

Die approbierte Originalversion dieser Dissertation ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (<http://www.ub.tuwien.ac.at>).

The approved original version of this thesis is available at the main library of the Vienna University of Technology (<http://www.ub.tuwien.ac.at/englweb/>).



TECHNISCHE
UNIVERSITÄT
WIEN
Vienna University of Technology

Iterierte Defektkorrektur für explizite und implizite Anfangswertprobleme erster und zweiter Ordnung

Dissertation

ausgeführt zum Zwecke der Erlangung des akademischen Grades
Doktor der technischen Wissenschaften

Betreuung:

Ao. Univ.-Prof. Dipl.-Ing. Dr. techn. Winfried Auzinger

Institutsnummer: E101

Institut für Analysis und Scientific Computing

eingereicht an der Technischen Universität Wien
Fakultät für Mathematik und Geoinformation

von

Mag. rer. nat. Dr. rer. nat. Martin Tutz

Matrikelnummer: 9625396

Diese Dissertation haben begutachtet:

.....
(Ao. Univ.-Prof. Dr. Winfried Auzinger)

.....
(Ao. Univ.-Prof. Dr. Mechthild Thalhammer)

Guntramsdorf, 21.2.2013

.....

Inhaltsverzeichnis

1	Explizite Differentialgleichungen erster Ordnung	6
1.1	Einleitung	6
1.2	Klassischer Defekt – IDeC	8
1.2.1	IDeC-EEUL	8
1.2.2	EEUL – Glattheit des Fehlers und seiner Differenzenquotienten	12
1.2.3	IDeC-EEUL – Fixpunktconvergenz	16
1.2.4	IDeC-EEUL – Konvergenz gegen die exakte Lösung	24
1.2.5	IDeC-Heun	26
1.2.6	IDeC-Heun – Fixpunktconvergenz	27
1.2.7	IDeC-EMR	29
1.2.8	IDeC-RK4	30
1.2.9	IDeC-IEUL	32
1.2.10	IDeC-ITR	33
1.2.11	IDeC-IMR	34
1.2.12	IDeC-Taylor2	35
1.2.13	IDeC-Taylor3	37
1.3	Integralmitteldefekt – IQDeC	38
1.3.1	IQDeC-EEUL	39
1.3.2	IQDeC-EEUL – Fixpunktconvergenz	40
1.3.3	IQDeC-EEUL – Konvergenz gegen die exakte Lösung	42
1.3.4	IQDeC-Heun	43
1.3.5	IQDeC-EMR	45
1.3.6	IQDeC-IEUL	45
1.3.7	IQDeC-ITR	46
1.3.8	IQDeC-IMR	48
1.4	Interpolierter Defekt – IPDeC	48
1.4.1	IPDeC-EEUL	49
1.4.2	IPDeC-EEUL – Fixpunktconvergenz	50
1.4.3	IPDeC-Heun	50
1.4.4	IPDeC-EMR	51
1.4.5	IPDeC-IEUL	52
1.4.6	IPDeC-ITR	52
1.4.7	IPDeC-IMR	53
2	Implizite Differentialgleichungen erster Ordnung	54
2.1	Klassischer Defekt – IIDeC	54
2.1.1	IIDeC-IEUL	54

2.1.2	IIDeC-IEUL – Fixpunktkonvergenz	55
2.1.3	IIDeC-ITR	58
2.2	Integralmitteldefekt – IIQDeC	59
2.2.1	IIQDeC-IEUL	59
2.2.2	IIQDeC-IEUL – Fixpunktkonvergenz	59
2.2.3	IIQDeC-ITR	63
2.3	Interpolierter Defekt – IIPDeC	64
2.3.1	IIPDeC-IEUL	64
2.3.2	IIPDeC-ITR	65
3	Explizite Differentialgleichungen zweiter Ordnung	66
3.1	Klassischer Defekt – IDeC2-DQ2	66
3.1.1	IDeC2-DQ2	66
3.1.2	IDeC2-DQ2-PC	69
3.2	Integralmitteldefekt – IQDeC2-DQ2	69
3.2.1	Diskretisierung der 2. Ableitung mittels lokaler Integration	69
3.2.2	IQDeC2-DQ2	71
3.2.3	IQDeC2-DQ2-PC	75
3.3	Interpolierter Defekt – IPDeC2-DQ2	76
3.3.1	IPDeC2-DQ2	76
3.3.2	IPDeC2-DQ2-PC	77
3.4	Lösung durch Rückführung auf ein System 1. Ordnung	77
3.4.1	IDeC2-Heun	78
3.4.2	IDeC2-EMR	80
3.4.3	IDeC2-Taylor2	80
3.4.4	IDeC2-ITR	83
3.4.5	IQDeC2-Heun	83
3.4.6	IQDeC2-EMR	85
3.4.7	IQDeC2-ITR	86
4	Implizite Differentialgleichungen zweiter Ordnung	88
4.1	Klassischer Defekt – IIDeC2-DQ2	88
4.2	Interpolierter Defekt – IIPDeC2-DQ2-Lobatto	90
4.3	Lösung durch Rückführung auf ein System 1. Ordnung	91
4.3.1	IIQDeC2-ITR	91
5	Fixpunkt-Kollokationsverfahren	94
5.1	IDeC-Fixpunkte	94
5.1.1	Euler-Kollokation	94
5.1.2	IMR-Kollokation	95
5.1.3	ITR-Kollokation	96
5.1.4	EMR-Kollokation	96
5.1.5	Heun-Kollokation	96
5.1.6	RK4-Kollokation	97
5.2	IQDeC-Fixpunkte	97
5.3	IPDeC-Fixpunkte	98
5.4	Kollokation für Differentialgleichungen 2. Ordnung	98

5.5	Konvergenz von (verallgemeinerten) Kollokationsverfahren	98
5.5.1	IEUL-Kollokation – Konvergenz	98
5.5.2	EEUL-Kollokation – Konvergenz	101
5.5.3	Heun-Kollokation – Konvergenz	103
5.5.4	Superkonvergenz	106
5.5.5	Gleichmäßige Beschränktheit des Kollokationspolynoms und dessen Ableitungen	106
6	Diskrete Defektkorrektur	111
6.1	Verfahren von Heun als diskrete Defektkorrektur	111
6.2	PC-Verfahren für Differentialgleichungen 2. Ordnung als diskrete Defektkor- rektur	113
7	Schlussbemerkungen und Ausblick	116
7.1	Grenzen der Defektkorrektur	116
7.1.1	IQDeC-RK4	116
7.1.2	IPDeC-RK4	117
7.2	Fehlende Algorithmen	118
7.3	Ausblick auf partielle Differentialgleichungen	118
7.3.1	Wärmeleitungsgleichung mit IQDeC-IEUL	119
7.3.2	Wellengleichung mit IQDeC2-IEUL	120
A	Maple-Code von IIPDeC2-DQ2-Lobatto	125

Danksagung

Mein besonderer Dank gilt Herrn Ao. Univ.-Prof. Dipl.-Ing. Dr. techn. Winfried Auzinger für die interessante Themenstellung sowie für die erstklassige Betreuung.

Ebenso gebührt mein Dank meiner Frau Tanja für das Korrekturlesen der vorliegenden Arbeit und das mir entgegengebrachte Verständnis.

Kurzfassung

Aufbauend auf bisher geleisteten Arbeiten auf dem Gebiet der iterierten Defektkorrektur, welche ausschließlich explizite Differentialgleichungen behandeln, wird eine Erweiterung dieser Idee auf bisher nicht betrachtete explizite sowie implizite Anfangswertprobleme erster und zweiter Ordnung mit allen drei Defektarten (klassischer, Integralmittel- und interpolierter Defekt) präsentiert. Der Schwerpunkt bei Anfangswertproblemen 2. Ordnung liegt in der Erstellung von Verfahren ohne Rückführung auf ein System von Differentialgleichungen 1. Ordnung. Für ausgewählte explizite und implizite Defektkorrekturverfahren sind Konvergenzbeweise beruhend auf „lokalen Überlegungen“ ausgeführt.

Weitere zentrale Punkte sind die Bestimmung und Analyse der Fixpunktverfahren, welche (verallgemeinerte) Kollokationsverfahren darstellen. Für eine Auswahl dieser Fixpunktverfahren sind die zugehörigen Konvergenzbeweise ausgeführt.

Ein kurzes Kapitel zeigt anhand der Analyse des Verfahrens von Heun und eines Predictor-Corrector-Verfahrens für Differentialgleichungen 2. Ordnung, dass Predictor-Corrector-Verfahren als (diskrete) Defektkorrekturverfahren interpretiert werden können.

Als Abschluss zeigt die Arbeit die Grenzen der Defektkorrektur für Integralmittel- und interpolierten Defekt auf und gibt mit der Anwendung der Defektkorrektur auf die (eindimensionale) Wärmeleitungs- und Wellengleichung einen Ausblick auf die Behandlung von partiellen Differentialgleichungen mittels Defektkorrektur.

Abstract

Based on previous works in the field of iterated defect correction, which treated explicit initial value problems only, this thesis expands these methods and applies them to untreated explicit as well as implicit first and second order initial value problems with the three different defects (classical, with quadrature and with interpolation). For second order initial value problems the focus is on methods avoiding reduction to a system of first order differential equations. Convergence proofs based on 'local considerations' of selected explicit and implicit defect correction methods are presented.

Other central topics are the determination and analysis of the fixed point methods, which are (generalized) collocation methods. For some of these fixed point methods convergence proofs are presented.

Using the analysis of Heun's method and a predictor corrector method for second order differential equations shows that predictor corrector methods may be interpreted as (discrete) defect correction methods.

Finally, the limitations of defect correction with quadrature and interpolation are demonstrated. By applying the defect correction method to the (one dimensional) heat and wave equation an outlook on the treatment of partial differential equations by means of defect correction is given.

Kapitel 1

Explizite Differentialgleichungen erster Ordnung

1.1 Einleitung

Die Idee der Defektkorrektur stammt vom argentinischen Astronomen P.E. Zadunaisky und wurde in [23] vorgestellt. Eine abstrakte Formulierung der iterierten Defektkorrektur bzw. des zugehörigen Fehlerschätzers findet sich in [22]¹, eine umfassende Beschreibung des Algorithmus sowie eine Konvergenzanalyse basierend auf der asymptotischen Entwicklung des globalen Fehlers in [15]. Diese und viele weitere Arbeiten, wie z.B. [3] behandeln explizite Randwertprobleme erster Ordnung, [2], [5] und [23] explizite Anfangswertprobleme erster Ordnung, [12] und [21] explizite Randwertprobleme zweiter Ordnung.

Die vorliegende Arbeit widmet sich der Erweiterung der Theorie auf implizite Anfangswertprobleme erster sowie auf explizite und implizite Anfangswertprobleme zweiter Ordnung mit allen drei bekannten Defektvarianten (klassischer, Integralmittel- und interpolierter Defekt).

- Kapitel 1 zeigt, neben einer kurzen Zusammenfassung der bisherigen Forschungsarbeit auf diesem Gebiet, Konvergenzbeweise für die Defektkorrektur mit klassischem Defekt und dem expliziten Euler-Verfahren sowie dem Verfahren von Heun als Basisverfahren basierend auf „lokalen Überlegungen“ ohne die vergleichsweise aufwändige asymptotische Entwicklung des globalen Fehlers (siehe [15]). Die Vorgangsweise wird durch einleitende Betrachtungen über die Glattheit des Fehlers und seiner Differenzenquotienten (beim expliziten Euler-Verfahren) motiviert.

Eine Vielzahl numerischer Experimente zeigt die Leistungsfähigkeit der Defektkorrektur mit allen „gängigen“ Einschrittverfahren erster und zweiter Ordnung sowie mit den klassischen Runge-Kutta-Verfahren als Basisverfahren. Hervorzuheben ist auch die Entwicklung des IDEC-Taylor-Algorithmus (siehe Abschnitt 1.2.12), welcher wesentlich ist für das Verständnis des Startschrittdefekts bei Anfangswertproblemen 2. Ordnung (siehe Kapitel 3.1.1).

- Kapitel 2 beschreibt ausführlich die Umformulierung der Defektkorrekturalgorithmen für die bisher nicht betrachtete direkte Behandlung impliziter Anfangswertprobleme

¹Die in der vorliegenden Arbeit vorgestellten Verfahren beziehen sich, in der Terminologie von [22], auf Version A. Dazu eng verwandte Verfahren vom Typ B werden in der vorliegenden Arbeit nicht betrachtet und sind in der Literaturliste nicht enthalten. Einen kurzen Überblick dazu gibt [9].

erster Ordnung (für alle drei in [2] vorgestellten Defektvarianten). Kernpunkt hierzu ist die Definition der Defekte für implizite Differentialgleichungen. Darüber hinaus sind für den klassischen Defekt und den Integralmitteldefekt die zugehörigen Konvergenzbeweise ausgeführt. Interessant erscheinen auch die entwickelten superkonvergenten IIPDeC-Algorithmen aus Abschnitt 2.3 am Ende des Kapitels.

- Kapitel 3 behandelt ausführlich die Anwendung der Defektkorrektur auf explizite Anfangswertprobleme zweiter Ordnung. Der Großteil des Kapitels widmet sich dabei der direkten Lösung solcher Probleme, d.h. ohne Rückführung auf Systeme erster Ordnung. Die dabei auftretenden Probleme betreffen vor allem die Umsetzung des Startschritts inkl. Defektdefinition, sowie den so genannten Sprungdefekt an den Intervallenden. Wiederum wurden für alle drei bekannten Defektarten Algorithmen entwickelt, welche die Leistungsfähigkeit der Defektkorrektur unter Beweis stellen. Für nicht steife Probleme werden Defektkorrekturverfahren mit Predictor-Corrector-Basisverfahren vorgestellt. Den Abschluss bilden Lösungsverfahren mit Rückführung auf Systeme erster Ordnung. Die Funktionsfähigkeit sämtlicher entwickelter Verfahren wird durch eine Reihe numerischer Experimente belegt.
- Kapitel 4 beleuchtet die Anwendung der Defektkorrektur auf implizite Anfangswertprobleme zweiter Ordnung. Dazu werden ein Verfahren mit klassischem Defekt und ein superkonvergentes Verfahren mit interpoliertem Defekt ohne Rückführung auf ein System erster Ordnung vorgestellt. Die wesentlichen Neuerungen dieses Kapitels betreffen wie in Kapitel 2 die Umformulierungen der Defekte für implizite Differentialgleichungen. Numerische Experimente zeigen wiederum die erfolgreiche Umsetzung der Algorithmen.
- Kapitel 5 widmet sich ausführlich den (verallgemeinerten) Fixpunkt-Kollokationsverfahren der verschiedenen Defektkorrekturverfahren. Für sämtliche Basisverfahren aus Kapitel 1 mit Ausnahme des Taylor-Verfahrens werden die zugehörigen Fixpunktbedingungen hergeleitet. Darüber hinaus sind für die explizite und die implizite Euler-Kollokation sowie für die Heun-Kollokation die zugehörigen Konvergenzbeweise ausgeführt.
- Kapitel 6 behandelt die diskrete Defektkorrektur und zeigt mit der Interpretation des Verfahrens von Heun und eines Predictor-Corrector-Verfahrens für Differentialgleichungen 2. Ordnung eine weitere Facette der Idee der Defektkorrektur.
- Kapitel 7 zeigt die Grenzen der Defektkorrektur auf und gibt mit der numerischen Lösung der eindimensionalen Wellen- und Wärmeleitungsgleichung einen Ausblick auf die Behandlung partieller Differentialgleichungen.

Sämtliche Konvergenzanalysen beschränken sich aus Transparenzgründen auf die Behandlung linearer Differentialgleichungen, obwohl sämtliche vorgestellte Algorithmen auch nichtlineare Probleme korrekt lösen, wie die numerischen Experimente am Ende der jeweiligen Abschnitte belegen. Diese zeigen anhand skalarer, aber nichtlinearer Probleme die globalen Fehler der Näherungslösungen der vorgestellten Defektkorrektur- bzw. (verallgemeinerten) Kollokationsverfahren bezüglich der exakten Lösung sowie die daraus berechneten (klassischen) Konvergenzordnungen. Die Übertragung der Beweise auf nichtlineare Fälle ist analog, wie in Anhang A des Preprints von [2] dargestellt, durchführbar. Der Schwerpunkt der

vorliegenden Arbeit liegt jedoch auf der Untersuchung einer großen Anzahl verschiedener Algorithmen.

1.2 Klassischer Defekt – IDeC

Die Abkürzung IDeC steht für *Iterated Defect Correction*. Im Folgenden soll der Algorithmus sowie die zugrunde liegende Idee der Defektkorrektur für Anfangswertprobleme erster Ordnung mit dem expliziten Euler-Verfahren als so genanntes Basisverfahren angelehnt an [2] kurz zusammengefasst wiedergegeben werden.

1.2.1 IDeC-EEUL

Gegeben sei ein Anfangswertproblem, also ein System von n gewöhnlichen Differentialgleichungen samt den zugehörigen Anfangsbedingungen

$$u'(t) = f(t, u(t)), \quad (1.1a)$$

$$u(t_0) = u_0, \quad (1.1b)$$

wobei u_0 den Startwert(-vektor) des Anfangswertproblems bezeichnet. Die exakte Lösung $u^*(t)$ ist somit eine Funktion von $[t_0, t_{end}] \rightarrow \mathbb{R}^n$. Weiters soll $f(t, u(t))$ Lipschitz-stetig im 2. Argument sein. Für lineare Anfangswertprobleme sei $f(t, u(t)) = C(t)u(t) + g(t)$ mit $C(t) \in \mathbb{R}^{n \times n}$ und $g(t) \in \mathbb{R}^n$.

Für die numerische Lösung von (1.1) wird für die vorliegende Arbeit das folgende Gitter zugrunde gelegt.

