\widehat{B}_{21}\widehat{B}_{12})^{-1}\widehat{B}_{21}(Q_{S_1^t}^{S_1}y)' \end{pmatrix}.
\end{aligned}
$$

Der nächste Term in der Berechnung ist $(\widehat{Q}\widehat{Q}_{\widehat{N}_1}^{\widehat{S}_1})'\hat{u}$,

$$
\begin{aligned}
= & \left[\begin{pmatrix} Q & 0 \\ 0 & I_\kappa \end{pmatrix}\begin{pmatrix} PQ_{S_1^t}^{S_1} & 0 \\ \left[(\widehat{B}_{21}\widehat{B}_{12})^{-1}\widehat{B}_{21}\right] & 0 \end{pmatrix}\right]'\begin{pmatrix} z \\ 0 \end{pmatrix} \\
= & \begin{pmatrix} 0 & 0 \\ \left[(\widehat{B}_{21}\widehat{B}_{12})^{-1}\widehat{B}_{21}\right]' & 0 \end{pmatrix}\begin{pmatrix} z \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ (\widehat{B}_{21}\widehat{B}_{12})^{-1}\widehat{B}_{21}'z \end{pmatrix}.
\end{aligned}
$$

Als Letztes berechnet man den Term $\widehat{Q}\widehat{Q}_{\widehat{N}_1}^{\widehat{S}_1}(\widehat{P}\hat{y})'$

$$
\begin{aligned}
= & \begin{pmatrix} 0 & 0 \\ (\widehat{B}_{21}\widehat{B}_{12})^{-1}\widehat{B}_{21} & 0 \end{pmatrix}\left[\begin{pmatrix} P & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} (Q+PQ_{S_1^t}^{S_1})(y+w) \\ 0 \end{pmatrix}\right]' \\
= & \begin{pmatrix} 0 & 0 \\ (\widehat{B}_{21}\widehat{B}_{12})^{-1}\widehat{B}_{21} & 0 \end{pmatrix}\begin{pmatrix} P(Q_{S_1^t}^{S_1}y)' \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ (\widehat{B}_{21}\widehat{B}_{12})^{-1}\widehat{B}_{21}(Q_{S_1^t}^{S_1}y)' \end{pmatrix}.
\end{aligned}
$$

Und schließlich setzt man die Gleichung (4.17) zusammen. Es folgt

$$
\begin{aligned}
& -\begin{pmatrix} 0 \\ (\widehat{B}_{21}\widehat{B}_{12})^{-1}\widehat{B}_{21}(Q_{S_1^t}^{S_1}y)' \end{pmatrix} + \begin{pmatrix} 0 \\ (\widehat{B}_{21}\widehat{B}_{12})^{-1}\widehat{B}_{21}'z \end{pmatrix} - \begin{pmatrix} 0 \\ (\widehat{B}_{21}\widehat{B}_{12})^{-1}\widehat{B}_{21}'z \end{pmatrix} \\
= & +\begin{pmatrix} 0 \\ \mu \end{pmatrix} - \begin{pmatrix} 0 \\ (\widehat{B}_{21}\widehat{B}_{12})^{-1}\widehat{B}_{21}(Q_{S_1^t}^{S_1}y)' \end{pmatrix},
\end{aligned}
$$

was äquivalent zu $\mu = 0$ ist.

Die Ergebnisse dieses Abschnittes können wie folgt zusammengefasst werden:

Schlussfolgerung 4.3.3 *Die kanonische Entkopplung für den GGL-Ansatz (4.14) ist nichts anderes als die Entkopplung längs S_1 mit dem Projektor $Q_{S_1^t}^{S_1} = \Psi(\Gamma^T B \Psi)^{-1}\Gamma^T B$ für (4.8) mit der zusätzlichen Gleichung $\mu = 0$.*

4.3.2 BDF-Entkopplung

Nach der detaillierten Berechnung der Entkopplung der GGL-Formulierung (4.14) wird für die BDF-Methode die Gültigkeit des Diagramms 4.6 untersucht. Für die Analyse werden eine homogene lineare Index-2-ADGl der Form

$$A(t)x'(t) + B(t)x(t) = 0 \tag{4.21}$$

und ihr GGL-Ansatz

$$\widehat{A}(t)\hat{x}'(t) + \widehat{B}(t)\hat{x}(t) = 0 \tag{4.22}$$

betrachtet. Hier sind die Gleichungskoeffizienten wie in 4.2 definiert.

Es wird das BDF-Verfahren für (4.22) mit dem Projektor $\widehat{Q}_{\widehat{N}_1}^{\widehat{S}_1}$ entkoppelt und schließlich mit der entsprechenden Diskretisierung der entkoppelten Gleichungen verglichen.

Im idealen Fall würde man die Gleichung (4.22) erst entkoppeln, bevor das numerische Verfahren angewendet wird. Die BDF-Methode, angewandt auf die Entkopplungsgleichungen von (4.22), nachdem man wie in 4.3.1 das System in die Variablen z, y, w, μ umschreibt, lautet

$$\begin{aligned}
\sum_{j=0}^{k} \alpha_j z_{i-j} - h((PP_{S_1}^{S_1^t})' PP_{S_1}^{S_1^t} + PP_{S_1}^{N_1} Q_{S_1^t}^{S_1\prime} PP_{S_1}^{S_1^t} & \qquad (4.23)\\
-PP_{S_1}^{N_1} G_2^{-1} B PP_{S_1}^{S_1^t})_i z_i &= 0,\\
w_i + (QQ_{N_1}^{S_1} Q_{S_1^t}^{S_1\prime} + TQP_{S_1}^{N_1} G_2^{-1} B PP_{S_1}^{N_1})_i z_i &= 0, \qquad (4.24)\\
(I + UQG_2^{-1} B PP_{S_1}^{N_1} Q_{S_1^t}^{S_1})_i y_i + (UQG_2^{-1} B PP_{S_1}^{N_1})_i z_i &= 0, \qquad (4.25)\\
\mu_i &= 0, \qquad (4.26)
\end{aligned}$$

für $i \geq k$.

Von (4.23) folgt für z_i

$$\left[\alpha_0 I - h((PP_{S_1}^{S_1^t})' PP_{S_1}^{S_1^t} + PP_{S_1}^{N_1} Q_{S_1^t}^{S_{1'}} PP_{S_1}^{S_1^t} - PP_{S_1}^{N_1} G_2^{-1} BPP_{S_1}^{S_1^t})\right]_i z_i = r_i^z, \tag{4.27}$$

wobei

$$r_i^z := -\sum_{j=1}^{k} \alpha_j z_{i-j}.$$

Diese Rekursion für die Variable z kann mit Hilfe der Projektoren $PP_{S_1}^{S_1^t}$, $PQ_{S_1^t}^{S_1}$ und Q folgendermaßen in die entsprechenden Komponenten zerlegt werden:

$$\begin{aligned}
(PP_{S_1}^{S_1^t})_i: \quad & \left[\alpha_0 I - h(PP_{S_1}^{N_1} Q_{S_1^t}^{S_1} - PP_{S_1}^{N_1} G_2^{-1} B)\right]_i (PP_{S_1}^{S_1^t} z)_i && (4.28)\\
& = (PP_{S_1}^{S_1^t} r^z)_i, \\
(PQ_{S_1^t}^{S_1})_i: \quad & (PQ_{S_1^t}^{S_1} z)_i = \frac{1}{\alpha_0} (PQ_{S_1^t}^{S_1})_i \left[h(PP_{S_1}^{S_1^t})' PP_{S_1}^{S_1^t} z_i - \sum_{j=0}^{k} \alpha_j z_{i-j}\right] && (4.29)\\
Q: \quad & Qz_i = 0. && (4.30)
\end{aligned}$$

Für die anderen Lösungskomponenten erhält man offensichtlich

$$\begin{aligned}
y_i &= (UQ + PQ_{N_1}^{S_1})_i y_i && (4.31)\\
&= -(I - UQG_2^{-1} BPP_{S_1}^{N_1} Q_{S_1^t}^{S_1})_i (UQG_2^{-1} BPP_{S_1}^{N_1})_i z_i, \\
w_i &= (TQ)_i w_i && (4.32)\\
&= -(QQ_{N_1}^{S_1} Q_{S_1^t}^{S_{1'}} + TQP_{S_1}^{N_1} G_2^{-1} BPP_{S_1}^{N_1})_i z_i.
\end{aligned}$$

Hier ist zu unterstreichen, dass für einen konstanten Projektor $PP_{S_1}^{S_1^t}$ die $PQ_{S_1^t}^{S_1}$-Komponente von z_i null ist.