Definition 1.1. Das Gesamtintervall $[t_0, t_{end}] = [\mathbf{t}_0, \mathbf{t}_N]$ wird in N gleich lange Teil- bzw. Interpolationsintervalle² $\mathbf{I}_j = [\mathbf{t}_{j-1}, \mathbf{t}_j]$ zerlegt. Jedes dieser Interpolationsintervalle \mathbf{I}_j enthält inklusive der Randknoten $m + 1$ Gitterpunkte $t_{j,l}$ ($j = 1, \dots, N, l = 0, \dots, m$). Die relative Lage der Gitterpunkte $t_{j,l}$ innerhalb eines Intervalls \mathbf{I}_j wird durch $m + 1$ paarweise verschiedene Parameter c_l , $0 = c_0 < c_1 < \dots < c_m = 1$, festgelegt. Für die Lage der (inneren) Knoten folgt somit

$$t_{j,l} = \mathbf{t}_{j-1} + c_l \mathbf{h} \text{ mit } \mathbf{h} = \mathbf{t}_j - \mathbf{t}_{j-1} \text{ und } j = 1, \dots, N, l = 0, \dots, m. \quad (1.2)$$

Für die *innere Schrittweite* $h_{j,l}$ gilt

$$h_{j,l} = t_{j,l} - t_{j,l-1} = \mathbf{h} (c_l - c_{l-1}) \text{ mit } j = 1, \dots, N, l = 1, \dots, m. \quad (1.3)$$

Zur Veranschaulichung eines auf diese Art definierten Gitters siehe Abbildung 1.1.

Die numerische Lösung von (1.1) erfolgt auf dem in Definition 1.1 konstruierten Gitter zunächst mittels eines beliebigen (expliziten oder impliziten) Einschritt-Verfahrens niedriger Ordnung³, dem so genannten *Basisverfahren*. Die Erklärung in diesem Abschnitt beschränkt

²Im Allgemeinen Fall können diese auch verschieden lang sein. Dieser Aspekt ist jedoch nicht wesentlich und diene lediglich der Vereinfachung beim Erstellen der Maple-Programme.

³Hiermit sind Verfahren der Ordnung 1 (explizites oder implizites Euler-Verfahren) oder 2 (explizite oder implizite Mittelpunktsregel, Verfahren von Heun = „explizite Trapezregel“ oder implizite Trapezregel) gemeint. Ob die Wahl auf ein explizites oder implizites Basisverfahren fällt, hängt von der Steifheit des Problems (1.1) ab.

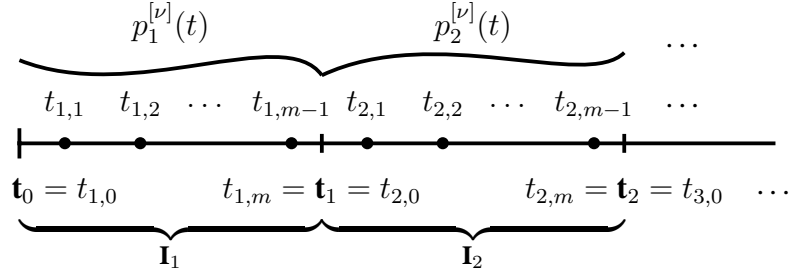


Abbildung 1.1: Lage der (inneren) Gitterpunkte $t_{j,l}$ in den Interpolationsintervallen \mathbf{I}_j .

sich zunächst auf das explizite Euler-Verfahren. Man diskretisiert wie folgt:

$$\frac{x_{j,l}^{[0]} - x_{j,l-1}^{[0]}}{h_{j,l}} = f\left(t_{j,l-1}, x_{j,l-1}^{[0]}\right), \quad j = 1, \dots, N, \quad l = 1, \dots, m, \quad (1.4a)$$

$$x_{1,0} = u_0 \quad (1.4b)$$

bzw.

$$x_{j,l}^{[0]} = x_{j,l-1}^{[0]} + h_{j,l} f\left(t_{j,l-1}, x_{j,l-1}^{[0]}\right), \quad (1.5a)$$

$$x_{1,0} = u_0 \quad (1.5b)$$

und interpoliert die so erhaltenen $x_{j,l}^{[0]}$ in jedem Intervall \mathbf{I}_j durch je ein Polynom $p_j^{[0]}(t)$ vom Grad $\leq m$, mittels derer man die stückweise stetige Polynomfunktion

$$p^{[0]}(t) := p_j^{[0]}(t), \quad t \in \mathbf{I}_j \quad (1.6)$$

definiert. Durch diese wird anschließend der so genannte *klassische Defekt*

$$d^{[0]}(t) = p^{[0]'}(t) - f(t, p^{[0]}(t)) \quad (1.7)$$

berechnet, mit welchem man das so genannte *Nachbarproblem* (inkl. Anfangsbedingung)

$$u'(t) = f(t, u(t)) + d^{[0]}(t), \quad (1.8a)$$

$$u(t_0) = u_0, \quad (1.8b)$$

aufstellen kann, dessen exakte Lösung $p^{[0]}(t)$ ist. Diskretisiert man (1.8) analog zu (1.1), erhält man

$$z_{j,l}^{[0]} = z_{j,l-1}^{[0]} + h_{j,l} \left(f\left(t_{j,l-1}, z_{j,l-1}^{[0]}\right) + d^{[0]}(t_{j,l-1}) \right), \quad (1.9a)$$

$$z_{1,0}^{[0]} = u_0 \quad (1.9b)$$

als Näherungslösung für das Nachbarproblem. Da die $d^{[0]}(t_{j,l-1})$ im Allgemeinen klein sind, liegt folgende heuristische Erwartung für den Fehlerschätzer $\varepsilon_{j,l}^{[0]}$ des tatsächlichen (globalen) Fehlers $e_{j,l}^{*,[0]}$ des zugrunde liegenden Basisverfahrens nahe⁴:

$$\varepsilon_{j,l}^{[0]} = \underbrace{z_{j,l}^{[0]}}_{\text{Näherungslösung des Nachbarproblems}} - \underbrace{p^{[0]}(t_{j,l})}_{\text{exakte Lösung des Nachbarproblems}} \approx \underbrace{x_{j,l}^{[0]}}_{\text{Näherungslösung des Originalproblems}} - \underbrace{u^*(t_{j,l})}_{\text{exakte Lösung des Originalproblems}} = e_{j,l}^{*,[0]}. \quad (1.10)$$

⁴Begründen kann man diese Erwartungshaltung damit, dass ein und dasselbe Lösungsverfahren bei zwei benachbarten (=„ähnlichen“) Differentialgleichungen bei gleichen Anfangsbedingungen „ähnlich“ genau arbeitet. D.h., dass die Diskretisierung bei der Lösung des Nachbarproblems bezüglich seiner exakten Lösung in etwa den gleichen Fehler verursacht, wie bei der Lösung der ungestörten Differentialgleichung.

Unter Berücksichtigung von $p^{[0]}(t_{j,l}) = x_{j,l}^{[0]}$ folgt

$$z_{j,l}^{[0]} - x_{j,l}^{[0]} \approx x_{j,l}^{[0]} - u^*(t_{j,l}), \quad (1.11)$$

was man als Ausgangsbasis zur (iterativen) Verbesserung der numerischen Näherungslösung von (1.1) heranziehen kann. Dazu ersetzt man in (1.11) die exakte Lösung $u^*(t_{j,l})$ durch $x_{j,l}^{[1]}$, das \approx -Zeichen durch ein $=$, und formt um

$$x_{j,l}^{[1]} := x_{j,l}^{[0]} - \left(z_{j,l}^{[0]} - x_{j,l}^{[0]} \right). \quad (1.12)$$

Allgemein gelangt man durch induktive Fortsetzung zu folgender Iteration: Zur Berechnung des Defekts werden in jedem Iterationsschritt die $x_{j,l}^{[\nu-1]}$ mittels $p^{[\nu-1]}(t)$ interpoliert. Man erhält den klassischen Defekt

$$d^{[\nu-1]}(t) = p^{[\nu-1]'}(t) - f(t, p^{[\nu-1]}(t)) \quad (1.13)$$

und das zugehörige Nachbarproblem

$$u'(t) = f(t, u(t)) + d^{[\nu-1]}(t), \quad (1.14a)$$

$$u(t_0) = u_0, \quad (1.14b)$$

dessen numerische Lösung zu

$$z_{j,l}^{[\nu-1]} = z_{j,l-1}^{[\nu-1]} + h_{j,l} \left(f(t_{j,l-1}, z_{j,l-1}^{[\nu-1]}) + d_{j,l-1}^{[\nu-1]} \right), \quad (1.15a)$$

$$z_{1,0}^{[\nu-1]} = u_0 \quad (1.15b)$$

mit

$$d_{j,l-1}^{[\nu-1]} \equiv d^{[\nu-1]}(t_{j,l-1}) = p^{[\nu-1]'}(t_{j,l-1}) - f(t_{j,l-1}, p^{[\nu-1]}(t_{j,l-1})) \quad (1.16)$$

führt. Mit der Lösung $z_{j,l}^{[\nu-1]}$ der Differenzgleichung (1.15a) wird die Näherungslösung $x_{j,l}^{[\nu-1]}$ mittels

$$x_{j,l}^{[\nu]} = x_{j,l}^{[0]} - \left(z_{j,l}^{[\nu-1]} - x_{j,l}^{[\nu-1]} \right) \quad (1.17)$$

korrigiert ($\nu = 1, 2, \dots$).

Der Fixpunkt der Folge der Näherungslösungen $x_{j,l}^{[\nu]}$, $\nu = 0, 1, \dots$ ist die (stückweise stetige) Kollokationslösung $p^{[C]}(t) := p_j^{[C]}(t)$ von (1.1), für die

$$d^{[C]}(t_{j,l-1}) = p^{[C]'}(t_{j,l-1}) - f(t_{j,l-1}, p^{[C]}(t_{j,l-1})) = 0, \quad l = 1, \dots, m \quad (1.18)$$

gilt. Kollokationsverfahren liefern auf Intervallen mit m (äquidistanten) Kollokationsknoten mindestens Konvergenzordnung m (siehe dazu Abschnitt 5 bzw. [17]). Es ist daher zu erwarten, dass die maximale Konvergenzordnung der iterierten Defektkorrektur bezüglich der exakten Lösung ebenfalls gleich m ist.

Bemerkung 1.2. Konvergenz gegen den Fixpunkt erhält man bei dem vorgestellten Verfahren mit *klassischem* Defekt auf dem beschriebenen Gitter dann, und nur dann, wenn die Schrittweite $h_{j,l}$ innerhalb der Intervalle \mathbf{I}_j konstant gehalten wird, d.h. $h \equiv h_{j,l} = \frac{t_j - t_{j-1}}{m}$. Eine Begründung, warum das Gitter trotzdem nicht äquidistant eingeführt wurde, liefert Abschnitt 1.3 für die Defektkorrektur mit Intergralmitteldefekt (IQDeC). Um Wiederholungen für IQDeC zu vermeiden werden die Basisverfahren und Nachbarprobleme in den Abschnitten 1.2.5 bis 1.2.13 ebenfalls für nicht äquidistante Gitter angesprochen.

h	EEUL	IDeC1	IDeC2	IDeC3	COLL
0,1	3,31E-02	1,84E-03	1,16E-05	6,75E-06	9,07E-06
0,05	1,65E-02	4,56E-04	1,91E-06	9,92E-07	1,14E-06
0,025	8,21E-03	1,13E-04	2,67E-07	1,33E-07	1,42E-07
0,0125	4,10E-03	2,83E-05	3,50E-08	1,72E-08	1,78E-08
0,1	1,01	2,01	2,60	2,77	3,00
0,05	1,00	2,01	2,84	2,90	3,00
0,025	1,00	2,00	2,93	2,95	3,00
0,0125					

Tabelle 1.1: Resultate zu Beispiel 1.3.

Beispiel 1.3. Gegeben sei das skalare nichtlineare Anfangswertproblem

$$u'(t) = \frac{u(t)}{1 + u^2(t)} - \sin(t) - \frac{\cos(t)}{1 + \cos^2(t)}, \quad (1.19a)$$

$$u(t_0 = 0) = 1, \quad (1.19b)$$

mit der exakten Lösung $u^*(t) = \cos(t)$. Löst man dieses auf einem äquidistanten Gitter mit 3 Kollokationsknoten (\Rightarrow Polynomgrad $m = 3$) gemäß Definition 1.1 ($c_1 = \frac{1}{3}$; $c_2 = \frac{2}{3}$; $c_3 = 1$) im Intervall $[t_0, t_{end}] = [\mathbf{t}_0, \mathbf{t}_N] = [0, 3]$, so erhält man die in Tabelle 1.1 gezeigten globalen Fehler bezüglich der exakten Lösung $e^{*,[\#]}(t_{end}) = p^{[\#]}(t_{end}) - u^*(t_{end})$, $\# = \nu, C$ und die daraus berechneten Konvergenzordnungen⁵ für $t_{end} = 3$. Details zur Schrittweite \mathbf{h} sind in der Tabelle ersichtlich. In der ersten Spalte sind die Schrittweiten aufgelistet, danach folgen die globalen Fehler für das Basisverfahren sowie für die einzelnen Iterationen der Defektkorrektur und das zugehörige Fixpunkt-Kollokationsverfahren (siehe dazu Abschnitt 5.1).

Sämtliche Berechnungen wurden mit Maple 14 durchgeführt. Die Rechengenauigkeit wurde in *Tools* \Rightarrow *Options* \Rightarrow *Precision* so gewählt, dass die gezeigten Ergebnisse de facto rundungsfehlerfrei sind⁶. Für die Lösung der in jedem Schritt zu lösenden nichtlinearen Gleichung wurde auf *fsolve* mit der Option *fulldigits* zurückgegriffen um volle Genauigkeit sicherzustellen.

Bemerkung 1.4. Bei sämtlichen Beispielen wird die globale (=passive) Lösungsstrategie angewandt. D.h., dass für das komplette Intervall $[t_0, t_{end}]$ zuerst die Lösung des Basisverfahrens $x_{j,l}^{[0]}$ berechnet wird. Erst danach werden hintereinander auf alle Intervalle $\mathbf{I}_j, j = 1, \dots, N$ die Iterationen des jeweiligen Defektkorrekturalgorithmus angewandt.

Im Gegensatz dazu wird bei der lokalen (=aktiven) Lösungsstrategie zuerst nur die Basislösung $x_{j,l}^{[0]}$ im Intervall \mathbf{I}_1 berechnet. Diese wird anschließend mittels Defektkorrektur bis zur gewünschten/maximal möglichen Ordnung verbessert. Erst danach folgt Intervall \mathbf{I}_2 . Als Anfangswert für dieses Intervall wird die bestmögliche Approximation des Endwertes von Intervall \mathbf{I}_1 herangezogen. Mit der gleichen Vorgangsweise wie für Intervall \mathbf{I}_1 (Basisverfahren plus Defektkorrektur) gelangt man zur bestmöglichen Approximation für Intervall \mathbf{I}_2 , usw..

⁵Für die numerische Bestimmung der Konvergenzordnung siehe [6].

⁶Ursprünglich war eine Implementierung in MATLAB geplant. Da man hier jedoch an die maximale Rechengenauigkeit von Double Precision gebunden ist, die hohen Ordnungen aber in „Reinform“ beobachtet werden sollten, erfolgte die (weitere) Software-Entwicklung zur Gänze in Maple 14.

h	EEUL	IDeC1	IDeC2	IDeC3	IDeC4	COLL
0,1	2,48E-02	1,03E-03	5,74E-06	8,61E-07	1,07E-07	1,11E-07
0,05	1,23E-03	2,55E-04	8,57E-07	5,26E-08	6,87E-09	7,04E-09
0,025	6,15E-03	6,37E-05	1,15E-07	3,25E-09	4,37E-10	4,43E-10
0,0125	3,07E-03	1,59E-05	1,49E-08	2,02E-10	2,76E-11	2,78E-11
0,1	1,01	2,01	2,74	4,03	3,96	3,98
0,05	1,00	2,00	2,89	4,02	3,97	3,99
0,025	1,00	2,00	2,94	4,01	3,98	4,00
0,0125	1,00	2,00	2,94	4,01	3,98	4,00

Tabelle 1.2: Resultate zu Beispiel 1.5.

h	EEUL	IDeC1	IDeC2	IDeC3	IDeC4	COLL
0,1	2,85E-02	7,66E-03	9,05E-04	2,21E-03	3,00E-03	1,91E-07
0,05	1,41E-02	3,96E-03	5,25E-04	1,11E-03	1,49E-03	1,21E-08
0,025	7,05E-03	2,01E-03	2,81E-04	5,53E-04	7,39E-04	7,62E-10
0,0125	3,52E-03	1,01E-03	1,45E-04	2,76E-04	3,69E-04	4,78E-11
0,1	1,01	0,95	0,78	1,00	1,01	3,98
0,05	1,00	0,98	0,90	1,00	1,01	3,99
0,025	1,00	0,99	0,95	1,00	1,00	4,00
0,0125	1,00	0,99	0,95	1,00	1,00	4,00

Tabelle 1.3: Resultate zu Beispiel 1.6 – Versagen der IDeC auf nicht äquidistantem Gitter.

Beispiel 1.5. Wie Beispiel 1.3 auf einem äquidistanten Gitter, jedoch mit 4 Kollokationsknoten. Ergebnisse siehe Tabelle 1.2.

Beispiel 1.6. Tabelle 1.3 zeigt mittels der Angabe aus Beispiel 1.3 auf einem „zufälligen“ *nicht* äquidistanten Gitter mit $m = 4$ Kollokationsknoten ($c_1 = 0,1234$; $c_2 = 0,5054$; $c_3 = 0,7134$; $c_3 = 1$) die Stagnation der Konvergenzordnungsfolge ab der ersten Defektkorrekturiteration, obwohl das zugehörige Fixpunkt-Kollokationsverfahren die höchstmögliche Konvergenzordnung (in diesem Fall 4) zeigt. Eine Begründung für das Versagen der klassischen Defektkorrektur und damit verbunden die Leistungsstärke des Integralmittel- und des interpolierten Defekts (siehe Abschnitte 1.3 und 1.4) liefert Abschnitt 1.2.2. Das Versagen von IDeC am nicht äquidistanten Gitter ist bereits in [2] dokumentiert.

1.2.2 EEUL – Glattheit des Fehlers und seiner Differenzenquotienten

In diesem Abschnitt wird anhand eines einfachen Beispiels plausibel gemacht, warum die Defektkorrektur mit klassischem Defekt auf Gittern mit nicht äquidistant verteilten Knoten nicht funktioniert. In Abschnitt 1.2.3 werden wir sehen, dass für den Beweis der Konvergenz solcher Defektkorrekturverfahren in jedem Iterationsschritt die asymptotischen Ordnungen

der entsprechenden dividierten Differenzen bzw. Differenzenquotienten⁷ der Lösungen der Fehlergleichung des vorangegangenen Schrittes benötigt werden. Die Ordnungen des Fehlers „vererben“ sich jedoch nur im äquidistanten Fall auf die Differenzenquotienten (höherer Ordnung).