Für die asymptotische Stabilitätsanalyse schreibt man die Rekursion (4.27) als eine Einschrittsrekursion. Sei

$$M_{h,i} := \left[\alpha_0 I - h((PP_{S_1}^{S_1^t})' PP_{S_1}^{S_1^t} + PP_{S_1}^{N_1} Q_{S_1^t}^{S_{1'}} PP_{S_1}^{S_1^t} - PP_{S_1}^{N_1} G_2^{-1} BPP_{S_1}^{S_1^t})\right]_i,$$

dann kann auch (4.27) als

$$\underbrace{\begin{pmatrix} z_i \\ z_{i-1} \\ \vdots \\ z_{i-k+1} \end{pmatrix}}_{Z_i} = \underbrace{\begin{pmatrix} -\alpha_1 M_{h,i}^{-1} & -\alpha_2 M_{h,i}^{-1} & \dots & -\alpha_k M_{h,i}^{-1} \\ I_m & & & \\ & \ddots & & \\ & & I_m & \end{pmatrix}}_{R_i} \underbrace{\begin{pmatrix} z_{i-1} \\ z_{i-2} \\ \vdots \\ z_{i-k} \end{pmatrix}}_{Z_{i-1}} \tag{4.33}$$

geschrieben werden.

Man bemerke, dass diese Rekursion genau dann in einem Integrationsschritt eine Kontraktion in einer bestimmten Norm darstellt, wenn der Spektralradius von R_i kleiner als 1 ist. Hier wird etwas Stärkeres angenommen, nämlich dass in einer vorgegebenen Skalarproduktnorm $\|R_i\| < 1$.

Jetzt wird realitätsnäher verfahren: Man diskretisiert zuerst (4.22) und dann entkoppelt man diese Diskretisierung mit Hilfe des kanonischen Projektors. Die BDF-Methode für (4.22) ist

$$\widehat{A}_i \sum_{j=0}^{k} \alpha_j \hat{x}_{i-j} + h\widehat{B}_i \hat{x}_i = 0, \quad i \geq k. \tag{4.34}$$

Um die Entkopplung durchzuführen, nimmt man die kanonischen Projektoren $\widehat{Q}_{\widehat{N}_1}^{\widehat{S}_1}$, $\widehat{P}_{\widehat{S}_1}^{\widehat{N}_1}$. Man skaliert zunächst (4.34) mit $(\widehat{G}_2^{-1})_i$ und es folgt (Abschnitt 3.1.1)

$$(\widehat{P}_{\widehat{S}_1}^{\widehat{N}_1})_i \widehat{P} \sum_{j=0}^{k} \alpha_j \hat{x}_{i-j} + h(\widehat{G}_2^{-1} \widehat{B} \widehat{P} \widehat{P}_{\widehat{S}_1}^{\widehat{N}_1} \hat{x})_i + h(\widehat{Q}_{\widehat{N}_1}^{\widehat{S}_1} \hat{x})_i + h\widehat{Q}\hat{x}_i = 0. \tag{4.35}$$

Als Nächstes multipliziert man (4.35) mit den Projektoren $(\widehat{P}\widehat{P}_{\widehat{S}_1}^{\widehat{N}_1})_i$, $(\widehat{U}\widehat{Q} + \widehat{P}\widehat{Q}_{\widehat{N}_1}^{\widehat{S}_1})_i$ und $(\widehat{T}\widehat{Q}\widehat{P}_{\widehat{S}_1}^{\widehat{N}_1})_i$, und man erhält

$$\sum_{j=0}^{k} \alpha_j (\widehat{P}\widehat{P}_{\widehat{S}_1}^{\widehat{N}_1} \hat{x})_{i-j} - \sum_{j=1}^{k} \alpha_j \left[(\widehat{P}\widehat{P}_{\widehat{S}_1}^{\widehat{N}_1})_{i-j} - (\widehat{P}\widehat{P}_{\widehat{S}_1}^{\widehat{N}_1})_i \right] \tag{4.36}$$

$$\times (\widehat{P}\widehat{P}_{\widehat{S}_1}^{\widehat{N}_1} \hat{x})_{i-j} + h(\widehat{P}\widehat{P}_{\widehat{S}_1}^{\widehat{N}_1} \widehat{G}_2^{-1} \widehat{B} \widehat{P} \widehat{P}_{\widehat{S}_1}^{\widehat{N}_1} \hat{x})_i = 0,$$

$$(\widehat{U}\widehat{Q}\widehat{G}_2^{-1} \widehat{B} \widehat{P} \widehat{P}_{\widehat{S}_1}^{\widehat{N}_1} \hat{x})_i + (\widehat{U}\widehat{Q} + \widehat{P}\widehat{Q}_{\widehat{N}_1}^{\widehat{S}_1})_i \hat{x}_i = 0, \tag{4.37}$$

$$-\widehat{Q} \sum_{j=0}^{k} \alpha_j \left((\widehat{Q}_{\widehat{N}_1}^{\widehat{S}_1})_i - (\widehat{Q}_{\widehat{N}_1}^{\widehat{S}_1})_{i-j} \right) (\widehat{P}\widehat{P}_{\widehat{S}_1}^{\widehat{N}_1} \hat{x})_{i-j} \tag{4.38}$$

$$+h(\widehat{T}\widehat{Q}\widehat{P}^{\widehat{N}_1}_{\widehat{S}_1}\widehat{G}_2^{-1}\widehat{B}\widehat{P}\widehat{P}^{\widehat{N}_1}_{\widehat{S}_1}\hat{x})_i + h(\widehat{T}\widehat{Q}\hat{x})_i \;=\; 0.$$

Diese Gleichungen stellen die diskrete kanonische Entkopplung für (4.34) dar.

Nun wird das System (4.36)-(4.38), analog zu 4.3.1 in dem kontinuierlichen Fall, mit den Variablen $\bar{z}_i := (PP^{S_1^t}_{S_1}x)_i$, $\bar{y}_i := (UQ + PQ^{S_1}_{N_1})_i x_i$, $\bar{w}_i := (TQx)_i$ und $\bar{\mu}_i$ umgeschrieben. Hier kennzeichnet man die numerischen Approximationen von (4.36)-(4.38) durch $\bar{z}_i$, um sie von den (4.23)-(4.26)-Approximationen zu unterscheiden.

Jetzt können die Berechnungen von 4.3.1 verwendet werden. Die Gleichung (4.37) ist nichts anderes als (4.18) an der Stelle t_i ausgewertet. Dann kann man ohne Weiteres sagen, dass (4.37) äquivalent zu

$$(I + UQG_2^{-1}BPP^{N_1}_{S_1}Q^{S_1}_{S_1^t})_i\bar{y}_i + (UQG_2^{-1}BPP^{N_1}_{S_1})_i\bar{z}_i \;=\; 0, \qquad (4.39)$$

$$(QQ^{S_1}_{N_1}Q^{S_{1'}}_{S_1^t} + TQG_2^{-1}BPP^{N_1}_{S_1})_i\bar{z}_i + \bar{w}_i \;=\; 0 \qquad (4.40)$$

ist und das sind genau die Entkopplungsgleichungen (4.25) und (4.24) an der Stelle t_i ausgewertet.

Jetzt werden die dynamischen Komponenten analysiert, also die Gleichung (4.36). Die einzigen Terme, die neu berechnet werden müssen, sind

$$\sum_{j=0}^{k}\alpha_j(\widehat{P}\widehat{P}^{\widehat{N}_1}_{\widehat{S}_1}\hat{x})_{i-j},$$

$$\sum_{j=1}^{k}\alpha_j\left[(\widehat{P}\widehat{P}^{\widehat{N}_1}_{\widehat{S}_1})_{i-j} - (\widehat{P}\widehat{P}^{\widehat{N}_1}_{\widehat{S}_1})_i\right](\widehat{P}\hat{x})_{i-j}.$$

Der Rest ist wie in 4.3.1 an der Stelle t_i ausgewertet.

Der erste Term ist

$$\sum_{j=0}^{k}\alpha_j(\widehat{P}\widehat{P}^{\widehat{N}_1}_{\widehat{S}_1}\hat{x})_{i-j} = \sum_{j=0}^{k}\alpha_j\begin{pmatrix} PP^{S_1^t}_{S_1}x \\ 0\end{pmatrix}_{i-j} = \sum_{j=0}^{k}\alpha_j\begin{pmatrix}\bar{z}\\0\end{pmatrix}_{i-j},$$

und der zweite

$$\sum_{j=1}^{k}\alpha_j\left[(\widehat{P}\widehat{P}^{\widehat{N}_1}_{\widehat{S}_1})_{i-j} - (\widehat{P}\widehat{P}^{\widehat{N}_1}_{\widehat{S}_1})_i\right](\widehat{P}\hat{x})_{i-j}$$

$$= \sum_{j=1}^{k}\alpha_j\left[\begin{pmatrix} PP^{S_1^t}_{S_1} & 0\\ 0 & 0\end{pmatrix}_{i-j} - \begin{pmatrix} PP^{S_1^t}_{S_1} & 0\\ 0 & 0\end{pmatrix}_{i}\right]\begin{pmatrix}Px\\0\end{pmatrix}_{i-j}$$

$$= \sum_{j=1}^{k} \alpha_j \left[\begin{pmatrix} PP_{S_1}^{S_1^t} & 0 \\ 0 & 0 \end{pmatrix}_{i-j} - \begin{pmatrix} PP_{S_1}^{S_1^t} & 0 \\ 0 & 0 \end{pmatrix}_{i} \right] \begin{pmatrix} \bar{z} \\ 0 \end{pmatrix}_{i-j}.$$

Nun kann man die Gleichung (4.36) zusammensetzen, man erhält

$$\begin{aligned} \sum_{j=0}^{k} \alpha_j \begin{pmatrix} \bar{z} \\ 0 \end{pmatrix}_{i-j} - \sum_{j=1}^{k} \alpha_j \left[\begin{pmatrix} PP_{S_1}^{S_1^t} & 0 \\ 0 & 0 \end{pmatrix}_{i-j} - \begin{pmatrix} PP_{S_1}^{S_1^t} & 0 \\ 0 & 0 \end{pmatrix}_{i} \right] \begin{pmatrix} \bar{z} \\ 0 \end{pmatrix}_{i-j} \\ + h \begin{pmatrix} (PP_{S_1}^{N_1} G_2^{-1} B - PP_{S_1}^{N_1} Q_{S_1^t}^{S_1}) PP_{S_1}^{S_1^t} \bar{z} \\ 0 \end{pmatrix}_i &= 0, \end{aligned}$$

und wenn man die erste Komponente dieses Vektors betrachtet, folgt

$$\begin{aligned} \sum_{j=0}^{k} \alpha_j \bar{z}_{i-j} - \sum_{j=1}^{k} \alpha_j ((PP_{S_1}^{S_1^t})_{i-j} - (PP_{S_1}^{S_1^t})_i) \bar{z}_{i-j} - h(PP_{S_1}^{N_1} Q_{S_1^t}^{S_1'} PP_{S_1}^{S_1^t} \\ - PP_{S_1}^{N_1} G_2^{-1} B PP_{S_1}^{S_1^t})_i \bar{z}_i &= 0. \end{aligned} \tag{4.41}$$

Hier sieht man, dass sich diese Diskretisierung nur um dem Term

$$\sum_{j=1}^{k} \alpha_j ((PP_{S_1}^{S_1^t})_{i-j} - (PP_{S_1}^{S_1^t})_i) \bar{z}_{i-j},$$

der an der Stelle von $h(PP_{S_1}^{S_1^t\prime} PP_{S_1}^{S_1^t} z)_i$ auftritt, von der "idealen" Diskretisierung (4.23) unterscheidet.

Aus (4.41) ergibt sich für die Variable $\bar{z}_i$ das lineare System

$$\left[\alpha_0 I - h(PP_{S_1}^{N_1} Q_{S_1^t}^{S_1} PP_{S_1}^{S_1^t} - PP_{S_1}^{N_1} G_2^{-1} B PP_{S_1}^{S_1^t}) \right]_i \bar{z}_i = \bar{r}_i^z, \tag{4.42}$$

wobei

$$\bar{r}_i^z := -(PP_{S_1}^{S_1^t})_i \sum_{j=1}^{k} \alpha_j \bar{z}_{i-j}.$$

An der Stelle ist darauf aufmerksam zu machen, dass $\bar{r}_i^z \in \operatorname{im} (PP_{S_1}^{S_1^t})_i$. Wenn man, wie bei der Diskretisierung (4.23), das Gleichungssystem in die Komponenten $PP_{S_1}^{S_1^t}$, $PQ_{S_1^t}^{S_1}$ und Q zerlegt, folgt

$$
\begin{aligned}
(PP_{S_1}^{S_1^t})_i \quad &: \quad \left[\alpha_0 I - h(PP_{S_1}^{N_1} Q_{S_1^t}^{S_{1'}} PP_{S_1}^{S_1^t} - PP_{S_1}^{N_1} G_2^{-1} BPP_{S_1}^{S_1^t})\right]_i \qquad (4.43)\\
&\quad \times (PP_{S_1}^{S_1^t}\bar{z})_i = \bar{r}_i^z,\\
(PQ_{S_1^t}^{S_1})_i \quad &: \quad (PQ_{S_1^t}^{S_1}\bar{z})_i = 0, \qquad (4.44)\\
Q \quad &: \quad (Q\bar{z})_i = 0. \qquad (4.45)
\end{aligned}
$$

Es werden die ersten Schlussfolgerungen gezogen. Die numerische BDF-Approximation (falls das System (4.43) eindeutig lösbar ist) liegt im im $(PP_{S_1}^{S_1^t})_i$, was als positiv zu werten ist. Zweitens ist die $PP_{S_1}^{S_1^t}$-Komponente von $\bar{z}_i$ Lösung eines linearen Systems mit gleicher Matrix wie bei der Diskretisierung (4.23).