Beispiel 1.7. Die Differentialgleichung

$$u'(t) = \cos(t) u(t) - \cos^2(t) - \sin(t), \quad (1.20a)$$

$$u(t_0 = 0) = 1 \quad (1.20b)$$

mit exakter Lösung $u^*(t) = \cos(t)$ wird mit dem expliziten Euler-Verfahren auf einem *äquidistanten* Gitter⁸ gelöst. Für den globalen Fehler

$$e_{j,l}^* = x_{j,l} - u^*(t_{j,l}) \quad (1.21)$$

der Euler-Lösung $x_{j,l}$ bezüglich der exakten Lösung $u^*(t_{j,l})$ sowie für den ersten und zweiten Differenzenquotienten

$$\partial e_{j,l}^* = \frac{e_{j,l}^* - e_{j,l-1}^*}{h_{j,l}}, \quad (1.22)$$

$$\partial^2 e_{j,l}^* = \frac{\partial_h e_{j,l}^* - \partial_h e_{j,l-1}^*}{\frac{1}{2}(h_{j,l} + h_{j,l-1})} \quad (1.23)$$

dieses Fehlers wurde Abbildung 1.2 erstellt. Tabelle 1.4 zeigt zusätzlich den (symmetrischen) dritten Differenzenquotienten

$$\partial^3 e_{j,l}^* = \frac{\partial_h^2 e_{j,l+1}^* - \partial_h^2 e_{j,l-1}^*}{h_{j,l} + h_{j,l-1}}. \quad (1.24)$$

Man sieht sowohl aus der Abbildung als auch aus der Tabelle, dass die Konvergenzordnung der Lösung im äquidistanten Fall bei der Bildung der Differenzenquotienten erhalten bleibt.

Beispiel 1.8. Löst man (1.20) auf einem sich periodisch wiederholenden *nicht äquidistanten* Gitter⁹, so erhält man bei analoger Vorgangsweise wie in Beispiel 1.7 die in Abbildung 1.3 und Tabelle 1.5 zusammengestellten Ergebnisse.

Bemerkung 1.9. In Tabelle 1.5 stagniert die Approximationsqualität des 2. Differenzenquotienten $\partial^2 e_{j,l}^*$. Man sieht, dass es ab dem zweiten Differenzenquotienten zu einer Ordnungsreduktion kommt. Für höhere Differenzenquotienten tritt sogar immer stärker ausgeprägte Divergenz für $h \rightarrow 0$ ein.

⁷Die k -te dividierte Differenz und der k -te Differenzenquotient unterscheiden sich um den Faktor $k!$.

⁸Das Gesamtintervall wurde für die gröbste Diskretisierung gemäß Definition 1.1 hierzu in 10 gleich lange Teilintervalle zerlegt. In jedem Teilintervall gilt für die relative Lage der Gitterpunkte: $c_0 = 0$; $c_1 = 0,25$; $c_2 = 0,5$; $c_3 = 0,75$; $c_4 = 1$.

⁹Das Gesamtintervall wurde für die gröbste Diskretisierung gemäß Definition 1.1 hierzu in 10 gleich lange Teilintervalle zerlegt. In jedem Teilintervall gilt für die relative Lage der Gitterpunkte: $c_0 = 0$; $c_1 = 0,1$; $c_2 = 0,5$; $c_3 = 0,7$; $c_4 = 1$.

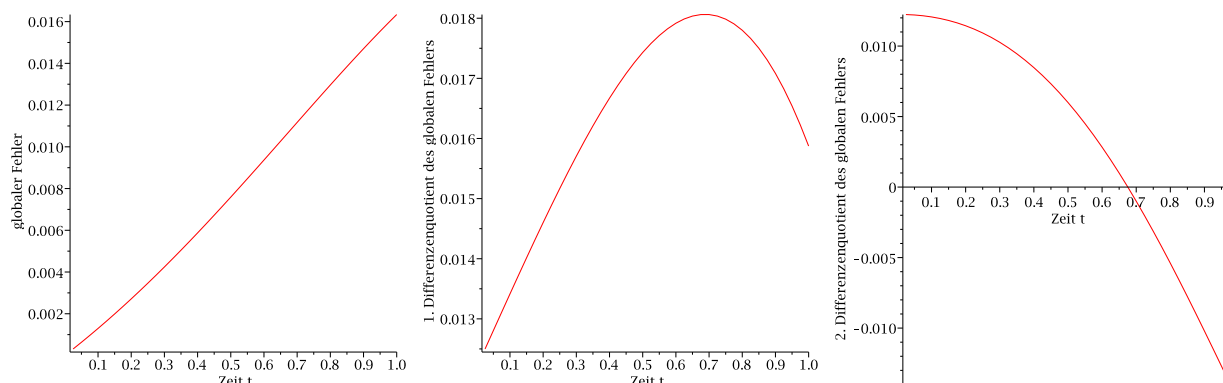


Abbildung 1.2: Darstellung des globalen Fehlers $e_{j,l}^*$ sowie des ersten und zweiten Differenzenquotienten $\partial e_{j,l}^*$ und $\partial^2 e_{j,l}^*$ (siehe (1.22) und (1.23)) auf einem äquidistanten Gitter mit $\mathbf{h} = 0, 1$ für Differentialgleichung (1.20).

\mathbf{h}	EEUL	DQ1	DQ2	DQ3
0,1	7,59E-03	1,74E-02	6,00E-03	1,34E-01
0,05	3,82E-03	8,79E-03	2,97E-03	5,77E-02
0,025	1,92E-03	4,41E-03	1,48E-03	2,91E-02
0,0125	9,60E-04	2,21E-03	7,36E-04	1,46E-02
0,1	0,99	0,99	1,01	0,98
0,05	1,00	0,99	1,01	0,99
0,025	1,00	1,00	1,00	0,99
0,0125	1,00	1,00	1,00	0,99

Tabelle 1.4: Die Tabelle zeigt die globalen Fehler $e_{j,l}^*$ gemäß (1.21) (oben) und deren erste, zweite und dritte Differenzenquotienten $\partial e_{j,l}^*$, $\partial^2 e_{j,l}^*$ und $\partial^3 e_{j,l}^*$ (siehe (1.22), (1.23) und (1.24)) sowie die daraus berechneten Konvergenzordnungen (unten) für Differentialgleichung (1.20) an der Stelle $t = 0, 5$ auf einem äquidistanten Gitter. In der ersten Spalte sind die Schrittweiten \mathbf{h} aufgelistet, danach folgen die Werte für das Euler-Verfahren sowie für die jeweiligen Differenzenquotienten.

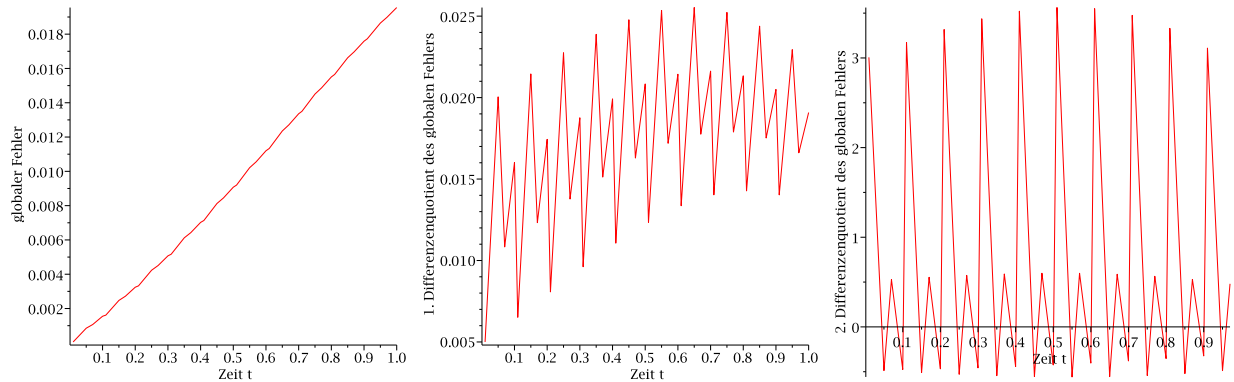


Abbildung 1.3: Darstellung des globalen Fehlers $e_{j,l}^*$ gemäß (1.21) sowie des ersten und zweiten Differenzenquotienten $\partial e_{j,l}^*$ und $\partial^2 e_{j,l}^*$ (siehe (1.22) und (1.23)) auf einem *nicht* äquidistanten Gitter mit $\mathbf{h} = 0, 1$ für Differentialgleichung (1.20).

\mathbf{h}	EEUL	DQ1	DQ2	DQ3
0,1	9,07E-03	2,08E-02	4,25E-01	8,36E±00
0,05	4,57E-03	1,05E-02	4,32E-01	1,72E+01
0,025	2,30E-03	5,29E-03	4,37E-01	3,48E+01
0,0125	1,15E-03	2,65E-03	4,37E-01	6,99E+01
0,1	0,99	0,99	-0,02	-1,04
0,05	0,99	0,99	-0,01	-1,02
0,025	1,00	1,00	-0,01	-1,01
0,0125				

Tabelle 1.5: Die Tabelle zeigt die globalen Fehler $e_{j,l}^*$ gemäß (1.21) (oben) und deren erste, zweite und dritte Differenzenquotienten $\partial e_{j,l}^*$, $\partial^2 e_{j,l}^*$ und $\partial^3 e_{j,l}^*$ (siehe (1.22), (1.23) und (1.24)) sowie die daraus berechneten Konvergenzordnungen (unten) für Differentialgleichung (1.20) an der Stelle $t = 0, 5$ auf einem *nicht* äquidistanten Gitter. In der ersten Spalte sind die Schrittweiten \mathbf{h} aufgelistet, danach folgen die Werte für das Euler-Verfahren sowie für die jeweiligen Differenzenquotienten.

1.2.3 IDeC-EEUL – Fixpunktconvergenz

Für die Defektkorrektur mit klassischem Defekt gilt folgende Konvergenzaussage:

Satz 1.10. (*Fixpunktconvergenz von IDeC-EEUL*) Löst man das Anfangswertproblem vom Typ (1.1) auf einem äquidistanten Gitter (d.h. $h \equiv h_{j,l}$) gemäß Definition 1.1 mittels iterierter Defektkorrektur mit klassischem Defekt und dem expliziten Euler-Verfahren als Basisverfahren, so besitzt der Fehler $e_{j,l}^{[\nu]}$ der Defektkorrektur-Lösungen $x_{j,l}^{[\nu]}$ der einzelnen Iterationen ν bezüglich des Fixpunktes $p_{j,l}^{[C]}$ (=Kollokationslösung auf demselben Gitter, Interpolationsgrad m) die Konvergenzordnung

$$\left\| e_{j,l}^{[\nu]} \right\| = \left\| x_{j,l}^{[\nu]} - p_{j,l}^{[C]} \right\| = \mathcal{O}(h^{\nu+1}), \quad \nu = 0, 1, \dots \quad (1.25)$$

Bemerkung 1.11. Satz 1.10 wurde bereits in [15] mittels der vergleichsweise aufwändigen asymptotischen Entwicklung des globalen Fehlers bewiesen. Der hier erbrachte Beweis beruht wie beschrieben auf einer lokalen Überlegung und ist somit als deutlich einfacher einzustufen. Die Beweisformulierung zeigt auch, wie die in Abschnitt 1.2.2 gezeigte Glattheit des Fehlers in das Argument einfließt.

Abstrakt formulierte Beweisorganisation, Herleitung der Fehlergleichungen für Satz 1.10 (lineare Analysis):

In einer an die Artikel [22] bzw. [7] angelehnten abstrakten (Operator-)Schreibweise soll die Beweisorganisation für die Fixpunktconvergenz (d.h. Konvergenz gegen die Kollokationslösung, siehe Abschnitt 5) bei klassischem Defekt für Differentialgleichungen erster Ordnung skizziert werden. Zur besseren Lesbarkeit wird auf Interpolationsoperatoren und die Angabe anderer einschränkender Bedingungen verzichtet. Die nachfolgenden Identitäten sind diskret, d.h. über dem zugrundeliegenden (Integrations-)Gitter zu verstehen. Im Anschluss wird aufbauend auf diesem abstrakten Schema der eigentliche Beweis geführt.

Identifiziert man Gleichung (1.1) mit einem linearen Differentialoperator \mathcal{L} und einer linearen rechten Seite \mathcal{F} , erhält man

$$\mathcal{L}u = \mathcal{F}(u). \quad (1.26)$$

Die numerische Lösung erhält man mittels des Basisverfahrens (1.4a). Dabei wird der Differentialoperator \mathcal{L} durch den Differenzenoperator L ersetzt,

$$Lx^{[0]} = \mathcal{F}(x^{[0]}). \quad (1.27)$$

Mit der so erhaltenen Näherungslösung $x^{[0]}$ wird gemäß (1.7) der Defekt

$$d^{[0]} = \mathcal{L}x^{[0]} - \mathcal{F}(x^{[0]}) \quad (1.28)$$

definiert. Mit Hilfe des Defektes erhält man für das diskretisierte Nachbarproblem (vgl. (1.9))

$$Lz^{[0]} = \mathcal{F}(z^{[0]}) + d^{[0]} \quad (1.29)$$

die Lösung $z^{[0]}$. Der Fehlerschätzer in Bezug auf den Fixpunkt lautet

$$z^{[0]} - x^{[0]} \approx x^{[0]} - p^{[C]}, \quad (1.30)$$

wobei $p^{[C]}$ die Kollokationslösung von (1.26) bezeichnet¹⁰. Für den Defekt der Kollokationslösung gilt

$$d^{[C]} = \mathcal{L}p^{[C]} - \mathcal{F}(p^{[C]}) = 0. \quad (1.31)$$

- Ausgehend von (1.30) wird gemäß (1.12) die Näherungslösung zu

$$x^{[1]} := x^{[0]} - (z^{[0]} - x^{[0]}) \quad (1.32)$$

korrigiert.

Wir definieren zunächst den Fehler bezüglich des Fixpunktes

$$e^{[\nu]} := x^{[\nu]} - p^{[C]}, \quad (1.33)$$

wobei $\nu = 0, 1, \dots$ den Iterationsschritt (beginnend mit 0 für das Basisverfahren) indiziert.

- $\nu = 0$: Die Fehlergleichung für das Basisverfahren erhält man, indem man von (1.27) auf beiden Seiten den Term $Lp^{[C]}$ abzieht und Gleichung (1.31) addiert sowie die Linearität der Differentialgleichung ausnützt:

$$Lx^{[0]} - Lp^{[C]} = \mathcal{F}(x^{[0]}) - Lp^{[C]} + (\mathcal{L}p^{[C]} - \mathcal{F}(p^{[C]})), \quad (1.34)$$

$$Le^{[0]} = \mathcal{F}_H(e^{[0]}) + \underbrace{\mathcal{L}p^{[C]} - Lp^{[C]}}_{=:r^{[0]}}. \quad (1.35)$$

\mathcal{F}_H bezeichnet dabei den *homogenen Anteil* von \mathcal{F} , d.h. für $\mathcal{F}(x) = Cx + g$ ist $\mathcal{F}_H(x) = Cx$.

- $\nu = 1$: Für die Herleitung der Fehlergleichung der gemäß (1.32) verbesserten Näherungslösung geht man wie folgt vor. Man subtrahiert auf beiden Seiten von (1.32) die Kollokationslösung $p^{[C]}$ und wendet den Operator L an. Somit erhält man

$$\underbrace{Lx^{[1]} - Lp^{[C]}}_{=:Le^{[1]}} = Lx^{[0]} - Lz^{[0]} + \underbrace{Lx^{[0]} - Lp^{[C]}}_{=:Le^{[0]}}, \quad (1.36)$$

wodurch man mittels Einsetzen von (1.27), (1.29) und (1.28)

$$Le^{[1]} = \mathcal{F}(x^{[0]}) - \mathcal{F}(z^{[0]}) - \mathcal{L}x^{[0]} + \mathcal{F}(x^{[0]}) + Le^{[0]} \quad (1.37)$$

erhält. Weiteres Umstellen der Terme und Ausnutzen der Linearität der Differentialgleichung ergibt

$$Le^{[1]} = \mathcal{F}(x^{[0]}) - (\mathcal{F}(z^{[0]}) - \mathcal{F}(x^{[0]})) - \mathcal{L}x^{[0]} + Le^{[0]}, \quad (1.38)$$

$$Le^{[1]} = \mathcal{F}\left(\underbrace{x^{[0]} - (z^{[0]} - x^{[0]})}_{=:x^{[1]}}\right) - \mathcal{L}x^{[0]} + Le^{[0]}. \quad (1.39)$$

¹⁰Vergleicht man (1.30) mit (1.11), so sieht man, dass die exakte Lösung u^* von letzterem Zusammenhang einfach durch die Kollokationslösung $p^{[C]}$ ersetzt wurde, was möglich ist, da die Kollokationslösung „sehr nahe“ bei der exakten Lösung liegt und die Gültigkeit der Fehlerschätzung (lediglich) durch eine heuristische Erwartungshaltung begründet wird.

Subtrahiert man davon die Gleichung (1.31), führt dies, wieder unter Ausnützung der Linearität, zur Fehlergleichung für die verbesserte Lösung

$$Le^{[1]} = \mathcal{F}(x^{[1]}) - \mathcal{L}x^{[0]} + Le^{[0]} + \mathcal{L}(p^{[C]}) - \mathcal{F}(p^{[C]}), \quad (1.40)$$

$$Le^{[1]} = \mathcal{F}_H(e^{[1]}) + \underbrace{Le^{[0]} - \mathcal{L}e^{[0]}}_{=:r^{[1]}}. \quad (1.41)$$

- Die induktive Fortsetzung der oben angestellten Überlegungen ergibt

$$Le^{[\nu]} = \mathcal{F}_H(e^{[\nu]}) + \underbrace{Le^{[\nu-1]} - \mathcal{L}e^{[\nu-1]}}_{=:r^{[\nu]}} \quad (1.42)$$

bzw. ausformuliert

$$\frac{e_{j,l}^{[\nu]} - e_{j,l-1}^{[\nu]}}{h} = C_{j,l-1} e_{j,l-1}^{[\nu]} + \underbrace{\frac{e_{j,l}^{[\nu-1]} - e_{j,l-1}^{[\nu-1]}}{h} - e_{j,l-1}^{[\nu-1]'}}_{=:r_{j,l}^{[\nu]}}. \quad (1.43)$$

Zu zeigen bleibt, dass die Restglieder $r^{[0]}$ bzw. $r^{[\nu]}$ in (1.35) und (1.42) die lokalen Ordnungen $\mathcal{O}(h)$ bzw. $\mathcal{O}(h^{\nu+1})$ für $\nu = 1, 2, \dots$ haben. Zusammen mit einem Stabilitätsargument und dem Lemma von Gronwall (siehe z.B. [4]) erhält man für das verwendete Basisverfahren die lokalen Ordnungen auch global.