Sei $\overline{M}_{h,i} := \left[\alpha_0 I - h(PP_{S_1}^{N_1} Q_{S_1^t}^{S_{1'}} PP_{S_1}^{S_1^t} - PP_{S_1}^{N_1} G_2^{-1} BPP_{S_1}^{S_1^t})\right]_i$, dann gilt die folgende Aussage:

Lemma 4.3.4 *Die Matrix $M_{h,i}$ ist genau dann regulär, wenn die Einschränkung von $\overline{M}_{h,i}$ auf* $\operatorname{im}(PP_{S_1}^{S_1^t})_i$ *eine Bijektion auf diesem Unterraum darstellt.*

Beweis: Z uerst ersieht man aus der Definition von $\overline{M}_{h,i}$, dass diese Anwendung im $(PP_{S_1}^{S_1^t})_i$ in sich selbst abbildet ($(\overline{M}_h PP_{S_1}^{S_1^t})_i = (PP_{S_1}^{S_1^t}\overline{M}_h)_i$).

Man betrachtet nun das Gleichungssystem

$$M_{h,i}x = 0.$$

Dieses System kann in die Komponenten $PP_{S_1}^{S_1^t}$, $PQ_{S_1^t}^{S_1}$ und Q zerlegt werden,

$$
\begin{aligned}
PP_{S_1}^{S_1^t}: \quad & (PP_{S_1}^{S_1^t})_i \left[\alpha_0 I - h((PP_{S_1}^{S_1^t})' PP_{S_1}^{S_1^t} + PP_{S_1}^{N_1} Q_{S_1^t}^{S_{1'}} PP_{S_1}^{S_1^t}\right.\\
& \left. - PP_{S_1}^{N_1} G_2^{-1} BPP_{S_1}^{S_1^t})\right]_i x = 0,\\
& (PP_{S_1}^{S_1^t})_i \left[\alpha_0 I - h(PP_{S_1}^{N_1} Q_{S_1^t}^{S_{1'}} PP_{S_1}^{S_1^t} - PP_{S_1}^{N_1} G_2^{-1} BPP_{S_1}^{S_1^t})\right]_i x = 0,\\
& (\overline{M}_h PP_{S_1}^{S_1^t})_i x = 0,
\end{aligned}
$$

$$
\begin{aligned}
PQ^{S_1}_{S_1^t}: \quad & (PQ^{S_1}_{S_1^t})_i \Big[\alpha_0 I - h(PP^{S_1^t}_{S_1'} PP^{S_1^t}_{S_1} + PP_{1,c} Q^{S_{1'}}_{S_1^t} PP^{S_1^t}_{S_1} \\
& -PP_{1,c} G_2^{-1} BPP^{S_1^t}_{S_1})\Big]_i x = 0, \\
& (PQ^{S_1}_{S_1^t})_i (\alpha_0 I - hPP^{S_1^t}_{S_1'} PP^{S_1^t}_{S_1})_i x = 0, \\
& (PQ^{S_1}_{S_1^t})_i x = \frac{h}{\alpha_0} (PQ^{S_1}_{S_1^t} PP^{S_1^t}_{S_1'} PP^{S_1^t}_{S_1})_i x,
\end{aligned}
$$

$$
\begin{aligned}
Q: \quad & Q_i \Big[\alpha_0 I - h(PP^{S_1^t}_{S_1'} PP^{S_1^t}_{S_1} + PP^{N_1}_{S_1} Q^{S_{1'}}_{S_1^t} PP^{S_1^t}_{S_1} - PP^{N_1}_{S_1} G_2^{-1} BPP^{S_1^t}_{S_1})\Big]_i x = 0, \\
& Qx = 0.
\end{aligned}
$$

Aus diesen Gleichungen ist es möglich, die Aussage des Lemmas herzuleiten. Sei jetzt

$$\overline{M}^-_{h,i} : \mathrm{im}\,(PP^{S_1^t}_{S_1})_i \to \mathrm{im}\,(PP^{S_1^t}_{S_1})_i$$

die Inverse der Anwendung $\overline{M}_{h,i}\Big|_{PS_{1,i}}$. Mit der gleichen Vorgehensweise wie der des letzten Lemmas folgt für $M^{-1}_{h,i}$ der Ausdruck

$$M^{-1}_{h,i} = (I + \frac{h}{\alpha_0}(PP^{S_1^t}_{S_1})' PP^{S_1^t}_{S_1})_i (\overline{M}^-_h PP^{S_1^t}_{S_1})_i + \frac{1}{\alpha_0}(PQ^{S_1}_{S_1^t} + Q)_i.$$

Lemma 4.3.5 *Die Matrix $\overline{M}_{h,i}$ ist genau dann invertierbar, wenn $M_{h,i}$ invertierbar ist.*

Beweis: A us der Regularität von $\overline{M}_{h,i}$ folgt wie im Lemma 4.3.4 die Regularität für $M_{h,i}$. Für die Umkehrung stellt man zuerst fest, dass diese Matrizen die Beziehung

$$\overline{M}_{h,i} = M_{h,i} + h(PP^{S_1^t}_{S_1})'_i (PP^{S_1^t}_{S_1})_i = M_{h,i}(I + hM^{-1}_h (PP^{S_1^t}_{S_1})' PP^{S_1^t}_{S_1})_i,$$

erfüllen. Nach dem Ausdruck von $M^{-1}_{h,i}$ hat man

$$M^{-1}_{h,i}(PP^{S_1^t}_{S_1})'_i (PP^{S_1^t}_{S_1})_i = \frac{1}{\alpha_0}(PQ^{S_1}_{S_1^t})_i (PP^{S_1^t}_{S_1})'_i (PP^{S_1^t}_{S_1})_i = \frac{1}{\alpha_0}(PP^{S_1^t}_{S_1})'_i (PP^{S_1^t}_{S_1})_i,$$

dann folgt

$$
\begin{aligned}
\overline{M}_{h,i} &= M_{h,i}(I + \frac{h}{\alpha_0}(PP^{S_1^t}_{S_1})' PP^{S_1^t}_{S_1})_i, \\
\overline{M}^{-1}_{h,i} &= (I - \frac{h}{\alpha_0}(PP^{S_1^t}_{S_1})' PP^{S_1^t}_{S_1})_i M^{-1}_{h,i}.
\end{aligned}
$$

Als Folge von Lemma 4.3.5 hat man

$$(PP_{S_1}^{S_1^t}\overline{M}_h^{-1})_i = (PP_{S_1}^{S_1^t}M_h^{-1})_i,$$

außerdem gilt wegen der Invarianz von $\operatorname{im}(PP_{S_1}^{S_1^t})_i$ unter $\overline{M}_{h,i}^{-1}$

$$(PP_{S_1}^{S_1^t}\overline{M}_h^{-1}PP_{S_1}^{S_1^t})_i = (\overline{M}_h^{-1}PP_{S_1}^{S_1^t})_i.$$

Andererseits erhält man durch die Zerlegung der Gleichung $\overline{M}_{h,i}x = b$ in die Teile $(PP_{S_1}^{S_1^t})_i$, $(PQ_{S_1^t}^{S_1})_i$ und Q für $\overline{M}_{h,i}^{-1}$ auch den Ausdruck

$$\overline{M}_{h,i}^{-1} = (\overline{M}_h^{-1}PP_{S_1}^{S_1^t})_i + (PQ_{S_1^t}^{S_1})_i + Q,$$

woraus folgt

$$\begin{aligned}(PP_{S_1}^{S_1^t}\overline{M}_h^{-1})_i &= (PP_{S_1}^{S_1^t}\overline{M}_h^{-1}PP_{S_1}^{S_1^t})_i,\\ (PP_{S_1}^{S_1^t}\overline{M}_h^{-1})_i &= (\overline{M}_h^{-1}PP_{S_1}^{S_1^t})_i.\end{aligned}$$