Bemerkung 1.12. Die Operatoren auf der rechten Seite von (1.26) und (1.27) unterscheiden sich für Differentialgleichungen erster Ordnung nicht. Für Gleichungen höherer Ordnung ist eine Differenzierung in \mathcal{F} und F notwendig, da \mathcal{F} i.A. Ableitungen von u enthält, die bei der Lösung durch das Basisverfahren durch Differenzenquotienten ersetzt werden müssen. Diese Diskretisierung wird durch die Bezeichnung F zum Ausdruck gebracht. Auf diesem abstrakten Niveau kann das vorgestellte „Rezept“ unter Berücksichtigung der eben beschriebenen Erweiterung grundsätzlich auch für Konvergenzbeweise von Differentialgleichungen höherer Ordnung und auch für den nichtlinearen Fall herangezogen werden.

IDeC-EEUL – Fixpunktkonvergenz – Beweis

In diesem Abschnitt soll der eben beschriebene abstrakte Formalismus „mit Leben erfüllt werden“.

Beweis von Satz 1.10 mit 4 Knoten, davon 3 Kollokationsknoten (lineare Analysis):

Es folgt eine detaillierte Beschreibung der Analyse für das lineare Anfangswertproblem

$$u'(t) = C(t)u(t) + g(t), \quad (1.44a)$$

$$u(t_0) = u_0, \quad (1.44b)$$

wobei $C(t) \in \mathbb{R}^{n \times n}$ und $g(t) \in \mathbb{R}^n$ hinreichend glatte Funktionen sind.

- $\nu = 0$: Man geht unter Zuhilfenahme der Abkürzungen $C_{j,l} \equiv C(t_{j,l})$, $e_{j,l}^{[0]} \equiv e^{[0]}(t_{j,l})$, etc. von Fehlergleichung (1.35) aus,

$$\frac{e_{j,l}^{[0]} - e_{j,l-1}^{[0]}}{h} = C_{j,l-1} e_{j,l-1}^{[0]} + \underbrace{p_{j,l-1}^{[C]'} - \frac{p_{j,l}^{[C]} - p_{j,l-1}^{[C]}}{h}}_{=r_{j,l}^{[0]}}. \quad (1.45)$$

Bei 4 Knoten, davon 3 Kollokationsknoten, ist der Maximalgrad des Kollokationspolynoms $p^{[C]}(t)$ gleich 3. D.h. es gilt $\frac{d^k p^{[C]}(t)}{dt^k} = 0$ für $k > 3$. Für die Ordnung des Restgliedes gilt

$$\begin{aligned} r_{j,l}^{[0]} &= p_{j,l-1}^{[C]'} - \frac{p_{j,l}^{[C]} - p_{j,l-1}^{[C]}}{h} \\ &= p_{j,l-1}^{[C]'} - \frac{p_{j,l-1}^{[C]} + h p_{j,l-1}^{[C]'} + \frac{h^2}{2} p_{j,l-1}^{[C]''} + \mathcal{O}(h^3) - p_{j,l-1}^{[C]}}{h} \\ &= -\frac{h}{2} p_{j,l-1}^{[C]''} + \mathcal{O}(h^2) \\ &= \mathcal{O}(h). \end{aligned} \quad (1.46)$$

Zusammen mit einem Stabilitätsargument und dem Lemma von Gronwall (siehe z.B. [4]) ist (1.25) für $\nu = 0$ und 4 Knoten bewiesen.

Bemerkung 1.13. Zu zeigen bleibt (an dieser Stelle erstmalig) die gleichmäßige Beschränktheit des Kollokationspolynoms $p^{[C]}(t)$ und dessen Ableitungen (siehe dazu Abschnitt 5.5.5).

- $\nu = 1$: Die Fehlergleichung (1.41) lautet ausformuliert

$$\frac{e_{j,l}^{[1]} - e_{j,l-1}^{[1]}}{h} = C_{j,l-1} e_{j,l-1}^{[1]} + \underbrace{\frac{e_{j,l}^{[0]} - e_{j,l-1}^{[0]}}{h} - e_{j,l-1}^{[0]'}}_{=r_{j,l}^{[1]}}. \quad (1.47)$$

Schreibt man die stückweise definierten Fehlerpolynome $e_j^{[\nu]}(t)$ (mit $e^{[\nu]}(t) := e_j^{[\nu]}(t)$) für jedes Intervall \mathbf{I}_j in Newtonscher Darstellung an, so erhält man

$$\begin{aligned} e_j^{[\nu]}(t) &= \sum_{l=0}^3 b_{j,l}^{[\nu]} \cdot \omega_l(t) \quad \text{mit} \quad \omega_0(t) \equiv 1 \quad \text{und} \quad \omega_l(t) = \prod_{i=0}^{l-1} (t - t_{j,i}) \quad (1.48) \\ &\quad \text{für} \quad l = 1, 2, 3 \quad \text{und} \quad \nu = 0, 1, \dots \end{aligned}$$

Die Koeffizienten $b_{j,l}^{[\nu]}$ sind die dividierten Differenzen der Fehlerpolynomwerte $e_{j,l}^{[\nu]}$ an

den Gitterpunkten $t_{j,l}$ (siehe z.B. [6, Kapitel 4])

$$b_{j,0}^{[\nu]} = e_{j,0}^{[\nu]}, \quad (1.49)$$

$$b_{j,1}^{[\nu]} = \frac{e_{j,1}^{[\nu]} - e_{j,0}^{[\nu]}}{t_{j,1} - t_{j,0}} = \frac{e_{j,1}^{[\nu]} - e_{j,0}^{[\nu]}}{h} =: e_j^{[\nu]}[t_{j,0}, t_{j,1}], \quad (1.50)$$

$$b_{j,2}^{[\nu]} = \frac{\frac{e_{j,2}^{[\nu]} - e_{j,1}^{[\nu]}}{t_{j,2} - t_{j,1}} - \frac{e_{j,1}^{[\nu]} - e_{j,0}^{[\nu]}}{t_{j,1} - t_{j,0}}}{t_{j,2} - t_{j,0}} = \frac{e_{j,2}^{[\nu]} - 2e_{j,1}^{[\nu]} + e_{j,0}^{[\nu]}}{2h^2} =: e_j^{[\nu]}[t_{j,0}, t_{j,1}, t_{j,2}], \quad (1.51)$$

$$\begin{aligned} b_{j,3}^{[\nu]} &= \frac{\frac{\frac{e_{j,3}^{[\nu]} - e_{j,2}^{[\nu]}}{t_{j,3} - t_{j,2}} - \frac{e_{j,2}^{[\nu]} - e_{j,1}^{[\nu]}}{t_{j,2} - t_{j,1}}}{t_{j,3} - t_{j,1}} - \frac{\frac{e_{j,2}^{[\nu]} - e_{j,1}^{[\nu]}}{t_{j,2} - t_{j,1}} - \frac{e_{j,1}^{[\nu]} - e_{j,0}^{[\nu]}}{t_{j,1} - t_{j,0}}}{t_{j,2} - t_{j,0}}}{t_{j,3} - t_{j,0}} = \frac{e_{j,3}^{[\nu]} - 3e_{j,2}^{[\nu]} + 3e_{j,1}^{[\nu]} - e_{j,0}^{[\nu]}}{6h^3} \\ &=: e_j^{[\nu]}[t_{j,0}, t_{j,1}, t_{j,2}, t_{j,3}]. \end{aligned} \quad (1.52)$$

O.B.d.A wird nun für $\nu = 1$ der erste (lokale) Schritt eines Teilpolynoms $e_j^{[0]}(t)$ (von $t_{j,0}$ nach $t_{j,1}$) analysiert¹¹. Wertet man (1.48) an $t_{j,1}$ und $t_{j,0}$ und die daraus berechnete erste Ableitung $e_j^{[0]'}(t)$ an der Stelle $t_{j,0}$ aus und setzt alle Terme in das Restglied von (1.47) ein, erhält man

$$\begin{aligned} r_{j,1}^{[1]} &= h b_{j,2}^{[0]} - 2h^2 b_{j,3}^{[0]} \\ &= h e_j^{[0]}[t_{j,0}, t_{j,1}, t_{j,2}] - 2h^2 e_j^{[0]}[t_{j,0}, t_{j,1}, t_{j,2}, t_{j,3}]. \end{aligned} \quad (1.53)$$

Für die Ordnung des Restgliedes bleibt zu zeigen, dass die dividierten Differenzen 2. Ordnung $e_j^{[0]}[t_{j,0}, t_{j,1}, t_{j,2}]$ und 3. Ordnung $e_j^{[0]}[t_{j,0}, t_{j,1}, t_{j,2}, t_{j,3}]$ von Werten $e_{j,l}^{[0]} \equiv e_j^{[0]}(t_{j,l})$ bzw. deren Differenzenquotienten $\frac{e_{j,l}^{[0]} - e_{j,l-1}^{[0]}}{h}$, deren Ordnung $\mathcal{O}(h)$ ist, ebenfalls $\mathcal{O}(h)$ sind. Für die Bestimmung der Ordnung von $e_j^{[0]}[t_{j,0}, t_{j,1}, t_{j,2}] = b_{j,2}^{[0]}$ geht man durch Einsetzen der rechten Seite von (1.45) in (1.51) und Umformen wie folgt vor:

$$\begin{aligned} b_{j,2}^{[0]} &= \frac{1}{2} \frac{\frac{e_{j,2}^{[0]} - e_{j,1}^{[0]}}{h} - \frac{e_{j,1}^{[0]} - e_{j,0}^{[0]}}{h}}{h} \\ &= \frac{1}{2} \left(\underbrace{\frac{C_{j,1} e_{j,1}^{[0]} - C_{j,0} e_{j,0}^{[0]}}{h}}_{=: {}_1 b_{j,2}^{[0]}} - \frac{1}{2} \underbrace{\left(p_{j,1}^{[C]''} - p_{j,0}^{[C]''} \right)}_{=: {}_2 b_{j,2}^{[0]}} + \mathcal{O}(h) \right). \end{aligned} \quad (1.54)$$

Anwenden des diskreten Analogons der Produktregel auf den ersten Term in der Klammer ergibt

$$\begin{aligned} {}_1 b_{j,2}^{[0]} &= \frac{C_{j,1} e_{j,1}^{[0]} \overbrace{-C_{j,1} e_{j,0}^{[0]} + C_{j,1} e_{j,0}^{[0]} - C_{j,0} e_{j,0}^{[0]}}{=0}}{h} \\ &= C_{j,1} \frac{e_{j,1}^{[0]} - e_{j,0}^{[0]}}{h} + \frac{C_{j,1} - C_{j,0}}{h} e_{j,0}^{[0]} \\ &= \mathcal{O}(1) \mathcal{O}(h) + \mathcal{O}(1) \mathcal{O}(h) = \mathcal{O}(h), \end{aligned} \quad (1.55)$$

¹¹Die Argumentation für die weiteren Teilschritte funktioniert in analoger Weise, lediglich mit abweichenden technischen Details (Koeffizienten).

da bei hinreichend glattem $C(t)$ unter Ausnutzung des Mittelwertsatzes für Matrizen

$$C_{j,1} = C_{j,0} + h \int_0^1 C'(t_{j,0} + \vartheta h) d\vartheta \quad (1.56)$$

und

$$\left\| \frac{C_{j,1} - C_{j,0}}{h} \right\| \leq \max_{t_{j,l} \leq \tau \leq t_{j,l-1}} \|C'(\tau)\| = \mathcal{O}(1) \quad (1.57)$$

gilt¹².

Für den zweiten Term liefert ebenfalls der Mittelwertsatz das gewünschte Ergebnis:

$${}^2b_{j,2}^{[0]} = p_{j,1}^{[C]''} - p_{j,0}^{[C]''} = h \int_0^1 p^{[C]'''}(t_{j,0} + \vartheta h) d\vartheta = \mathcal{O}(h). \quad (1.58)$$

Es folgt somit $b_{j,2}^{[0]} = \mathcal{O}(h)$. Analoges Vorgehen liefert die Ordnung der dritten dividier- ten Differenz $b_{j,3}^{[0]}$ ausgehend von (1.52) durch Einsetzen von (1.45) und anschließender Umformung

$$\begin{aligned} b_{j,3}^{[0]} &= \frac{1}{6} \frac{\frac{e_{j,3}^{[0]} - e_{j,2}^{[0]}}{h} - \frac{e_{j,2}^{[0]} - e_{j,1}^{[0]}}{h}}{h} - \frac{\frac{e_{j,2}^{[0]} - e_{j,1}^{[0]}}{h} - \frac{e_{j,1}^{[0]} - e_{j,0}^{[0]}}{h}}{h} \\ &= \frac{1}{6} \left(\underbrace{B_{j,2} \frac{e_{j,2}^{[0]} - e_{j,1}^{[0]}}{h} - B_{j,1} \frac{e_{j,1}^{[0]} - e_{j,0}^{[0]}}{h}}_{=: {}^1b_{j,3}^{[0]}} + \underbrace{\frac{B_{j,2} - B_{j,1}}{h} e_{j,1}^{[0]} - \frac{B_{j,1} - B_{j,0}}{h} e_{j,0}^{[0]}}_{=: {}^2b_{j,3}^{[0]}} \right. \\ &\quad \left. - \frac{h p_{j,2}^{[C]''} - 2p_{j,1}^{[C]''} + p_{j,0}^{[C]''}}{2h^2} - \frac{h^2 p_{j,2}^{[C]'''} - 2p_{j,1}^{[C]'''} + p_{j,0}^{[C]'''}}{6h^2} \right). \end{aligned} \quad (1.59)$$

Die Terme ${}^1b_{j,3}^{[0]}$ bis ${}^4b_{j,3}^{[0]}$ werden nun in gleicher Weise wie bei $b_{j,2}^{[0]}$ getrennt behandelt:

$$\begin{aligned} {}^1b_{j,3}^{[0]} &= \frac{B_{j,2} \frac{e_{j,2}^{[0]} - e_{j,1}^{[0]}}{h} - B_{j,2} \frac{e_{j,1}^{[0]} - e_{j,0}^{[0]}}{h} + B_{j,2} \frac{e_{j,1}^{[0]} - e_{j,0}^{[0]}}{h} - B_{j,1} \frac{e_{j,1}^{[0]} - e_{j,0}^{[0]}}{h}}{h} \\ &= B_{j,2} \frac{e_{j,2}^{[0]} - 2e_{j,1}^{[0]} + e_{j,0}^{[0]}}{h^2} + \frac{B_{j,2} - B_{j,1}}{h} \frac{e_{j,1}^{[0]} - e_{j,0}^{[0]}}{h} \\ &= \mathcal{O}(1) \mathcal{O}(h) + \mathcal{O}(1) \mathcal{O}(h) = \mathcal{O}(h), \end{aligned} \quad (1.60)$$

$$\begin{aligned} {}^2b_{j,3}^{[0]} &= \frac{\frac{B_{j,2} - B_{j,1}}{h} e_{j,1}^{[0]} - \frac{B_{j,1} - B_{j,0}}{h} e_{j,1}^{[0]} + \frac{B_{j,1} - B_{j,0}}{h} e_{j,1}^{[0]} - \frac{B_{j,1} - B_{j,0}}{h} e_{j,0}^{[0]}}{h} \\ &= \frac{B_{j,2} - 2B_{j,1} + B_{j,0}}{h^2} e_{j,1}^{[0]} + \frac{B_{j,1} - B_{j,0}}{h} \frac{e_{j,1}^{[0]} - e_{j,0}^{[0]}}{h} \\ &= \mathcal{O}(1) \mathcal{O}(h) + \mathcal{O}(1) \mathcal{O}(h) = \mathcal{O}(h). \end{aligned} \quad (1.61)$$

¹²Die Norm bezeichnet hier die zur Vektornorm gehörige Matrixnorm.

Für ${}^3b_{j,3}^{[0]}$ gilt: Da $p^{[C]}(t)$ ein Polynom 3. Grades ist, hat die zweite Ableitung $p^{[C]''}(t)$ den Grad 1. Der symmetrische zweite Differenzenquotient von $p^{[C]''}(t)$ liefert für die Berechnung der 4. Ableitung von $p^{[C]}(t)$ ein exaktes Ergebnis. Dieses ist jedoch für ein Polynom von besagtem Grad Null. Da eine analoge Argumentation für ${}^4b_{j,3}^{[0]}$ gilt, ist insgesamt $b_{j,3}^{[0]} = \mathcal{O}(h)$.

Setzt man die Ordnungen von $b_{j,2}^{[0]}$ und $b_{j,3}^{[0]}$ in (1.53) ein, erhält man

$$r_{j,1}^{[1]} = h \mathcal{O}(h) + h^2 \mathcal{O}(h) = \mathcal{O}(h^2), \quad (1.62)$$

wodurch zusammen mit einem Stabilitätsargument in Kombination mit dem Lemma von Gronwall (siehe z.B. [4]) Beziehung (1.25) für $\nu = 1$ und 4 Knoten bewiesen ist.