Wenn man jetzt wieder die Rekursion (4.42) als ein Einschrittverfahren umschreibt, folgt

$$\underbrace{\begin{pmatrix}\bar{z}_i\\ \bar{z}_{i-1}\\ \vdots\\ \bar{z}_{i-k+1}\end{pmatrix}}_{\overline{Z}_i} = \underbrace{\begin{pmatrix} -\alpha_1(\overline{M}_h^{-1}PP_{S_1}^{S_1^t})_i & -\alpha_2(\overline{M}_h^{-1}PP_{S_1}^{S_1^t})_i & \dots & -\alpha_k(\overline{M}_h^{-1}PP_{S_1}^{S_1^t})_i\\ I_m & & & \\ & \ddots & & \\ & & & I_m\end{pmatrix}}_{\overline{R}_i} \times \underbrace{\begin{pmatrix}\bar{z}_{i-1}\\ \bar{z}_{i-2}\\ \vdots\\ \bar{z}_{i-k}\end{pmatrix}}_{\overline{Z}_{i-1}}, \tag{4.46}$$

Weiterhin kann die Matrix $\overline{R}_i$ als

$$\overline{R}_i = \begin{pmatrix} -\alpha_1(PP_{S_1}^{S_1^t}M_h^{-1})_i & -\alpha_2(PP_{S_1}^{S_1^t}M_h^{-1})_i & \dots & -\alpha_k(PP_{S_1}^{S_1^t}M_h^{-1})_i\\ I_m & & & \\ & \ddots & & \\ & & & I_m\end{pmatrix}$$

$$
\begin{aligned}
&= \begin{pmatrix} (PP_{S_1}^{S_1^t})_i & & & \\ & I_m & & \\ & & \ddots & \\ & & & I_m \end{pmatrix} \begin{pmatrix} -\alpha_1 M_{h,i}^{-1} & -\alpha_2 M_{h,i}^{-1} & \dots & -\alpha_k M_{h,i}^{-1} \\ I_m & & & \\ & \ddots & & \\ & & I_m & \end{pmatrix} \\
&= \begin{pmatrix} (PP_{S_1}^{S_1^t})_i & & & \\ & I_m & & \\ & & \ddots & \\ & & & I_m \end{pmatrix} R_i,
\end{aligned}
$$

ausgedrückt werden. Aus dieser Beziehung zwischen $\overline{R}_i$ und R_i folgt die Ungleichung

$$
\left\|\overline{R}_i\right\| \leq \left\| \begin{pmatrix} (PP_{S_1}^{S_1^t})_i & & & \\ & I_m & & \\ & & \ddots & \\ & & & I_m \end{pmatrix} \right\| \|R_i\|, \tag{4.47}
$$

wobei $\|\cdot\|$ eine beliebige Norm in $\underbrace{I\!R^m \times I\!R^m \times \cdots \times I\!R^m}_{k}$ repräsentiert.

Schließlich will man (4.38) betrachten. Hier ist ausschließlich der Term

$$
\widehat{Q} \sum_{j=0}^{k} \alpha_j \left((\widehat{Q}_{\widehat{N}_1}^{\widehat{S}_1})_i - (\widehat{Q}_{\widehat{N}_1}^{\widehat{S}_1})_{i-j} \right) (\widehat{P}\widehat{P}_{\widehat{S}_1}^{\widehat{N}_1} \hat{x})_{i-j}
$$

zu berechnen:

$$
\begin{aligned}
&\widehat{Q} \sum_{j=0}^{k} \alpha_j \left((\widehat{Q}_{\widehat{N}_1}^{\widehat{S}_1})_i - (\widehat{Q}_{\widehat{N}_1}^{\widehat{S}_1})_{i-j} \right) (\widehat{P}\widehat{P}_{\widehat{S}_1}^{\widehat{N}_1} \hat{x})_{i-j} \\
&= \begin{pmatrix} Q & 0 \\ 0 & I \end{pmatrix} \sum_{j=0}^{k} \alpha_j \begin{pmatrix} PQ_{S_1^t}^{S_1} & 0 \\ (\widehat{B}_{21}\widehat{B}_{21}^T)^{-1}\widehat{B}_{21} & 0 \end{pmatrix}_i \begin{pmatrix} PP_{S_1}^{S_1^t} x \\ 0 \end{pmatrix}_{i-j} \\
&= \begin{pmatrix} Q & 0 \\ 0 & I \end{pmatrix} \sum_{j=1}^{k} \alpha_j \begin{pmatrix} (PQ_{S_1^t}^{S_1})_i (PP_{S_1}^{S_1^t} x)_{i-j} \\ \left[(\widehat{B}_{21}\widehat{B}_{21}^T)^{-1}\widehat{B}_{21} \right]_i (PP_{S_1}^{S_1^t} x)_{i-j} \end{pmatrix} \\
&= \sum_{j=1}^{k} \alpha_j \begin{pmatrix} 0 \\ \left[(\widehat{B}_{21}\widehat{B}_{21}^T)^{-1}\widehat{B}_{21} \right]_i \bar{z}_{i-j} \end{pmatrix}.
\end{aligned}
$$

Jetzt setzt man die Gleichung (4.38) zusammen und es folgt

$$
-\sum_{j=0}^{k} \alpha_j \begin{pmatrix} 0 \\ \left[(\widehat{B}_{21}\widehat{B}_{21}^T)^{-1}\widehat{B}_{21} \right]_i \bar{z}_{i-j} \end{pmatrix} + h \begin{pmatrix} 0 \\ \bar{\mu}_i \end{pmatrix}
$$

$$-h \begin{pmatrix} 0 \\ \left[(\widehat{B}_{21}\widehat{B}_{21}^T)^{-1}\widehat{B}_{21}'\right]_i \bar{z}_i \end{pmatrix} = 0,$$

also

$$\bar{\mu}_i = \left[(\widehat{B}_{21}\widehat{B}_{21}^T)^{-1}\right]_i \left[(\widehat{B}_{21})_i' \bar{z}_i + (\widehat{B}_{21})_i \frac{1}{h} \sum_{j=0}^{k} \alpha_j \bar{z}_{i-j}\right]. \tag{4.48}$$

Die Variable $\bar{\mu}_i$ enthält einen Diskretisierungsfehler, der aber keinen Einfluss auf die anderen Komponenten hat.

Auf Grund der Analyse dieser Sektion können die folgenden Aussagen getroffen werden:

Theorem 4.3.6 *Sei ein Schritt bei der Iteration (4.33) eine Kontraktion in einer Skalarproduktnorm $\|\cdot\|_\Sigma$ in*

$$\underbrace{I\!R^m \times I\!R^m \times \cdots \times I\!R^m}_{k}$$

in dem restriktiven Sinne, dass $\|R_i\|_\Sigma < 1$, wobei

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \dots & \Sigma_{1k} \\ \Sigma_{12}^T & \Sigma_{22} & \dots & \Sigma_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{1k}^T & \Sigma_{2k}^T & \dots & \Sigma_{kk} \end{pmatrix}.$$

Wenn $\operatorname{im}\Psi(t)$ *in dem GGL-Ansatz (4.14) orthogonal zu $S_1(t)$ nach dem Skalarprodukt*

$$\langle x, x\rangle_{\Sigma_{11}} := (x^T \Sigma_{11} x)$$

gewählt wird, dann ist der entsprechende Iterationsschritt (4.42) auch eine Kontraktion in der Norm $\langle\cdot,\cdot\rangle_\Sigma$.