- Für $\nu = 2$ gilt die (ausformulierte) Fehlergleichung (1.42)

$$\frac{e_{j,l}^{[2]} - e_{j,l-1}^{[2]}}{h} = C_{j,l-1} e_{j,l-1}^{[2]} + \underbrace{\frac{e_{j,l}^{[1]} - e_{j,l-1}^{[1]}}{h} - e_{j,l-1}^{[1]'}}_{=r_{j,l}^{[2]}}. \quad (1.63)$$

Für die Bestimmung der Ordnung des Restgliedes $r_{j,l}^{[2]}$ stellt man das Fehlerpolynom $e_j^{[1]}(t)$ gemäß (1.48) dar. Man erhält für den ersten Schritt ($l = 1$) im j -ten Teilintervall

$$r_{j,1}^{[2]} = h b_{j,2}^{[1]} - 2h^2 b_{j,3}^{[1]}. \quad (1.64)$$

Für $b_{j,2}^{[1]}$ gilt mit Hilfe von (1.47) und (1.48)

$$\begin{aligned} b_{j,2}^{[1]} &= \frac{1}{2} \frac{\frac{e_{j,2}^{[1]} - e_{j,1}^{[1]}}{h} - \frac{e_{j,1}^{[1]} - e_{j,0}^{[1]}}{h}}{h} \\ &= \frac{1}{2} \frac{B_{j,1} e_{j,1}^{[1]} + \frac{e_{j,2}^{[0]} - e_{j,1}^{[0]}}{h} - e_{j,1}^{[0]'} - \left(B_{j,0} e_{j,0}^{[1]} + \frac{e_{j,1}^{[0]} + e_{j,0}^{[0]}}{h} - e_{j,0}^{[0]'} \right)}{h} \\ &= \frac{1}{2} \frac{B_{j,1} e_{j,1}^{[1]} - B_{j,0} e_{j,0}^{[1]}}{h} + \frac{1}{2} \frac{h b_{j,2}^{[0]} + h^2 b_{j,3}^{[0]} - \left(h b_{j,2}^{[0]} - 2h^2 b_{j,3}^{[0]} \right)}{h} \\ &= \frac{1}{2} \underbrace{\frac{B_{j,1} e_{j,1}^{[1]} - B_{j,0} e_{j,0}^{[1]}}{h}}_{=: {}^1b_{j,2}^{[1]}} + \frac{1}{2} \underbrace{3h b_{j,3}^{[0]}}_{=: {}^2b_{j,2}^{[1]}}. \end{aligned} \quad (1.65)$$

Analysiert man ${}^1b_{j,2}^{[1]}$ gleich wie ${}^1b_{j,2}^{[0]}$, so erhält man unter Berücksichtigung der bereits bekannten Ergebnisse $e_{j,l}^{[1]} = \mathcal{O}(h^2)$ bzw. $\frac{e_{j,l}^{[1]} - e_{j,l-1}^{[1]}}{h} = \mathcal{O}(h^2)$ für ${}^1b_{j,2}^{[1]} = \mathcal{O}(h^2)$. Für ${}^2b_{j,2}^{[1]}$ gilt

$${}^2b_{j,2}^{[1]} = 3h b_{j,3}^{[0]} = 3h \mathcal{O}(h) = \mathcal{O}(h^2). \quad (1.66)$$

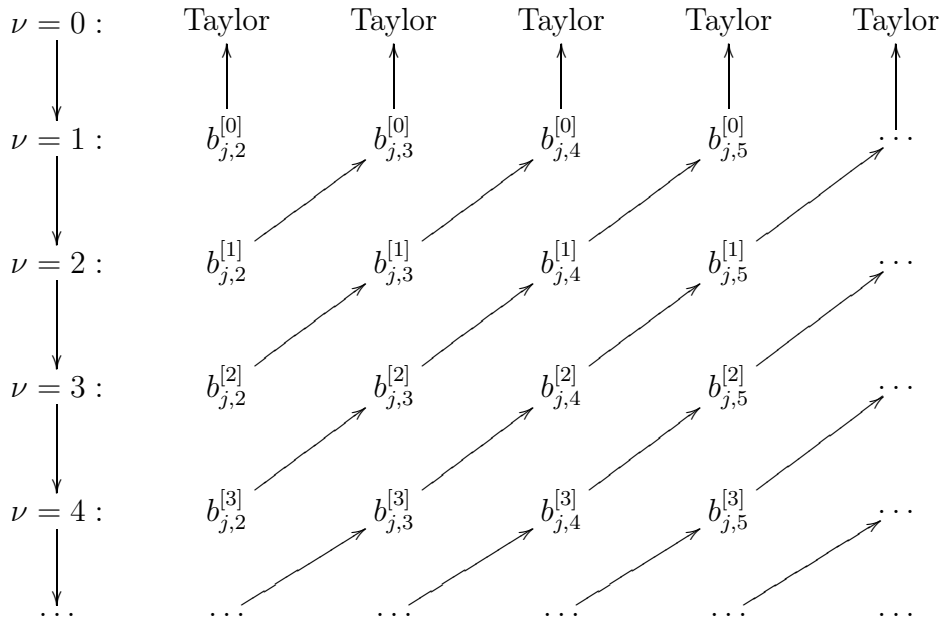
Für die verbleibende Analyse von $b_{j,3}^{[1]}$ erhält man bei gleicher Vorgangsweise wie für $b_{j,3}^{[0]}$

$$\begin{aligned}
b_{j,3}^{[1]} &= \frac{1}{6} \frac{\frac{e_{j,3}^{[1]} - e_{j,2}^{[1]}}{h} \frac{e_{j,2}^{[1]} - e_{j,1}^{[1]}}{h} - \frac{e_{j,2}^{[1]} - e_{j,1}^{[1]}}{h} \frac{e_{j,1}^{[1]} - e_{j,0}^{[1]}}{h}}{h} \\
&= \frac{1}{6} \left(\underbrace{\frac{B_{j,2} \frac{e_{j,2}^{[1]} - e_{j,1}^{[1]}}{h} - B_{j,1} \frac{e_{j,1}^{[1]} - e_{j,0}^{[1]}}{h}}{h}}_{=: 1b_{j,3}^{[1]}} + \underbrace{\frac{\frac{B_{j,2} - B_{j,1}}{h} e_{j,1}^{[1]} - \frac{B_{j,1} - B_{j,0}}{h} e_{j,0}^{[1]}}{h}}_{=: 2b_{j,3}^{[1]}} \right). \quad (1.67)
\end{aligned}$$

Analoge Analyse von $1b_{j,3}^{[1]}$ und $2b_{j,3}^{[1]}$ zu $1b_{j,3}^{[0]}$ und $2b_{j,3}^{[0]}$ (siehe (1.60) und (1.61)) ergibt die Ordnungen $\mathcal{O}(h^2)$. Daraus folgt für $b_{j,3}^{[1]} = \mathcal{O}(h^2)$.

Setzt man die Ordnungen in (1.64) ein, erhält man $r_{j,1}^{[1]} = \mathcal{O}(h^3)$. In Kombination mit einem Stabilitätsargument und dem Lemma von Gronwall (siehe z.B. [4]) ist (1.25) für $\nu = 2$ und 4 Knoten bewiesen.

Induktive Fortsetzung des gezeigten Weges liefert ausgehend von (1.43) bei Erweiterung auf m Knoten einen kompletten Beweis durch vollständige Induktion (für ν beliebig groß). Das folgende abstrakte Schema erläutert die Abhängigkeiten, die man erhält, wenn man die Fehlerterme mittels Newtonscher Polynomdarstellung analysiert und sukzessive für die darin enthaltenen dividierten Differenzen einsetzt.



- $\nu = 0$: Wie in (1.46) gezeigt, gelangt man mit einer Taylorreihenentwicklung des Restgliedes (1.43) zur erwarteten Ordnung des Basisverfahrens bezüglich des Fixpunktes.
- $\nu = 1$: In der Newtondarstellung des Restgliedes für m Knoten ($m = 3$ siehe (1.53)) kommen alle dividierten Differenzen des Basisverfahrens von $b_{j,2}^{[0]}$ bis $b_{j,m}^{[0]}$ vor. Deren Ordnungen müssen (getrennt) analysiert werden.

h	EEUL	IDeC1	IDeC2	...	IDeC12	IDeC13
0,1	2,48E-02	1,03E-03	5,63E-06	...	2,36E-20	5,84E-21
0,05	1,23E-03	2,56E-04	8,50E-07	...	8,64E-24	6,19E-25
0,025	6,15E-03	6,37E-05	1,15E-07	...	1,57E-27	4,78E-29
0,0125	3,07E-03	1,59E-05	1,49E-08	...	2,29E-31	3,25E-33
0,1	1,01	2,01	2,73	...	11,41	13,20
0,05	1,00	2,00	2,88	...	12,42	13,66
0,025	1,00	2,00	2,95	...	12,74	13,84
0,0125						

Tabelle 1.6: Ordnungen des Basisverfahrens und der einzelnen IDeC-Iterationen bezüglich der Fixpunktkollokationslösung $p^{[C]}(t)$ zu Beispiel 1.15. Man beachte die letzten beiden Spalten (12. und 13. Iteration).

Bemerkung 1.14. Würde man nur die Ordnung für den ersten Schritt ($\nu = 1$) zeigen, könnte man die höheren dividierten Differenzen aus (1.53) als Linearkombination von 2. dividierten Differenzen darstellen. Somit müsste man die Ordnungen der 3. bis m -ten dividierten Differenzen nicht zeigen. Dies ist jedoch nicht zielführend im Hinblick auf den allgemeinen Induktionsschritt, da man für die nächste Iteration jeweils die Ordnung aller um 1 höheren dividierten Differenzen des Fehlers benötigt. D.h., für $\nu = 2$ benötigt man die Ordnungen aller dividierten Differenzen ab der 3. dividierten Differenz des Fehlers des Basisverfahrens ($\nu = 0$).

- $\nu - 1 \rightarrow \nu$: Sieht man das Schema an, erkennt man, dass man für den ν -ten Schritt die Ordnungen aller dividierten Differenzen von $b_{j,3}^{[\nu-1]}$ bis $b_{j,\nu+1}^{[\nu-1]}$ des Fehlers benötigt.

Beispiel 1.15. Tabelle 1.6 zeigt für Beispiel 1.5, dass man die entsprechenden Ordnungen der numerischen Lösung bezüglich des Fixpunktes (=Kollokationslösung $p^{[C]}(t)$) auch für $\nu > m - 1$ erhält.

1.2.4 IDeC-EEUL – Konvergenz gegen die exakte Lösung

Der Vollständigkeit halber sei auch die Konvergenz der numerischen Lösungen gegen die exakte Lösung kurz diskutiert. Der folgende Satz folgt auch aus der Konvergenz der Kollokation.

Satz 1.16. (*IDeC-Konvergenz gegen die exakte Lösung*) Löst man das Anfangswertproblem vom Typ (1.1) auf einem äquidistanten Gitter (d.h. $h \equiv h_{j,l}$) mit $m + 1$ inneren Knoten gemäß Definition 1.1 mittels iterierter Defektkorrektur mit klassischem Defekt, so erfüllt der Fehler $e_{j,l}^{*,[\nu]}$ der Defektkorrektur-Lösungen $x_{j,l}^{[\nu]}$ der einzelnen Iterationen bezüglich der exakten Lösung $u^*(t_{j,l})$ folgende Bedingung

$$\left\| e_{j,l}^{*,[\nu]} \right\| = \left\| x_{j,l}^{[\nu]} - u^*(t_{j,l}) \right\| = \mathcal{O}(h^{\nu+1}), \quad \nu = 0, 1, \dots, m - 1. \quad (1.68)$$

IDeC-EEUL – Abstrakt formulierte Beweisorganisation für Satz 1.16 (lineare Analysis):

Die Beweisorganisation erfolgt analog zu jener in Abschnitt 1.2.3. Die folgenden Identitäten gelten wiederum über dem diskreten Gitter. Es sei q jenes Polynom, welches die exakte Lösung u^* an den Knoten des verwendeten Gitters interpoliert. Daraus folgt für den Defekt der exakten Lösung

$$d^* = \mathcal{L}u^* - \mathcal{F}(u^*) = \mathcal{L}(u^* - q + q) - \mathcal{F}(q) = \mathcal{L}(q) - \mathcal{F}(q) + \mathcal{L}(u^* - q) = 0. \quad (1.69)$$

Die Fehlergleichung für den Fehler $e^{*,[\nu]} = x^{[\nu]} - u^* = x^{[\nu]} - q$ der ν -ten Iteration erhält man wie folgt: Ausgehend von der Verbesserung der Näherungslösung (1.12)

$$x^{[\nu]} = x^{[0]} - (z^{[\nu-1]} - x^{[\nu-1]}) \quad (1.70)$$

subtrahiert man auf beiden Seiten die exakte Lösung $u^* = q$ (an den Kollokationsknoten) und wendet den Differenzenoperator L an,

$$Lx^{[\nu]} - Lq = Lx^{[0]} - Lz^{[\nu-1]} + Lx^{[\nu-1]} - Lq. \quad (1.71)$$

Nützt man auf beiden Seiten die Linearität von L aus, und setzt man auf der rechten Seite (1.27) und (1.29) mit (1.28) für die $(\nu - 1)$ -te Iteration ein, erhält man

$$Le^{*,[\nu]} = \mathcal{F}(x^{[0]}) - \mathcal{F}(z^{[\nu-1]}) - \mathcal{L}x^{[\nu-1]} + \mathcal{F}(x^{[\nu-1]}) + Le^{*,[\nu-1]}, \quad (1.72)$$

$$Le^{*,[\nu]} = \underbrace{\mathcal{F}(x^{[0]}) - (\mathcal{F}(z^{[\nu-1]}) - \mathcal{F}(x^{[\nu-1]}))}_{=\mathcal{F}(x^{[\nu]})} - \mathcal{L}x^{[\nu-1]} + Le^{*,[\nu-1]}. \quad (1.73)$$

Zur Fehlergleichung gelangt man, indem man (1.69) addiert und die Linearität von \mathcal{F} ausnutzt

$$Le^{*,[\nu]} = \mathcal{F}(x^{[\nu]}) - \mathcal{L}x^{[\nu-1]} + Le^{*,[\nu-1]} + \mathcal{L}(q) - \mathcal{F}(q) + \mathcal{L}(u^* - q), \quad (1.74)$$

$$Le^{*,[\nu]} = \mathcal{F}(e^{*,[\nu]}) + Le^{*,[\nu-1]} - \mathcal{L}e^{*,[\nu-1]} + \mathcal{L}(u^* - q). \quad (1.75)$$

Man sieht, dass sich (1.42) und (1.75) in ihrer Struktur nur durch den letzten Term unterscheiden, welcher den Fehler der Ableitungen von exakter Lösung u^* und der Interpolierenden q wiedergibt. Für $Lu = u'$ und Interpolationsgrad m gilt

$$\|u^* - q\| = \mathcal{O}(h^{m+1}) \quad (1.76)$$

und

$$\|\mathcal{L}(u^* - q)\| = \mathcal{O}(h^m), \quad (1.77)$$

siehe [6, Kapitel 4].

In den folgenden Abschnitten wird gezeigt, dass die Defektkorrektur auch mit anderen Basisverfahren funktioniert. Der Schwerpunkt ist dabei auf den Einbau des Defekts bzw. der Defekte in das jeweilige Nachbarproblem gelegt. Für Erweiterungen des Defektkorrekturformalismus auf allgemeine Runge-Kutta-Basisverfahren siehe auch [13]. Ersetzt man das explizite Euler-Verfahren durch ein Basisverfahren höherer Ordnung, so erhält man schnellere Konvergenz gegen den Fixpunkt. Die jeweiligen Fixpunkte für die einzelnen Verfahren werden in Kapitel 5 diskutiert.

h	Heun	IDeC1	IDeC2	Heun-COLL
0,1	4,30E-04	1,16E-07	7,85E-08	7,84E-08
0,05	1,06E-04	7,22E-09	4,90E-09	4,90E-09
0,025	2,65E-05	4,51E-10	3,06E-10	3,06E-10
0,0125	6,60E-06	2,82E-11	1,91E-11	1,91E-11
0,1	2,01	4,00	4,00	4,00
0,05	2,01	4,00	4,00	4,00
0,025	2,00	4,00	4,00	4,00
0,0125				

Tabelle 1.7: Resultate zu Beispiel 1.17.

1.2.5 IDeC-Heun

Ersetzt man unter Beibehaltung sämtlicher anderer algorithmischer Details¹³ das Basisverfahren (1.4a) und das Nachbarproblem (1.15a) durch geeignete Identitäten, so erhält man ein Defektkorrekturverfahren, dessen Konvergenzordnung bei jeder Iteration um 2 steigt, d.h. man erhält die Ordnungfolge 2,4,6,... Die Konvergenz der Heun-Kollokation (=Fixpunkt des Defektkorrekturverfahrens mit Heun-Basisverfahren) wird ausführlich in Kapitel 5 diskutiert.

Um den Einbau des Defekts in das Nachbarproblem verstehen zu können, wird das Basisverfahren (Verfahren von Heun) als 1-stufiges Runge-Kutta-Verfahren angeschrieben:

$$X_1^{[0]} = x_{j,l-1}^{[0]} + h_{j,l} f(t_{j,l-1}, x_{j,l-1}^{[0]}), \quad (1.78a)$$

$$x_{j,l}^{[0]} = x_{j,l-1}^{[0]} + \frac{h_{j,l}}{2} \left[f(t_{j,l-1}, x_{j,l-1}^{[0]}) + f(t_{j,l}, X_1^{[0]}) \right]. \quad (1.78b)$$

Mit dem Defekt (1.16) wird das Nachbarproblem wie folgt definiert:

$$Z_1^{[\nu]} = z_{j,l-1}^{[\nu]} + h_{j,l} \left[f(t_{j,l-1}, z_{j,l-1}^{[\nu]}) + d_{j,l-1}^{[\nu]} \right], \quad (1.79a)$$

$$z_{j,l}^{[\nu]} = z_{j,l-1}^{[\nu]} + \frac{h_{j,l}}{2} \left[f(t_{j,l-1}, z_{j,l-1}^{[\nu]}) + f(t_{j,l}, Z_1^{[\nu]}) + d_{j,l-1}^{[\nu]} + d_{j,l}^{[\nu]} \right]. \quad (1.79b)$$

Beispiel 1.17. Wie Beispiel 1.3, jedoch mit dem Verfahren von Heun als Basisverfahren. Tabelle 1.7 zeigt die Resultate.

Beispiel 1.18. Wie Beispiel 1.5, jedoch mit dem Verfahren von Heun als Basisverfahren. Tabelle 1.8 zeigt die Resultate.

Bemerkung 1.19. Man sieht, dass man bei ungeradem Polynomgrad sowohl bei IDeC-Heun, als auch beim zugehörigen Fixpunktverfahren Heun-COLL (siehe Abschnitt 5.1) eine Konvergenzordnung „geschenkt“ bekommt (Symmetrie). Vergleicht man die Resultate der Beispiele 1.17 und 1.18, so erscheint es vom Rechenaufwand effizienter, mit einer ungeraden Anzahl von Kollokationsknoten zu rechnen. Wie weitere Beispiele zeigen, gilt dies mit Ausnahme des Taylor-Verfahrens 2. Ordnung (siehe Abschnitt 1.2.12) für alle Basisverfahren der Ordnung 2 (vergleiche z.B. die Beispiele 1.27 und 1.28).

¹³Bemerkung: Es werden nur die $x_{j,l}^{[\nu]}$ interpoliert, nicht aber die Stufen $X_1^{[\nu]}$.

h	Heun	IDeC1	IDeC2	Heun-COLL
0,1	2,41E-04	4,67E-08	3,88E-08	3,88E-08
0,05	5,97E-05	2,93E-09	2,42E-09	2,42E-09
0,025	1,49E-05	1,83E-10	1,52E-10	1,52E-10
0,0125	3,71E-06	1,15E-11	9,47E-12	9,47E-12
0,1	2,01	4,00	4,00	4,00
0,05	2,01	4,00	4,00	4,00
0,025	2,00	4,00	4,00	4,00
0,0125	2,00	4,00	4,00	4,00

Tabelle 1.8: Resultate zu Beispiel 1.18.

1.2.6 IDeC-Heun – Fixpunktkonvergenz

Satz 1.20. (*Fixpunktkonvergenz von IDeC-Heun*) Löst man das Anfangswertproblem vom Typ (1.1) auf einem äquidistanten Gitter (d.h. $h \equiv h_{j,l}$) gemäß Definition 1.1 mittels iterierter Defektkorrektur mit klassischem Defekt und dem Verfahren von Heun als Basisverfahren, so erfüllt der Fehler $e_{j,l}^{[\nu]}$ der Defektkorrektur-Lösungen $x_{j,l}^{[\nu]}$ der einzelnen Iterationen bezüglich des Fixpunktes $p_{j,l}^{[C]}$ (=Heun-Kollokationslösung gemäß Abschnitt 5.1.5 auf demselben Gitter) folgende Bedingung

$$\|e_{j,l}^{[\nu]}\| = \|x_{j,l}^{[\nu]} - p_{j,l}^{[C]}\| = \mathcal{O}(h^{2\nu+2}), \quad \nu = 0, 1, \dots \quad (1.80)$$

Beweis von Satz 1.20 – Herleitung der Fehlergleichungen (lineare Analysis):

Für die folgenden Ausführungen gilt $h \equiv h_{j,l}$, da das Verfahren mit klassischem Defekt ohnehin nur am äquidistanten Gitter konvergiert.

- $\nu = 0$: Subtrahiert man vom umgeformten Basisverfahren (1.78) (passend zu (1.44))

$$\frac{x_{j,l}^{[0]} - x_{j,l-1}^{[0]}}{h} = \frac{1}{2} \left[C_{j,l-1} x_{j,l-1}^{[0]} + g_{j,l-1} + C_{j,l} + h X_1^{[0]} + g_{j,l} \right], \text{ mit} \quad (1.81a)$$

$$X_1^{[0]} = x_{j,l-1}^{[0]} + h \left(C_{j,l-1} x_{j,l-1}^{[0]} + g_{j,l-1} \right) \quad (1.81b)$$

links und rechts den Term $\frac{p_{j,l}^{[C]} - p_{j,l-1}^{[C]}}{h}$, wobei $p_{j,l}^{[C]}$ die zugehörige Fixpunkt-Kollokationslösung (siehe Abschnitt 5.1.5) bezeichnet, erhält man mit $p^{[0]}(t)$ als Interpolante der $x_{j,l}^{[0]}$ unter Ausnützung von $e_{j,l}^{[0]} = p_{j,l}^{[0]} - p_{j,l}^{[C]}$:

$$\begin{aligned} \frac{e_{j,l}^{[0]} - e_{j,l-1}^{[0]}}{h} &= \frac{1}{2} \left[C_{j,l-1} x_{j,l-1}^{[0]} + g_{j,l-1} + C_{j,l} \left(x_{j,l-1}^{[0]} + h \left(C_{j,l-1} x_{j,l-1}^{[0]} + g_{j,l-1} \right) \right) + g_{j,l} \right] \\ &\quad - \frac{p_{j,l}^{[C]} - p_{j,l-1}^{[C]}}{h}. \end{aligned} \quad (1.82)$$

Durch Addition der linearen Variante der Heun-Fixpunkt-Kollokationsbedingung (5.11),

$$\begin{aligned} 0 &= \frac{1}{2} \left[p_{j,l-1}^{[C]'} - C_{j,l-1} p_{j,l-1}^{[C]} - g_{j,l-1} + p_{j,l}^{[C]'} - C_{j,l} p_{j,l}^{[C]} - g_{j,l} \right] \\ &\quad + \frac{1}{2} C_{j,l} h \left(p_{j,l-1}^{[C]'} - C_{j,l-1} p_{j,l-1}^{[C]} - g_{j,l-1} \right), \end{aligned} \quad (1.83)$$

und $C_{j,l} p_{j,l-1}^{[C]} - C_{j,l} p_{j,l-1}^{[C]} = 0$ gelangt man zur Heun-Fehlergleichung für $\nu = 0$:

$$\begin{aligned} \frac{e_{j,l}^{[0]} - e_{j,l-1}^{[0]}}{h} &= \frac{1}{2} \left[C_{j,l-1} e_{j,l-1}^{[0]} + C_{j,l} \left(e_{j,l-1}^{[0]} + h \left(C_{j,l-1} e_{j,l-1}^{[0]} \right) \right) \right] \\ &+ \frac{1}{2} \underbrace{\left[p_{j,l-1}^{[C]'} + p_{j,l}^{[C]'} \right]}_{=\bar{r}_{j,l}^{1,[0]}} - \frac{p_{j,l}^{[C]} - p_{j,l-1}^{[C]}}{h} \\ &- \frac{1}{2} C_{j,l} \underbrace{\left[p_{j,l}^{[C]} - p_{j,l-1}^{[C]} - h p_{j,l-1}^{[C]'} \right]}_{=\bar{r}_{j,l}^{2,[0]}}. \end{aligned} \quad (1.84)$$

Zu zeigen bleiben die Ordnungen der beiden Restglieder $\bar{r}_{j,l}^{1,[0]}$ und $\bar{r}_{j,l}^{2,[0]}$. Mittels Taylor-Entwicklung folgt bei gleichmäßiger Beschränktheit der Ableitungen (für $h \rightarrow 0$) von $p^{[C]}(t)$ (siehe Abschnitt 5.5.5)

$$\bar{r}_{j,l}^{1,[0]} = \frac{h^2}{12} p_{j,l-1}^{[C]'''} + \mathcal{O}(h^3) = \mathcal{O}(h^2) \quad (1.85)$$

bzw.

$$\bar{r}_{j,l}^{2,[0]} = -\frac{1}{2} C_{j,l} \left[\frac{h^2}{2} p_{j,l-1}^{[C]''} + \mathcal{O}(h^3) \right] = \mathcal{O}(h^2). \quad (1.86)$$

Setzt man die Ordnungen der Restglieder (1.85) und (1.86) in (1.84) ein, so erhält man zusammen mit einem Stabilitätsargument und dem Lemma von Gronwall (siehe z.B. [4]) die Ordnung $\mathcal{O}(h^2)$ auch global.

- $\nu - 1 \rightarrow \nu$: Ausgehend vom allgemeinen Iterationsschritt (1.17) gelangt man durch Subtraktion der Heun-Kollokationslösung $p_{j,l}^{[C]}$ auf beiden Seiten und Bildung des ersten Differenzenquotienten mit $h \equiv h_{j,l}$ zu

$$\frac{e_{j,l}^{[\nu]} - e_{j,l-1}^{[\nu]}}{h} = \frac{x_{j,l}^{[0]} - x_{j,l-1}^{[0]}}{h} - \frac{z_{j,l}^{[\nu-1]} - z_{j,l-1}^{[\nu-1]}}{h} + \frac{e_{j,l}^{[\nu-1]} - e_{j,l-1}^{[\nu-1]}}{h}. \quad (1.87)$$

Setzt man die linearen Beziehungen für das Basisverfahren (1.78) (gemäß (1.44))

$$\begin{aligned} \frac{x_{j,l}^{[0]} - x_{j,l-1}^{[0]}}{h} &= \frac{1}{2} \left[C_{j,l-1} x_{j,l-1}^{[0]} + g_{j,l-1} \right. \\ &\left. + C_{j,l} \left(x_{j,l-1}^{[0]} + h \left(C_{j,l-1} x_{j,l-1}^{[0]} + g_{j,l-1} \right) \right) + g_{j,l} \right] \end{aligned} \quad (1.88)$$

und das Nachbarproblem (1.79) für $\nu - 1$

$$\begin{aligned} \frac{z_{j,l}^{[\nu-1]} - z_{j,l-1}^{[\nu-1]}}{h} &= \frac{1}{2} \left[C_{j,l-1} z_{j,l-1}^{[\nu-1]} + g_{j,l-1} \right] \\ &+ \frac{1}{2} \left[C_{j,l} \left(z_{j,l-1}^{[\nu-1]} + h \left(C_{j,l-1} z_{j,l-1}^{[\nu-1]} + g_{j,l-1} + d_{j,l-1}^{[\nu-1]} \right) \right) + g_{j,l} \right] \\ &+ \frac{1}{2} \left[d_{j,l-1}^{[\nu-1]} + d_{j,l}^{[\nu-1]} \right] \end{aligned} \quad (1.89)$$

in (1.87) ein, erhält man

$$\begin{aligned}
& \frac{e_{j,l}^{[\nu]} - e_{j,l-1}^{[\nu]}}{h} \\
&= \frac{1}{2} \left[C_{j,l-1} x_{j,l-1}^{[0]} + g_{j,l-1} + C_{j,l} \left(x_{j,l-1}^{[0]} + h \left(C_{j,l-1} x_{j,l-1}^{[0]} + g_{j,l-1} \right) \right) + g_{j,l} \right] \\
&- \frac{1}{2} \left[C_{j,l-1} z_{j,l-1}^{[\nu-1]} + g_{j,l-1} + C_{j,l} \left(z_{j,l-1}^{[\nu-1]} + h \left(C_{j,l-1} z_{j,l-1}^{[\nu-1]} + g_{j,l-1} \right) \right) + g_{j,l} \right] \\
&+ \frac{1}{2} C_{j,l} h d_{j,l-1}^{[\nu-1]} - \frac{1}{2} \left[d_{j,l-1}^{[\nu-1]} + d_{j,l}^{[\nu-1]} \right] \\
&+ \frac{e_{j,l}^{[\nu-1]} - e_{j,l-1}^{[\nu-1]}}{h}. \tag{1.90}
\end{aligned}$$

Setzt man weiters für die Defektterme (1.16) ein und subtrahiert die linearisierte Heun-Fixpunkt-Kollokationsbedingung (1.83), gelangt man durch Umstellen und Zusammenfassen der Terme zur IDeC-Heun-Fehlergleichung

$$\begin{aligned}
\frac{e_{j,l}^{[\nu]} - e_{j,l-1}^{[\nu]}}{h} &= \frac{1}{2} \left[C_{j,l-1} e_{j,l-1}^{[\nu]} + C_{j,l} \left(e_{j,l-1}^{[\nu]} + h \left(C_{j,l-1} e_{j,l-1}^{[\nu]} \right) \right) \right] \\
&- \underbrace{\left(\frac{1}{2} \left[e_{j,l-1}^{[\nu-1]'} + e_{j,l}^{[\nu-1]'} \right] - \frac{e_{j,l}^{[\nu-1]} - e_{j,l-1}^{[\nu-1]}}{h} \right)}_{=\bar{r}_{j,l}^{1,[\nu]}} \\
&+ \underbrace{\frac{1}{2} C_{j,l} \left[e_{j,l}^{[\nu-1]} - e_{j,l-1}^{[\nu-1]} - h e_{j,l-1}^{[\nu-1]'} \right]}_{=\bar{r}_{j,l}^{2,[\nu]}}. \tag{1.91}
\end{aligned}$$

Für die Restglieder gilt induktiv bei analoger Vorgangsweise zu jener aus Abschnitt 1.2.3

$$\bar{r}_{j,l}^{1,[\nu]} = \frac{e_{j,l}^{[\nu-1]} - e_{j,l-1}^{[\nu-1]}}{h} - \frac{1}{2} \left[e_{j,l-1}^{[\nu-1]'} + e_{j,l}^{[\nu-1]'} \right] = \mathcal{O}(h^{2\nu+2}), \quad \nu = 1, 2, \dots \tag{1.92}$$

bzw.

$$\bar{r}_{j,l}^{2,[\nu]} = \frac{1}{2} C_{j,l} h \left[\frac{e_{j,l}^{[\nu-1]} - e_{j,l-1}^{[\nu-1]}}{h} - e_{j,l-1}^{[\nu-1]'} \right] = \mathcal{O}(h^{2\nu+2}), \quad \nu = 1, 2, \dots \tag{1.93}$$

Zusammen mit einem Stabilitätsargument und dem Lemma von Gronwall (siehe z.B. [4]) ist Satz 1.20 bewiesen.

1.2.7 IDeC-EMR

Analog zum Verfahren von Heun kann auch die explizite Mittelpunktsregel als Basisverfahren verwendet werden. Es gilt

$$X_{\frac{1}{2}}^{[0]} = x_{j,l-1}^{[0]} + \frac{h_{j,l}}{2} f \left(t_{j,l-1}, x_{j,l-1}^{[0]} \right), \tag{1.94a}$$

$$x_{j,l}^{[0]} = x_{j,l-1}^{[0]} + h_{j,l} f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, X_{\frac{1}{2}}^{[0]} \right) \tag{1.94b}$$

h	EMR	IDeC1	IDeC2	EMR-COLL
0,1	1,34E-04	1,38E-09	5,05E-08	5,05E-08
0,05	3,39E-04	6,58E-11	3,09E-09	3,09E-09
0,025	8,54E-05	8,91E-12	1,01E-10	1,91E-10
0,0125	2,14E-06	7,08E-13	1,19E-11	1,19E-11
0,1	1,98	4,39	4,03	4,03
0,05	1,99	2,88	4,02	4,02
0,025	1,99	3,65	4,01	4,01
0,0125				

Tabelle 1.9: Resultate zu Beispiel 1.21.

h	EMR	IDeC1	IDeC2	EMR-COLL
0,1	7,58E-05	8,89E-09	2,58E-08	2,58E-08
0,05	1,92E-05	5,48E-10	1,63E-09	1,63E-09
0,025	4,81E-06	3,40E-11	1,02E-10	1,02E-10
0,0125	1,21E-06	2,12E-12	6,40E-12	6,40E-12
0,1	1,98	4,02	3,99	3,99
0,05	1,99	4,01	3,99	3,99
0,025	2,00	4,00	4,00	4,00
0,0125				

Tabelle 1.10: Resultate zu Beispiel 1.22.

für das Basisverfahren. Mit (1.16) und

$$d_{\frac{1}{2}}^{[\nu]} := d^{[\nu]} \left(t_{j,l-1} + \frac{h_{j,l}}{2} \right) = p^{[\nu]'} \left(t_{j,l-1} + \frac{h_{j,l}}{2} \right) - f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, p^{[\nu]} \left(t_{j,l-1} + \frac{h_{j,l}}{2} \right) \right) \quad (1.95)$$

für die Defekte folgt für das Nachbarproblem

$$Z_{\frac{1}{2}}^{[\nu]} = z_{j,l-1}^{[0]} + \frac{h_{j,l}}{2} \left[f \left(t_{j,l-1}, z_{j,l-1}^{[0]} \right) + d_{j,l-1}^{[\nu]} \right], \quad (1.96a)$$

$$z_{j,l}^{[0]} = z_{j,l-1}^{[0]} + h_{j,l} \left[f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, Z_{\frac{1}{2}}^{[\nu]} \right) + d_{\frac{1}{2}}^{[\nu]} \right]. \quad (1.96b)$$

Beispiel 1.21. Wie Beispiel 1.3, jedoch mit der expliziten Mittelpunktsregel als Basisverfahren. Tabelle 1.9 zeigt die Resultate.

Beispiel 1.22. Wie Beispiel 1.5, jedoch mit der expliziten Mittelpunktsregel als Basisverfahren. Tabelle 1.10 zeigt die Resultate.

1.2.8 IDeC-RK4

Mit dem gleichen System wie beim Verfahren von Heun bzw. bei der expliziten Mittelpunktsregel, nur aufwändiger in der Implementierung, kann auch das klassische (4-stufige)

Runge-Kutta-Verfahren als Basisverfahren verwendet werden. Es gilt

$$X_1^{[0]} = x_{j,l-1}^{[0]}, \quad (1.97a)$$

$$X_2^{[0]} = x_{j,l-1}^{[0]} + \frac{h_{j,l}}{2} f \left(t_{j,l-1}, X_1^{[0]} \right), \quad (1.97b)$$

$$X_3^{[0]} = x_{j,l-1}^{[0]} + \frac{h_{j,l}}{2} f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, X_2^{[0]} \right), \quad (1.97c)$$

$$X_4^{[0]} = x_{j,l-1}^{[0]} + h_{j,l} f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, X_3^{[0]} \right), \quad (1.97d)$$

und

$$\begin{aligned} x_{j,l}^{[0]} = x_{j,l-1}^{[0]} &+ h_{j,l} \left[\frac{1}{6} f \left(t_{j,l-1}, X_1^{[0]} \right) + \frac{1}{3} f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, X_2^{[0]} \right) \right. \\ &\left. + \frac{1}{3} f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, X_3^{[0]} \right) + \frac{1}{6} f \left(t_{j,l}, X_4^{[0]} \right) \right]. \end{aligned} \quad (1.98)$$

Mit den Defekten (1.16) und (1.95) gelangt man zum RK4-Nachbarproblem

$$Z_1^{[0]} = z_{j,l-1}^{[0]}, \quad (1.99a)$$

$$Z_2^{[0]} = z_{j,l-1}^{[0]} + \frac{h_{j,l}}{2} \left[f \left(t_{j,l-1}, Z_1^{[0]} \right) + d_{j,l-1}^{[\nu]} \right], \quad (1.99b)$$

$$Z_3^{[0]} = z_{j,l-1}^{[0]} + \frac{h_{j,l}}{2} \left[f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, Z_2^{[0]} \right) + d_{\frac{1}{2}}^{[\nu]} \right], \quad (1.99c)$$

$$Z_4^{[0]} = z_{j,l-1}^{[0]} + h_{j,l} \left[f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, Z_3^{[0]} \right) + d_{\frac{1}{2}}^{[\nu]} \right], \quad (1.99d)$$

und

$$\begin{aligned} z_{j,l}^{[0]} = z_{j,l-1}^{[0]} &+ h_{j,l} \left[\frac{1}{6} \left[f \left(t_{j,l-1}, Z_1^{[0]} \right) + d_{j,l-1}^{[\nu]} \right] + \frac{1}{3} \left[f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, Z_2^{[0]} \right) + d_{\frac{1}{2}}^{[\nu]} \right] \right. \\ &\left. + \frac{1}{3} \left[f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, Z_3^{[0]} \right) + d_{\frac{1}{2}}^{[\nu]} \right] + \frac{1}{6} \left[f \left(t_{j,l}, Z_4^{[0]} \right) + d_{j,l}^{[\nu]} \right] \right]. \end{aligned} \quad (1.100)$$

Bemerkung 1.23. Wirklich sinnvoll ist das klassische Runge-Kutta-Verfahren nur in Kombination mit einem Polynom 8. Grades. Damit wäre nach einer Iteration Konvergenzordnung 8 erreicht.