Beweis: S ei der Projektor P orthogonal nach dem Skalarprodukt $\langle\cdot,\cdot\rangle_{\Sigma_{11}}$. Unter den Annahmen ist $Q_{S_1^t}^{S_1} = \Psi(\Gamma^T B\Psi)^{-1}\Gamma^T B$ auch ein Orthoprojektor in dem euklidischen Raum $\left\{I\!R^m, \langle\cdot,\cdot\rangle_{\Sigma_{11}}\right\}$. Demzufolge sind die Projektoren

$$\begin{pmatrix} P & & & \\ & I_m & & \\ & & \ddots & \\ & & & I_m \end{pmatrix}, \quad \begin{pmatrix} I - Q_{S_1^t i}^{S_1} & & & \\ & I_m & & \\ & & \ddots & \\ & & & I_m \end{pmatrix},$$

im $\left\{\mathbb{R}^{m\times k}, \langle\cdot,\cdot\rangle_\Sigma\right\}$ orthogonal und besitzen Norm 1. Es folgt dann aus der Ungleichung (4.47)

$$\begin{aligned}
\left\|\overline{R}_i\right\|_\Sigma &\leq \left\| \begin{pmatrix} (PP_{S_1}^{S_1^t})_i & & & \\ & I_m & & \\ & & \ddots & \\ & & & I_m \end{pmatrix} \right\|_\Sigma \|R_i\|_\Sigma \\
&\leq \left\| \begin{pmatrix} P & & & \\ & I_m & & \\ & & \ddots & \\ & & & I_m \end{pmatrix} \right\|_\Sigma \left\| \begin{pmatrix} I - Q_{S_1^t i}^{S_1} & & & \\ & I_m & & \\ & & \ddots & \\ & & & I_m \end{pmatrix} \right\|_\Sigma \|R_i\|_\Sigma \\
&= \|R_i\|_\Sigma < 1,
\end{aligned}$$

und damit ist der Beweis erbracht.

Bemerkung 4.3.7

1. Die Aussage dieses Theorems hat die Einschränkung, dass der GGL-Ansatz davon abhängig ist, in welcher Norm Kontraktivität für (4.33) vorliegt. Der im Beweis konstruierte GGL-Ansatz kann als optimal bezeichnet werden. Die Aussage gilt ebenfalls für eine Klasse von GGL-Ansätzen, bei denen im $\Psi(t_i)$ nur in dem Maße von der optimalen Wahl abweicht, dass die Bedingung

   $$\left\| diag((PP_{S_1}^{S_1^t})_i, I_m, \ldots, I_m) R_i \right\| < 1$$

   noch erfüllt ist.

2. Die Orthogonalität zwischen im $\Psi(t)$ und $S_1(t)$ war gleichermaßen eine Voraussetzung im Theorem 4.2.9.

3. Das Theorem steht ohne Bezug zu analytischer Kontraktivität. Man kennt zwar Aussagen über die Kontraktivität der GGL-Formulierung (Theoreme 4.2.8, 4.2.9) und sogar von (4.16), ([Hanke et al., 1998], [März and Rodríguez Santiesteban, 1999]), sie sind aber insofern von wenig Nutzen in diesem Kontext, als die Kontraktionseigenschaft von (4.33) sich nicht mit der Kontraktivität von (4.22) in Verbindung setzen lässt. Es kann gezeigt werden, dass (4.20) kontraktiv auf im $PP_{S_1}^{S_1^t}$ ist,

wenn für (4.22) Kontraktivität auf $\widehat{M}_1(t)$ vorliegt. Dennoch ist dieser Sachverhalt nicht ausreichend, da die $PQ_{S_1^t}^{S_1 6}$-Komponente der diskreten Lösung z nicht trivial ist, (4.29). Man bemerke, dass sich sogar eine Kontraktivität von (4.22) auf $M_0(t)$ als nutzlos erweist, da auf $M_0(t)$ die $PQ_{S_1^t}^{S_1 6}$-Komponente null sein muss.

4. Die durchgeführte Analyse ähnelt sehr dem G-Stabilitätsbegriff, [Hairer and Wanner, 1991]. Die G-Stabilität kann als das Analogon zur B-Stabilität für die so genannten öne-legVerfahren (allgemeinere Klasse als die BDF-Verfahren) betrachtet werden. Sie verlangt, dass ihre Anwendung auf jede kontraktive Gleichung eine ebenfalls kontraktive Rekursion für die Variable Z_i, also nachdem man die Diskretisierung als ein Einschrittverfahren umgeschrieben hat, in einer ausgewählte Norm $\|\cdot\|_G$ darstellt. Es ist durchaus denkbar, nach dem gleichen Schema wie bei den Runge-Kutta-Verfahren und der B-Stabilität auch bei den BDF-Verfahren und der G-Stabilität vorzugehen. Auf diese Art und Weise kommt eine Definition von (P)-Kontraktivität der Anwendung eines BDF-Verfahrens auf eine ADGl zu Stande. Allerdings dürfte eine Aussage wie das Theorem atGGL-¿IRK(DAE)-Kontraktivität höchstens für die ein- und zweischrittige BDF-Verfahren gelten, da im Fall der GDGln die BDF-Methoden ab 3 Schritten nicht G-stabil sind. In dieser Arbeit wurde ein ander Weg gewählt, nämmlich der der Entkopplungsanalyse.

Wie in dem IRK(DAE)-Fall stellt sich die Frage, ob die Voraussetzung über die Orthogonalität zwischen $\operatorname{im}\Psi(t)$ und $S_1(t)$ nur beweistechnisch bedingt ist. Hier sprechen zudem die numerischen Experimente dafür, dass dies der Fall sein könnte. Die Abbildung 4.7 zeigt die numerischen Ergebnisse für die GGL-Formulierung des Beispieles 4.2.12 mit dem einzigen Unterschied, dass statt des Euler-Verfahrens die BDF-2 verwendet wurde.

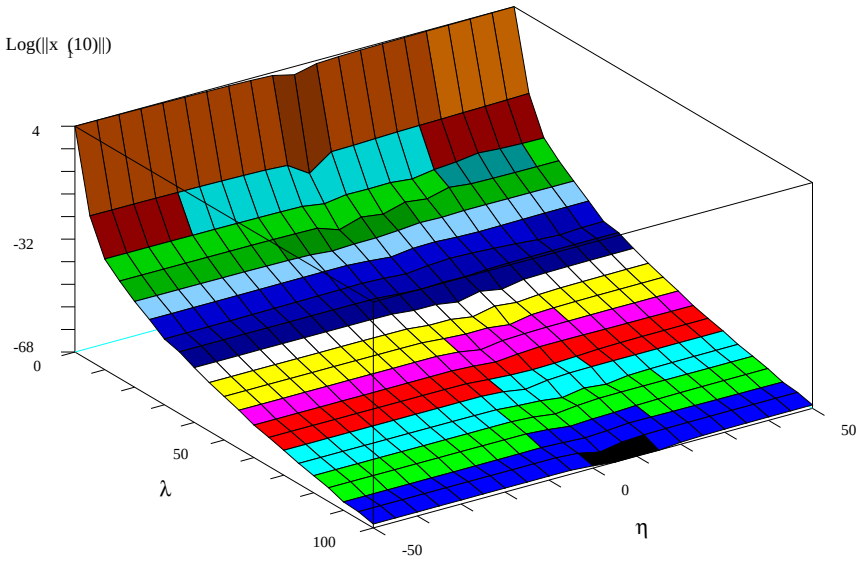

Abbildung 4.7: Numerische Ergebnisse für das Beispiel 3.0.8 unter Verwendung eines nicht-orthogonalen GGL-Ansatzes und der BDF-2. Das Bild zeigt den Logarithmus des Betrages von $x_1(10)$ für verschiedene Werte der Parameter λ und η. Die Schrittweite betrug $h = 0.1$ und $\theta = \pi/3$.