Beispiel 1.24. Die experimentelle Überprüfung von IDeC-RK4 wird für das lineare Anfangswertproblem

$$u'(t) = \cos(t) u(t) - \sin(t) - \cos(t)^2, \quad (1.101a)$$

$$u(t_0 = 0) = 1, \quad (1.101b)$$

bei Polynomgrad 8 auf einem äquidistanten Gitter gezeigt. Die exakte Lösung lautet $u^*(t) = \cos(t)$. Tabelle 1.11 zeigt die Resultate für $t_{end} = 3$.

h	RK4	IDeC1	IDeC2	RK4-COLL
0,1	3,70E-10	1,66E-19	2,32E-19	2,32E-19
0,05	2,32E-11	6,48E-22	9,04E-22	9,04E-22
0,025	1,45E-12	2,53E-24	3,53E-24	3,53E-24
0,0125	9,05E-14	9,89E-27	1,38E-26	1,38E-26
0,1	4,00	8,00	8,00	8,00
0,05	4,00	8,00	8,00	8,00
0,025	4,00	8,00	8,00	8,00
0,0125	4,00	8,00	8,00	8,00

Tabelle 1.11: Resultate zu Beispiel 1.24.

h	IEUL	IDeC1	IDeC2	IDeC3	COLL
0,1	3,23E-02	1,77E-03	2,49E-05	1,13E-05	9,09E-06
0,05	1,63E-02	4,47E-04	2,75E-06	1,28E-06	1,14E-06
0,025	8,15E-03	1,12E-04	3,19E-07	1,51E-07	1,42E-07
0,0125	4,08E-03	2,81E-05	3,82E-08	1,83E-08	1,78E-08
0,1	0,99	1,98	3,18	3,14	3,00
0,05	1,00	1,99	3,11	3,08	3,00
0,025	1,00	2,00	3,06	3,04	3,00
0,0125	1,00	2,00	3,06	3,04	3,00

Tabelle 1.12: Resultate zu Beispiel 1.25.

1.2.9 IDeC-IEUL

Ersetzt man unter Beibehaltung sämtlicher anderer algorithmischer Details das Basisverfahren (1.4a) durch

$$\frac{x_{j,l}^{[0]} - x_{j,l-1}^{[0]}}{h_{j,l}} = f\left(t_{j,l}, x_{j,l}^{[0]}\right), \quad (1.102)$$

den Defekt (1.16) durch

$$d_{j,l}^{[\nu]} = p_{j,l}^{[\nu]'} - f\left(t_{j,l}, p_{j,l}^{[\nu]}\right), \quad (1.103)$$

und das Nachbarproblem (1.15a) durch

$$\frac{z_{j,l}^{[\nu]} - z_{j,l-1}^{[\nu]}}{h_{j,l}} = f\left(t_{j,l}, z_{j,l}^{[\nu]}\right) + d_{j,l}^{[\nu]}, \quad (1.104)$$

so erhält man ein implizites Defektkorrekturverfahren mit der Ordnungsfolge 1,2,...

Beispiel 1.25. Wie Beispiel 1.3, jedoch mit dem impliziten Euler-Verfahren als Basisverfahren. Tabelle 1.12 zeigt die Resultate.

Beispiel 1.26. Wie Beispiel 1.5, jedoch mit dem impliziten Euler-Verfahren als Basisverfahren. Tabelle 1.13 zeigt die Resultate.

h	IEUL	IDeC1	IDeD2	IDeC3	IDeC4	COLL
0,1	2,42E-02	1,00E-03	9,87E-06	7,74E-07	1,28E-07	1,18E-07
0,05	1,22E-02	2,52E-04	1,11E-06	4,99E-08	7,52E-09	7,25E-09
0,025	6,12E-03	6,33E-05	1,32E-07	3,16E-09	4,58E-10	4,50E-10
0,0125	3,06E-03	1,58E-05	1,60E-08	1,99E-10	2,82E-11	2,80E-11
0,1	0,99	1,99	3,15	3,96	4,09	4,02
0,05	1,00	2,00	3,08	3,98	4,04	4,01
0,025	1,00	2,00	3,04	3,99	4,02	4,01
0,0125						

Tabelle 1.13: Resultate zu Beispiel 1.26.

h	ITR	IDeC1	IDeC2	ITR-COLL
0,1	3,48E-04	1,96E-07	7,83E-08	7,84E-08
0,05	8,71E-05	1,22E-08	4,90E-09	4,90E-09
0,025	2,18E-05	7,65E-10	3,06E-10	3,06E-10
0,0125	5,44E-06	4,78E-11	1,91E-11	1,91E-11
0,1	2,00	4,00	4,00	4,00
0,05	2,00	4,00	4,00	4,00
0,025	2,00	4,00	4,00	4,00
0,0125	2,00	4,00	4,00	4,00

Tabelle 1.14: Resultate zu Beispiel 1.27.

1.2.10 IDeC-ITR

Mit der impliziten Trapezregel als Basisverfahren erhält man ein Defektkorrekturverfahren mit der Ordnungsfolge 2,4,... Für das Basisverfahren und das Nachbarproblem mit dem Defekt (1.16) gelten

$$\frac{x_{j,l}^{[0]} - x_{j,l-1}^{[0]}}{h_{j,l}} = \frac{1}{2} \left[f \left(t_{j,l-1}, x_{j,l-1}^{[0]} \right) + f \left(t_{j,l}, x_{j,l}^{[0]} \right) \right], \quad (1.105)$$

und

$$\frac{z_{j,l}^{[\nu]} - z_{j,l-1}^{[\nu]}}{h_{j,l}} = \frac{1}{2} \left[f \left(t_{j,l-1}, z_{j,l-1}^{[\nu]} \right) + f \left(t_{j,l}, z_{j,l}^{[\nu]} \right) + d_{j,l-1}^{[\nu]} + d_{j,l}^{[\nu]} \right]. \quad (1.106)$$

Beispiel 1.27. Wie Beispiel 1.3, jedoch mit der impliziten Trapezregel als Basisverfahren. Tabelle 1.14 zeigt die Resultate.

Beispiel 1.28. Wie Beispiel 1.5, jedoch mit der impliziten Trapezregel als Basisverfahren. Tabelle 1.15 zeigt die Resultate.

h	ITR	IDeC1	IDeC2	ITR-COLL
0,1	1,96E-04	7,67E-08	3,88E-08	3,88E-08
0,05	4,90E-04	4,79E-09	2,42E-09	2,42E-09
0,025	1,22E-05	2,99E-10	1,52E-10	1,52E-10
0,0125	3,06E-06	1,87E-11	9,47E-12	9,47E-12
0,1	2,00	4,00	4,00	4,00
0,05	2,00	4,00	4,00	4,00
0,025	2,00	4,00	4,00	4,00
0,0125	2,00	4,00	4,00	4,00

Tabelle 1.15: Resultate zu Beispiel 1.28.

h	IMR	IDeC1	IDeC2	IMR-COLL
0,1	2,11E-04	5,93E-08	6,86E-08	6,86E-08
0,05	5,27E-05	3,71E-09	4,29E-09	4,29E-09
0,025	1,32E-05	2,32E-10	2,68E-10	2,68E-10
0,0125	3,29E-06	1,45E-11	1,67E-11	1,67E-11
0,1	2,00	4,00	4,00	4,00
0,05	2,00	4,00	4,00	4,00
0,025	2,00	4,00	4,00	4,00
0,0125	2,00	4,00	4,00	4,00

Tabelle 1.16: Resultate zu Beispiel 1.29.

1.2.11 IDeC-IMR

Analog zu IDeC-ITR kann man auch die implizite Mittelpunktsregel als Basisverfahren verwenden

$$\frac{x_{j,l}^{[0]} - x_{j,l-1}^{[0]}}{h_{j,l}} = f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, \frac{x_{j,l-1}^{[0]} + x_{j,l}^{[0]}}{2} \right). \quad (1.107)$$

Setzt man den Defekt (1.95) ins Nachbarproblem

$$\frac{z_{j,l}^{[\nu]} - z_{j,l-1}^{[\nu]}}{h_{j,l}} = f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, \frac{z_{j,l-1}^{[\nu]} + z_{j,l}^{[\nu]}}{2} \right) + d_{\frac{1}{2}}^{[\nu]} \quad (1.108)$$

ein, wird das Verfahren komplett.

Beispiel 1.29. Wie Beispiel 1.3, jedoch mit der impliziten Mittelpunktsregel als Basisverfahren. Tabelle 1.16 zeigt die Resultate.

Beispiel 1.30. Wie Beispiel 1.5, jedoch mit der impliziten Mittelpunktsregel als Basisverfahren. Tabelle 1.17 zeigt die Resultate.

h	IMR	IDeC1	IDeC2	IMR-COLL
0,1	1,19E-04	2,42E-08	2,63E-08	2,63E-08
0,05	2,96E-05	1,52E-09	1,64E-09	1,64E-09
0,025	7,41E-06	9,47E-11	1,03E-10	1,03E-10
0,0125	1,85E-06	5,92E-12	6,42E-12	6,42E-12
0,1	2,00	4,00	4,00	4,00
0,05	2,00	4,00	4,00	4,00
0,025	2,00	4,00	4,00	4,00
0,0125	2,00	4,00	4,00	4,00

Tabelle 1.17: Resultate zu Beispiel 1.30.

1.2.12 IDeC-Taylor2

Es sei festgehalten, dass das Taylorverfahren zum Lösen von Differentialgleichungssystemen aufgrund des vergleichsweise hohen Aufwands zur Bestimmung der Ableitungen nur bedingt geeignet ist. Die Vorgangsweise ist jedoch relevant im Hinblick auf Anfangswertprobleme 2. (oder höherer) Ordnung, da man für den Startschritt einen Taylorschritt passender Ordnung benötigt (siehe Kapitel 3 und 4).

Für ein Anfangswertproblem erster Ordnung der Bauart (1.1) gilt mit $u \equiv u(t)$ und $f \equiv f(t, u)$

$$u' = f, \quad (1.109a)$$

$$u'' = f_t + f_u u' = f_t + f_u f, \quad (1.109b)$$

$$u''' = f_{tt} + 2f_{tu}f + f_u f_t + f_{uu}f^2 + f_u^2 f. \quad (1.109c)$$

Somit lautet das Taylor-Basisverfahren der Ordnung 2

$$\begin{aligned} x_{j,l}^{[0]} &= x_{j,l-1}^{[0]} + h_{j,l} f(t_{j,l-1}, x_{j,l-1}^{[0]}) \\ &+ \frac{h_{j,l}^2}{2} \left[f_t(t_{j,l-1}, x_{j,l-1}^{[0]}) + f_u(t_{j,l-1}, x_{j,l-1}^{[0]}) f(t_{j,l-1}, x_{j,l-1}^{[0]}) \right]. \end{aligned} \quad (1.110)$$

Der „Gesamtdefekt“ teilt sich dabei in 2 Anteile A und B

$$d_{j,l-1}^{A,[\nu]} = p_{j,l-1}^{[\nu]'} - f(t_{j,l-1}, p_{j,l-1}^{[\nu]}), \quad (1.111a)$$

$$d_{j,l-1}^{B,[\nu]} = p_{j,l-1}^{[\nu]''} - \left[f_t(t_{j,l-1}, p_{j,l-1}^{[\nu]}) + f_u(t_{j,l-1}, p_{j,l-1}^{[\nu]}) f(t_{j,l-1}, p_{j,l-1}^{[\nu]}) \right]. \quad (1.111b)$$

Setzt man diese in das zugehörige Nachbarproblem ein, erhält man

$$\begin{aligned} z_{j,l}^{[\nu]} &= z_{j,l-1}^{[\nu]} + h_{j,l} f(t_{j,l-1}, z_{j,l-1}^{[\nu]}) \\ &+ \frac{h_{j,l}^2}{2} \left[f_t(t_{j,l-1}, z_{j,l-1}^{[\nu]}) + f_u(t_{j,l-1}, z_{j,l-1}^{[\nu]}) f(t_{j,l-1}, z_{j,l-1}^{[\nu]}) \right] \\ &+ h_{j,l} d_{j,l-1}^{A,[\nu]} + \frac{h_{j,l}^2}{2} d_{j,l-1}^{B,[\nu]}. \end{aligned} \quad (1.112)$$

Das Verfahren liefert die Ordnungsfolge 2,4,...

h	Taylor2	IDeC1	IDeC2	COLL
0,1	6,88E-04	2,92E-06	3,38E-06	9,08E-06
0,05	1,73E-04	3,71E-07	4,01E-07	1,14E-06
0,025	4,34E-04	4,69E-08	4,87E-08	1,42E-07
0,0125	1,09E-05	5,89E-09	6,00E-09	1,77E-08
0,1	1,99	2,97	3,08	3,00
0,05	2,00	2,99	3,04	3,00
0,025	2,00	2,99	3,02	3,00
0,0125				

Tabelle 1.18: Resultate zu Beispiel 1.31.

h	Taylor2	IDeC1	IDeC2	COLL
0,1	3,88E-04	3,88E-09	1,44E-07	1,18E-07
0,05	9,75E-05	3,27E-10	9,02E-09	7,25E-09
0,025	2,44E-05	2,30E-11	5,65E-10	4,50E-10
0,0125	6,12E-06	1,52E-12	3,54E-11	2,80E-11
0,1	1,99	3,57	3,99	4,02
0,05	2,00	3,83	4,00	4,01
0,025	2,00	3,92	4,00	4,01
0,0125				

Tabelle 1.19: Resultate zu Beispiel 1.32.

Beispiel 1.31. Wie Beispiel 1.3, jedoch mit dem Taylor-Verfahren 2. Ordnung als Basisverfahren. Tabelle 1.18 zeigt die Resultate, wobei die letzte Spalte nicht die Resultate der Taylor2-Fixpunktkollokation, sondern der IEUL-Kollokation (=Standard-Kollokation) zum Vergleich (siehe Abschnitt 5.1) zeigt. Man sieht, dass man anders als bei allen anderen Basisverfahren der Ordnung 2 bei ungeradem Polynomgrad keine Ordnung zusätzlich erhält.

Beispiel 1.32. Wie Beispiel 1.5, jedoch mit dem Taylor-Verfahren 2. Ordnung als Basisverfahren. Tabelle 1.19 zeigt die Resultate, wobei die letzte Spalte nicht die Resultate der Taylor2-Fixpunktkollokation, sondern der IEUL-Kollokation (=Standard-Kollokation) zum Vergleich (siehe Abschnitt 5.1) zeigt.

1.2.13 IDeC-Taylor3

Führt man die Taylorreihenentwicklung gemäß (1.109c) um einen Term weiter aus, gelangt man zum Taylorverfahren der Ordnung 3

$$\begin{aligned}
x_{j,l}^{[0]} &= x_{j,l-1}^{[0]} + h_{j,l} f \left(t_{j,l-1} x_{j,l-1}^{[0]} \right) \\
&+ \frac{h_{j,l}^2}{2} \left[f_t \left(t_{j,l-1} x_{j,l-1}^{[0]} \right) + f_u \left(t_{j,l-1} x_{j,l-1}^{[0]} \right) f \left(t_{j,l-1} x_{j,l-1}^{[0]} \right) \right] \\
&+ \frac{h_{j,l}^3}{6} \left[f_{tt} \left(t_{j,l-1} x_{j,l-1}^{[0]} \right) \right. \\
&\quad + 2 f_{tu} \left(t_{j,l-1} x_{j,l-1}^{[0]} \right) f \left(t_{j,l-1} x_{j,l-1}^{[0]} \right) \\
&\quad + f_u \left(t_{j,l-1} x_{j,l-1}^{[0]} \right) f_t \left(t_{j,l-1} x_{j,l-1}^{[0]} \right) \\
&\quad + f_{uu} \left(t_{j,l-1} x_{j,l-1}^{[0]} \right) f^2 \left(t_{j,l-1} x_{j,l-1}^{[0]} \right) \\
&\quad \left. + f_u^2 \left(t_{j,l-1} x_{j,l-1}^{[0]} \right) f \left(t_{j,l-1} x_{j,l-1}^{[0]} \right) \right]. \tag{1.113}
\end{aligned}$$

Für das zugehörige Nachbarproblem benötigt man zusätzlich zu den Defekten (1.111) einen C -Defektanteil

$$\begin{aligned}
d_{j,l-1}^{C,[\nu]} &= p_{j,l-1}^{[\nu]'''} - \left[f_{tt} \left(t_{j,l-1}, p_{j,l-1}^{[\nu]} \right) \right. \\
&\quad + 2 f_{tu} \left(t_{j,l-1}, p_{j,l-1}^{[\nu]} \right) f \left(t_{j,l-1}, p_{j,l-1}^{[\nu]} \right) \\
&\quad + f_u \left(t_{j,l-1}, p_{j,l-1}^{[\nu]} \right) f_t \left(t_{j,l-1}, p_{j,l-1}^{[\nu]} \right) \\
&\quad + f_{uu} \left(t_{j,l-1}, p_{j,l-1}^{[\nu]} \right) f^2 \left(t_{j,l-1}, p_{j,l-1}^{[\nu]} \right) \\
&\quad \left. + f_u^2 \left(t_{j,l-1}, p_{j,l-1}^{[\nu]} \right) f \left(t_{j,l-1}, p_{j,l-1}^{[\nu]} \right) \right]. \tag{1.114}
\end{aligned}$$

h	Taylor3	IDeC1	IDeC2	COLL
0,1	3,82E-07	9,65E-13	1,05E-12	4,05E-12
0,05	4,75E-08	1,49E-14	1,62E-14	6,24E-14
0,025	5,93E-09	2,33E-16	2,51E-16	9,68E-16
0,0125	7,40E-10	3,63E-18	3,91E-18	1,51E-17
0,1	3,01	6,01	6,02	6,02
0,05	3,00	6,01	6,01	6,01
0,025	3,00	6,00	6,00	6,01
0,0125	3,00	6,00	6,00	6,01

Tabelle 1.20: Resultate zu Beispiel 1.34.