Schlusswort

Die Algebro-Differentialgleichungen stellen ein aktives Forschungsgebiet der Angewandten Mathematik dar, das neuartige Fragestellungen aufwirft. Insbesondere die Untersuchung asymptotischer Lösungseigenschaften sowohl aus analytischer als auch aus numerischer Sicht, lässt - im Gegensatz zu der Gewöhnlichen Differentialgleichungstheorie - noch viele Fragen offen. Diese Schrift versucht Klarheit in dieser Problematik zu schaffen.

Das Kapitel 2 leistet einen Beitrag zur Formalisierung einer Kontraktivitätstheorie und führt Verallgemeinerungen beziehungsweise Erweiterungen existierender Begriffe oder Aussagen ein. Die neuen Konzepte sind nicht an die Lösungsmannigfaltigkeit gebunden, sondern ermöglichen die Betrachtung der Kontraktivität außerhalb dieser Mannigfaltigkeit. Diese Eigenschaft ist besser für die Situation bei den numerischen Verfahren geeignet. Aus diesem Teil der Arbeit ist das Theorem 2.3.5 besonders hervorzuheben. Diese Aussage stellt die notwendigen Bedingungen, damit die Anwendung eines IRK(DAE)-Verfahrens auf eine ADGl (P)-kontraktiv ist. Die gestellten Bedingungen sind allgemeiner Art, stellen aber ein wichtiges Werkzeug für weitere Untersuchungen dar, wie in dem weiteren Verlauf der Schrift festzustellen ist. Das Kapitel schlägt unter Anderem eine Betrachtung der Kontraktivitätsungleichung in allen Komponenten der Lösung vor (Kontraktivität). Das resultierende Konzept zeigt sich als restriktiver als die P-Kontraktivität (die Kontraktivitätsungleichung wird nur für die P-Komponente verlangt), jedoch ermöglicht es Aussagen zu beweisen, die sich auf die gesamte Lösung beziehen. Allerdings fehlen hier Beispiele, die den Nutzen der Kontraktivität belegen können.

Im Kapitel 3 wurden die BDF- und IRK-Verfahren auf ihr asymptotisches Verhalten bei Index-2-ADGln untersucht. Die Analyse stützt sich einerseits auf Entkopplungen des Systems in den zugrundeliegenden Komponenten, andererseits auf die Index-Reduktion durch Differentiation. Die Entkopplungstechniken wurden dabei verfeinert beziehungsweise erweitert, was

zu Ergebnissen führte, die die bis jetzt bekannten Aussagen verbessern. Es wurde gezeigt, dass sich die Verfahren wie bei den Gewöhnlichen Differentialgleichungen verhalten, wenn einer der Räume $N_1(t)$ oder $S_1(t)$ konstant ist. Die Ergebnisse dieses Kapitels finden Anwendung für eine Klasse von ADGl-Systemen, die in der Modellierung elektrischer Netzwerke vorkommen: die so genannten MNA-Gleichungen.

Das letzte Kapitel der Arbeit befasst sich mit Stabilisierungstechniken, die verwendet werden sollen, wenn die Aussagen von Kapitel 3 nicht anwendbar sind. Der Ansatz nach Gear-Gupta-Leimkuhler (GGL) wird sachgerecht auf vollimplizite lineare Systeme verallgemeinert und es wird gezeigt, dass der Ansatz die Kontraktivität der ursprünglichen Aufgabe auf der Lösungsmannigfaltigkeit ($M_1(t)$) beibehält. Bei geeigneter Wahl (eine Orthogonalitätsbedingung muss erfüllt sein) dieses Ansatzes erhält man nunmehr ein System, das Kontraktivität auf $M_0(t)$ vorweist. Aus der letzten Aussage lässt sich Kontraktivität der IRK(DAE)-Verfahren ableiten. Unter den gleichen Orthogonalitätsbedingungen wurde ein Kontraktivitätssatz für die BDF-Verfahren angewandt auf die GGL-Formulierung gezeigt, allerdings weniger aussagekräftig als für die IRK(DAE)-Verfahren. Numerische Experimente deuten an, dass sowohl für IRK(DAE)- als auch für BDF-Verfahren die vorausgesetzte Orthogonalitätsbedingung nicht notwendig ist, sondern dass nur Transversalität, die immer bei der GGL-Formulierung gilt, ausreichend ist. Eine Untersuchung in dieser Richtung ist auf jeden Fall einen Versuch Wert, der dann die noch offene Frage nach einer numerischen Umsetzung des GGL-Ansatzes vereinfachen würde.

Literaturverzeichnis

[Ascher and Lin, 1993] Ascher, U. and Lin, P. (1993). Sequencial regularization methods for higher index daes with constraint singularities: I linear index-2 case. Technical Report 93-24, Department of computer science University of British Columbia, Vancouver, B.C., Canada.

[Brenan et al., 1989] Brenan, K. E., Campbell, S. L., and Petzold, L. R. (1989). *The Numerical Solution of Initial Value Problems in Ordinary Differential-Algebraic Equations*. North Holland Publishing Co.

[Campbell, 1985] Campbell, S. L. (1985). The numerical solution of higher index linear time varying singular systems of differential equations. *SIAM J. Sci. Stat. Comput.*, 6:334–348.

[Campbell, pear] Campbell, S. L. (to appear). A computational method for general higher index singular systems of differential equations. In *IMACS Transactions on Scientific Computing 1988*.

[Campbell and Moore, 1995] Campbell, S. L. and Moore, E. (1995). Constraint preserving integrators for general nonlinear higher index daes. *Numerische Mathematik*, (69):383–399.

[Celayeta, 1998] Celayeta, B. (1998). *Stability for differential algebraic equations*. PhD thesis, Departamento de Matemáticas e Informática, Universidad Pública de Navarra.

[Dekker and Verwer, 1984] Dekker, K. and Verwer, J. G. (1984). *Stability of Runge-Kutta methods for stiff nonlinear differential equations*. Number 2. Centre for Mathematics and Computer Science, North-Holland. CWI Monographs.

[Eich et al., 1990] Eich, E., Führer, C., Leimkuhler, B., and Reich, S. (1990). Stabilization and Projection Methods for Multibody Dynamics. Technical report, Helsinki University of Technology, Finland.

[Eich-Soellner and Führer, 1998] Eich-Soellner, E. and Führer, C. (1998). *Numerical methods in multibody dynamics*. B. G. Teubner, Stuttgart.

[Estévez Schwarz, 2000] Estévez Schwarz, D. (2000). *Consistent initialization for index-2 differential algebraic equations and its application to circuit simulation*. PhD thesis, Humboldt-Univ. Berlin.

[Estévez Schwarz and Tischendorf, 1998] Estévez Schwarz, D. and Tischendorf, C. (1998). Structural analysis for electric circuits and consequences for mna. Technical Report 98-21, Humboldt-Univ. zu Berlin, Institut für Mathematik.

[Gear, 1971] Gear, C. W. (1971). Simultaneous numerical solution of differential-algebraic equations. *IEEE Trans. Circuit Theory*, CT-18(1):89–95.

[Gear, 1988] Gear, C. W. (1988). Differential-algebraic equation index transformations. *SIAM J. Sci. Stat. Comput.*, 9(1).

[Gear et al., 1985] Gear, C. W., Gupta, G. K., and Leimkuhler, B. J. (1985). Automatic integration of the Euler-Lagrange equations with constraints. *J. Comp. Appl. Math.*, 12,13:77–90.

[Griepentrog and März, 1986] Griepentrog, E. and März, R. (1986). *Differential-Algebraic Equations and Their Numerical Treatment*. Teubner-Texte zur Mathematik No. 88. BSB B.G. Teubner Verlagsgesellschaft, Leipzig.

[Hairer et al., 1987] Hairer, E., Nørsett, S. P., and Wanner, G. (1987). *Solving Ordinary Differential Equations I: Nonstiff Problems*. Springer Series in Computational Mathematics 8. Springer-Verlag, Berlin, Heidelberg.

[Hairer and Wanner, 1991] Hairer, E. and Wanner, G. (1991). *Solving Ordinary Differential Equations II: Stiff and differential-algebraic problems*. Springer Series in Computational Mathematics 14. Springer-Verlag, Berlin, Heidelberg.

[Hale, 1980] Hale, J. K. (1980). *Ordinary Differential equation*. Krieger, Malabar, 2nd edition.

[Hanke, 1991] Hanke, M. (1991). On the asymptotic representation of a regularization approach to nonlinear semiexplicit higher-index differential-algebraic equations. *IMA Journal of Applied Mathematics*, (46):225–245.

[Hanke, 1992] Hanke, M. (1992). Regularizations of differential-algebraic equations revisited. Technical Report 92-19, Fachbereich Mathematik, Humboldt Universität zu Berlin.

[Hanke, 1994] Hanke, M. (1994). Asymptotic expansions for regularization methods of linear fully implicit differential-algebraic equations. *Journal for Analysis and its Applications*, 13(3):513–535.

[Hanke et al., 1998] Hanke, M., Izquierdo Macana, E., and März, R. (1998). On asymptotics in case of linear index-2 differential-algebraic equations. *SIAM J. Numer. Anal.*, 35(4):1326–1346.

[Higueras and Celayeta, 1999] Higueras, I. and Celayeta, B. G. (1999). Logarithmic norms for matrix pencils. *SIAM J. Matrix. Anal.*, 20(3):646–666.

[Higueras and Garcia-Celayeta, 1997] Higueras, I. and Garcia-Celayeta, B. (1997). Irk methods: a new approach. Preprint 14, Sección 1. Departamento de Matemática e informática, Universidad publica de Navarra.

[Kalachev and O'malley, 1996] Kalachev, L. V. and O'malley, R. E. (1996). Regularization of nonlinear differential-algebraic equations. *SIAM J. Math. Anal.*, 27(1):258–273.

[Macana, 1993] Macana, E. I. (1993). *Numerische Approximation von Algebro-Differentialgleichungen mit Index 2 mittels impliziter Runge-Kutta-Verfahren*. PhD thesis, Humboldt-Univ., Fachbereich Mathematik, Berlin.

[März, 1989] März, R. (1989). Index-2 differential–algebraic equations. *Results in Mathematics*, pages 148–171.

[März, 1992] März, R. (1992). Numerical methods for differential-algebraic equations. *Acta Numerica*, pages 141–198.

[März, 1994] März, R. (1994). Numerical stability criteria for differential-algebraic equations. *Int. Series of Num. Math.*, 117:73–81.

[März, 1995] März, R. (1995). On linear differential-algebraic equations and linearizations. *APNUM*, 18:267–292.

[März, 1997] März, R. (1997). EXTRA-ordinary differential equations. attempts to an analysis of differential-algebraic systems. Technical Report 97-8, Institut für Mathemaik, Humboldt-Univ. zu Berlin.

[März and Rodríguez Santiesteban, 1999] März, R. and Rodríguez Santiesteban, A. (1999). Analyzing the stability behaviour of DAE solutions and their approximations. Technical Report 99-2, Fachbereich Mathematik, Humboldt-Univ. zu Berlin.

[März, 1993] März, R. (1993). Canonical projectors for linear algebraic differential equations. Technical Report 17, Fachbereich Mathematik, Humboldt Universität zu Berlin.

[März, 1998] März, R. (1998). Analysis and numerics of daes. Special lecture, Humboldt Universität zu Berlin 1998.

[O'malley and Kalachev, 1994] O'malley, R. E. and Kalachev, L. V. (1994). Regularization of nonlinear differential-algebraic equations. *SIAM J. Math. Anal.*, 25(2):615–629.

[Petzold, 1986] Petzold, L. R. (1986). Order results for implicit runge-kutta methods applied to differential algebraic systems. *SIAM J. Numer. Anal.*, (23):837–852.

[Reich, 92] Reich, S. (92). Existence and uniqueness results for nonlinear differential-algebraic equations. In E. Griepentrog, M. Hanke, R. M., editor, *Berlin Seminar on Differential-Algebraic Equations*, pages 61–81. Fachbereich Mathematik der Humboldt-Universität zu Berlin, Berlin.

[Rheinboldt, 1984] Rheinboldt, W. C. (1984). Differential-algebraic systems as differential equations on manifolds. *Math. Comp.*, 43:473–482.

[Shampine, 1986] Shampine, L. F. (1986). Conservation laws and the numerical solution of ODEs. *Comp. and Math. with Appl., Part B.*, 12.

[Tischendorf, 1994] Tischendorf, C. (1994). On the stability of solutions of autonomous index-1 tractable and quasilinear index-2 tractable DAEs. *Circuits Systems Signal Process.*, 13(2-3):139–154.

[Tischendorf, 1996] Tischendorf, C. (1996). *Solution of index-2 differential algebraic equations and its application in circuit simulation.* PhD thesis, Humboldt-Univ. Berlin.

[Wensch et al., 1995] Wensch, J., Weiner, R., and Strehmel, K. (1995). Stability investigations for index-2 systems. Halle.

Lebenslauf

Name:	Antonio R. Rodríguez S.
Geburtstag,	
Geburtsort:	09.02.1968 in Havanna, Kuba
Staatsangehörigkeit:	Kuba
Familienstand:	verheiratet
Schulbildung:	
1972-1979	Grundschule, 0.-6. Klasse
1979-1982	Sekundarstufe I
1982-1985	Abitur
Ausbildung,	
Berufserfahrung:	
1985-1990	Diplom-Mathematik-Studium an der Universität von Havanna Fachrichtung: Angewandte Mathematik Durchschnittsnote: 4.83 (von 5) Diplomnote: 5 (von 5) (Numerische Mathematik)
1990-1994	Wissenschaftlicher Mitarbeiter am Institut für Informatik, Mathematik und Physik, Akademie der Wissenschaften Kubas
Forschungsgebiet:	Numerische Mathematik, Modellierung, Simulation
1994-1997	Promotionsstipendium im Rahmen des Graduiertenkollegs am Institut für Mathematik der Humboldt-Universität zu Berlin Promotionsthema: Asymptotisches Stabilitätsverhalten der numerischen Lösung von Algebro-Differential-Gleichungen
01.09.1997- 31.03.1998	Wissenschaftlicher Mitarbeiter am Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS) Berlin Arbeitsthema: Implementierung eines bekannten Algorithmus für die Berechnung konsistenter Anfangswerte von großen, schwach besetzten Algebro-Differential-Systemen der chemischen Anlagen-Prozess-Simulation
01.04.1998- 31.08.1999	Fortsetzung der Promotionsarbeit am Institut für Mathematik der Humboldt-Universität zu Berlin. Von 01.01.1999-22.05.1999 angestellt als wissenschaftlicher Mitarbeiter am o.g. Institut im Rahmen des BMBF-Projektes Mathematische Verfahren zur Lösung von Problemstellungen in Industrie und Wirtschaft”
01.09.99- Gegenwart:	Softwareentwickler bei der Firma DResearch Digital Media System GmbH

Selbständigkeitserklärung

Hiermit erkläre ich, die vorliegende Arbeit selbständig ohne fremde Hilfe verfaßt zu haben und nur die angegebene Literatur und Hilfsmittel verwendet zu haben.

Antonio R. Rodríguez S.
17. November 2000