Für das Taylor-Nachbarproblem der Ordnung 3 erhält man

$$\begin{aligned}
z_{j,l}^{[\nu]} &= z_{j,l-1}^{[\nu]} + h_{j,l} f \left(t_{j,l-1} z_{j,l-1}^{[\nu]} \right) \\
&+ \frac{h_{j,l}^2}{2} \left[f_t \left(t_{j,l-1} z_{j,l-1}^{[\nu]} \right) + f_u \left(t_{j,l-1} z_{j,l-1}^{[\nu]} \right) f \left(t_{j,l-1} z_{j,l-1}^{[\nu]} \right) \right] \\
&+ \frac{h_{j,l}^3}{6} \left[f_{tt} \left(t_{j,l-1} z_{j,l-1}^{[\nu]} \right) \right. \\
&\quad + 2 f_{tu} \left(t_{j,l-1} z_{j,l-1}^{[\nu]} \right) f \left(t_{j,l-1} z_{j,l-1}^{[\nu]} \right) \\
&\quad + f_u \left(t_{j,l-1} z_{j,l-1}^{[\nu]} \right) f_t \left(t_{j,l-1} z_{j,l-1}^{[\nu]} \right) \\
&\quad + f_{uu} \left(t_{j,l-1} z_{j,l-1}^{[\nu]} \right) f^2 \left(t_{j,l-1} z_{j,l-1}^{[\nu]} \right) \\
&\quad \left. + f_u^2 \left(t_{j,l-1} z_{j,l-1}^{[\nu]} \right) f \left(t_{j,l-1} z_{j,l-1}^{[\nu]} \right) \right] \\
&+ h_{j,l} d_{j,l-1}^{A,[\nu]} + \frac{h_{j,l}^2}{2} d_{j,l-1}^{B,[\nu]} + \frac{h_{j,l}^3}{6} d_{j,l-1}^{C,[\nu]}. \tag{1.115}
\end{aligned}$$

Bemerkung 1.33. Das Verfahren ist nur dann sinnvoll anzuwenden, wenn der Grad der (stückweise definierten) Polynome $p^{[\nu]}(t)$ größer gleich 6 ist. Aufgrund der hohen Anzahl von $f_{(*)}$ -Auswertungen und dem damit verbundenen Rechenaufwand ist von der Anwendung des Verfahrens (für Systeme) jedoch grundsätzlich abzuraten.

Beispiel 1.34. Wie Beispiel 1.3, jedoch mit dem Taylor-Verfahren 3. Ordnung als Basisverfahren und Polynomgrad 6. Tabelle 1.20 zeigt die Resultate, wobei die letzte Spalte nicht die Resultate der Taylor3-Fixpunktkollokation, sondern der IEUL-Kollokation (=Standard-Kollokation) zum Vergleich (siehe Abschnitt 5.1) zeigt.

1.3 Integralmitteldefekt – IQDeC

Im Folgenden wird eine weitere Defektvariante, welche in [2] erstmals vorgestellt wurde, kurz zusammengefasst wiedergegeben: IQDeC, engl.: *Iterated Defect Correction with Defect*

Quadrature = Iterierte Defektkorrektur mit Integralmitteldefekt. Sie hat den Vorteil, dass sie auf nicht äquidistanten Gittern anwendbar ist und somit superkonvergente Verfahren konstruierbar sind (z.B. mittels Radau-Knoten).

1.3.1 IQDeC-EEUL

Die Konstruktion des Integralmitteldefekts beruht auf der Tatsache, dass für den Fixpunkt der Defektkorrektur, das Kollokationspolynom $p^{[C]}(t)$ vom Grad m , folgende Identitäten gelten:

$$\frac{1}{h_{j,l}} \int_{t_{j,l-1}}^{t_{j,l}=t_{j,l-1}+h_{j,l}} p^{[C]'}(s) ds = \frac{p_{j,l}^{[C]} - p_{j,l-1}^{[C]}}{h_{j,l}} = \sum_{\mu=1}^m \alpha_{l,\mu} f(t_{j,\mu}, p_{j,\mu}^{[C]}). \quad (1.116)$$

Die Darstellung der gegebenen Differentialgleichung als Volterra-Integralgleichung bezeichnet W. Auzinger als *exaktes Differenzenschema*. Darauf aufbauend lässt sich obige Identität begründen. Für die Quadraturkoeffizienten $\alpha_{l,i}$ gilt für beliebige Polynome $q(s)$ vom Maximalgrad m

$$\alpha_{l,i} = \frac{1}{c_l - c_{l-1}} \int_{c_{l-1}}^{c_l} \Lambda_i(s) ds, \quad (1.117)$$

wobei $\Lambda_i(s)$ die Lagrange-Polynome bezeichnet (siehe Abschnitt 5) und für $i, l = 1, \dots, m$ gilt. D.h., dass alle Polynome bis zum Grad m exakt integriert werden

$$\sum_{\mu=1}^m \alpha_{l,\mu} q(c_\mu) = \frac{1}{c_l - c_{l-1}} \int_{c_{l-1}}^{c_l} q(s) ds. \quad (1.118)$$

Darüber hinaus gilt für die Quadraturkoeffizienten

$$\sum_{\mu=1}^m \alpha_{l,\mu} = 1. \quad (1.119)$$

Basierend auf der rechten Gleichung von (1.116) kann man den so genannten *Integralmitteldefekt* definieren, der am Fixpunkt verschwindet¹⁴. Legt man durch die Lösungen $x_{j,i}^{[\nu]}$ der einzelnen Defektkorrekturschritte stückweise definierte Polynome $p^{[\nu]}(t)$, so ergibt sich mit $p_{j,l}^{[\nu]} \equiv p^{[\nu]}(t_{j,l})$ die Definition

$$\bar{d}_{j,l}^{[\nu]} := \frac{p_{j,l}^{[\nu]} - p_{j,l-1}^{[\nu]}}{h_{j,l}} - \sum_{\mu=1}^m \alpha_{l,\mu} f(t_{j,\mu}, p_{j,\mu}^{[\nu]}). \quad (1.120)$$

Für die Kollokationslösung gilt

$$\bar{d}_{j,l}^{[C]} = \frac{p_{j,l}^{[C]} - p_{j,l-1}^{[C]}}{h_{j,l}} - \sum_{\mu=1}^m \alpha_{l,\mu} f(t_{j,\mu}, p_{j,\mu}^{[C]}) = 0, \quad (1.121)$$

siehe (1.116). Der in (1.120) definierte Defekt ersetzt den klassischen Defekt (1.16), der Rest des Algorithmus ist ident mit jenem aus Abschnitt 1.2.

¹⁴In Abschnitt 3 von [2] wird ausführlich der Zusammenhang mit der zugehörigen Runge-Kutta- bzw. Kollokationslösung diskutiert.

h	EEUL	IQDeC1	IQDeC2	IQDeC3	IQDeC4	COLL
0,1	2,24E-03	2,19E-04	2,60E-06	2,07E-08	5,54E-09	5,55E-09
0,05	1,08E-03	5,36E-05	3,41E-07	1,44E-09	3,45E-10	3,47E-10
0,025	5,33E-04	1,32E-05	4,35E-08	9,47E-11	2,16E-11	2,17E-11
0,0125	2,64E-04	3,29E-06	5,49E-09	6,06E-12	1,35E-12	1,35E-12
0,1	1,05	2,03	2,93	3,84	4,00	4,00
0,05	1,02	2,01	2,97	3,92	4,00	4,00
0,025	1,01	2,00	2,99	3,97	4,00	4,00
0,0125						

Tabelle 1.21: Resultate zu Beispiel 1.36.

Bemerkung 1.35. Der Zusammenhang (1.120) kann als Quadratur für das Integralmittel des klassischen Defekts (1.16) aufgefasst werden. Es gilt

$$\begin{aligned} \bar{d}_{j,l}^{[\nu]} &= \sum_{\mu=1}^m \alpha_{l,\mu} d_{j,\mu}^{[\nu]} = \sum_{\mu=1}^m \alpha_{l,\mu} \left[p^{[\nu]'}(t_{j,\mu}) - f(t_{j,\mu}, p_{j,\mu}^{[\nu]}) \right] \\ &\approx \frac{1}{h_{j,l}} \int_{t_{j,l-1}}^{t_{j,l}=t_{j,l-1}+h_{j,l}} \left(p^{[\nu]'}(s) - f(s, p^{[\nu]'}(s)) \right) ds, \end{aligned} \quad (1.122)$$

was durch Auflösen der eckigen Klammer und Ausführen der numerischen Integration in (1.120) übergeführt werden kann. Dieses Detail ist wichtig im Hinblick auf die Erweiterung des Algorithmus auf implizite Differentialgleichungen (siehe Abschnitt 2.2).

Beispiel 1.36. Gegeben sei das skalare nichtlineare Anfangswertproblem

$$u'(t) = \frac{u^2(t)}{1+u^2(t)} - \sin(t) - \frac{\cos^2(t)}{1+\cos^2(t)}, \quad (1.123a)$$

$$u(t_0 = 0) = 1, \quad (1.123b)$$

mit der exakten Lösung $u^*(t) = \cos(t)$. Die IQDeC-EEUL-Lösung erfolgt auf einem *nicht* äquidistanten Gitter mit 4 Kollokationsknoten bzw. Polynomgrad 4 gemäß Definition 1.1 ($c_1 = 0,1234$, $c_2 = 0,5054$, $c_3 = 0,7134$, $c_4 = 1$) im Intervall $[t_0, t_{end}] = [\mathbf{t}_0, \mathbf{t}_N] = [0, 3]$. Die Resultate, also die globalen Fehler bezüglich der exakten Lösung $e^{*,[\#]}(t_{end}) = p^{[\#]}(t_{end}) - u^*(t_{end})$, $\# = \nu, C$ und die daraus berechneten Konvergenzordnungen an der Stelle $t_{end} = 3$ für das Basisverfahren, IQDeC und das zugehörige Fixpunkt-Kollokationsverfahren (=IEUL-Kollokation), sind in Tabelle 1.21 aufgelistet.

Beispiel 1.37. Wie Beispiel 1.36, jedoch auf einem Radau-Gitter mit 3 Kollokationsknoten ($c_1 = \frac{4-\sqrt{6}}{10}$, $c_2 = \frac{4+\sqrt{6}}{10}$, $c_3 = 1$). Ergebnisse siehe Tabelle 1.22.

1.3.2 IQDeC-EEUL – Fixpunktconvergenz

Satz 1.38. (*Fixpunktconvergenz von IQDeC-EEUL*) Löst man das Anfangswertproblem vom Typ (1.1) auf einem beliebigen Gitter gemäß Definition 1.1 mittels iterierter Defektkorrektur

h	EEUL	IQDeC1	IQDeC2	IQDeC3	IQDeC4	IQDeC5	COLL
0,1	3,06E-03	4,09E-04	6,27E-06	4,97E-08	2,96E-11	5,12E-11	3,87E-11
0,05	1,48E-03	9,95E-05	8,40E-07	3,82E-09	8,29E-12	1,53E-12	1,39E-12
0,025	7,26E-04	2,46E-05	1,08E-07	2,61E-10	3,75E-13	4,78E-14	4,61E-14
0,0125	3,60E-04	6,10E-06	1,37E-08	1,70E-11	1,35E-14	1,51E-15	1,48E-15
0,1	1,05	2,04	2,90	3,70	1,84	5,06	4,80
0,05	1,03	2,02	2,95	3,87	4,47	5,00	4,91
0,025	1,01	2,01	2,98	3,94	4,79	4,99	4,96
0,0125							

Tabelle 1.22: Resultate zu Beispiel 1.37.

mit Integralmittelfekt, so erfüllt der Fehler $e_{j,l}^{[\nu]}$ der Lösungen $x_{j,l}^{[\nu]}$ der einzelnen Iterationen bezüglich des Fixpunktes $p_{j,l}^{[C]}$ (=Kollokationslösung auf demselben Gitter) die Abschätzung

$$\left\| e_{j,l}^{[\nu]} \right\| = \left\| x_{j,l}^{[\nu]} - p_{j,l}^{[C]} \right\| = \mathcal{O} \left(h^{\nu+1} \right), \quad \nu = 0, 1, \dots \quad (1.124)$$

Für den Beweis dieses Satzes sei auf den Anhang von [2] verwiesen. Dort ist der erste Schritt ausführlich dargestellt. Der folgende Abschnitt beschränkt sich als Vorbereitung auf den Fall impliziter Differentialgleichungen erster Ordnung (siehe Abschnitt 2.2.2) auf die abstrakte Beweisorganisation für den ν -ten Iterationsschritt.

IQDeC-EEUL – Abstrakt formulierte Beweisorganisation für Satz 1.38

Die Notation der folgenden Ausführungen orientiert sich an jener von Abschnitt 1.2.3, auf Interpolationsoperatoren und dgl. wird zugunsten einfacherer Lesbarkeit verzichtet. Es sei

$$\mathcal{L}u = \mathcal{F}(u) \quad (1.125)$$

die Differentialgleichung (1.1) mit exakter Lösung u^* . Für diese gilt das exakte Differenzenschema, welches abstrakt durch

$$Lu^* = \mathcal{G}(u^*) \quad (1.126)$$

dargestellt werden kann, wobei L den zu \mathcal{L} gehörenden Differenzenoperator darstellt und \mathcal{G} den Integraloperator auf der linken Seite von (1.116) bezeichnet. Die Approximation des exakten Differenzschemas mittels einer geeigneten Quadraturformel wird mittels

$$Lp^{[C]} = G(p^{[C]}) \quad (1.127)$$

abstrahiert (G bezeichnet die Quadratur für \mathcal{G}). Für die Kollokationslösung $p^{[C]}$ von (1.125) (am entsprechenden Gitter) ist die Quadratur exakt. Mittels (1.127) kann man den Integralmittelfekt (1.120) für den ν -ten Defektkorrekturschritt allgemein anschreiben als

$$\bar{d}^{[\nu-1]} = Lx^{[\nu-1]} - G(x^{[\nu-1]}). \quad (1.128)$$

Am Fixpunkt gilt

$$\bar{d}^{[C]} = Lp^{[C]} - G(p^{[C]}) = 0. \quad (1.129)$$

Ausgehend von der Verbesserung der Näherungslösung gemäß (1.17)

$$x^{[\nu]} = x^{[0]} - (z^{[\nu-1]} - x^{[\nu-1]}), \quad (1.130)$$

und (1.33), d.h. $e^{[\nu]} = x^{[\nu]} - p^{[C]}$, kann nun die Fehlergleichung hergeleitet werden. Eine analoge Vorgangsweise zu jener aus Abschnitt 1.2.3 liefert

$$Lx^{[\nu]} - Lp^{[C]} = Lx^{[0]} - Lz^{[\nu-1]} + Lx^{[\nu-1]} - Lp^{[C]}, \quad (1.131)$$

$$Le^{[\nu]} = \mathcal{F}(x^{[0]}) - \mathcal{F}(z^{[\nu-1]}) - \bar{d}^{[\nu-1]} + L(e^{[\nu-1]}). \quad (1.132)$$

Setzt man für den Defekt die Identität (1.128) ein, erhält man

$$Le^{[\nu]} = \mathcal{F}(x^{[0]}) - \mathcal{F}(z^{[\nu-1]}) - Lx^{[\nu-1]} + G(x^{[\nu-1]}) + L(e^{[\nu-1]}). \quad (1.133)$$

Für den linearen Fall kann noch weiter zusammengefasst werden: Addieren des verschwindenden Defekts am Fixpunkt gemäß (1.129) und Vereinfachen liefert

$$Le^{[\nu]} = \mathcal{F}(x^{[0]}) - \mathcal{F}(z^{[\nu-1]}) + G_H(e^{[\nu-1]}). \quad (1.134)$$

Die Fehlergleichung ergibt sich durch Addition von $\mathcal{F}_H(e^{[\nu-1]}) - \mathcal{F}_H(e^{[\nu-1]}) = 0$ und Ausnutzen der Linearität der Differentialgleichung

$$Le^{[\nu]} = \underbrace{\mathcal{F}(x^{[0]}) - \mathcal{F}(z^{[\nu-1]}) + \mathcal{F}_H(e^{[\nu-1]})}_{=\mathcal{F}_H(e^{[\nu]})} + \underbrace{G_H(e^{[\nu-1]}) - \mathcal{F}_H(e^{[\nu-1]})}_{=\Delta\mathcal{F}_H(e^{[\nu-1]})}, \quad (1.135)$$

$$Le^{[\nu]} = \mathcal{F}_H(e^{[\nu]}) + \Delta\mathcal{F}_H(e^{[\nu-1]}), \quad (1.136)$$

wobei der Index H anzeigt, dass es sich hierbei um den homogenen Anteil von F bzw. G handelt. Die entsprechenden Ordnungen erhält man durch Analyse des Terms $\Delta\mathcal{F}_H(e^{[\nu-1]})$. An dieser Stelle wird jedoch darauf verzichtet und auf Abschnitt 2.2.2 verwiesen, wo bei der Beweisführung für implizite Differentialgleichungen derselbe Term auftritt.

1.3.3 IQDeC-EEUL – Konvergenz gegen die exakte Lösung

Satz 1.39. (*IQDeC-EEUL – Konvergenz gegen die exakte Lösung*) Löst man das Anfangswertproblem vom Typ (1.1) auf einem beliebigen Gitter mit m Kollokationsknoten bei insgesamt $m + 1$ Gitterpunkten gemäß Definition 1.1 mittels iterierter Defektkorrektur mit Integralmitteldefekt, so erfüllt der Fehler $e_{j,l}^{*,[\nu]}$ der Euler-Lösungen $x_{j,l}^{[\nu]}$ der einzelnen Iterationen bezüglich der exakten Lösung $u^*(t_{j,l})$ folgende Bedingung

$$\left\| e_{j,l}^{*,[\nu]} \right\| = \left\| x_{j,l}^{[\nu]} - u^*(t_{j,l}) \right\| = \mathcal{O}(h^{\nu+1}), \quad \nu = 0, 1, \dots, m-1. \quad (1.137)$$

IQDeC-EEUL – Aufstellen der Fehlergleichung für Satz 1.39

Aufbauend auf Abschnitt 1.3.2 gilt folgendes: Setzt man die exakte Lösung u^* von (1.125) in den Integralmitteldefekt \bar{d} ein, erhält man den Quadraturfehler

$$\Delta G(u^*) = Lu^* - G(u^*). \quad (1.138)$$

Analoge Überlegungen zu jenen des vorigen Abschnitts führen zur Fehlergleichung (vergleiche mit (1.136))

$$Le^{[\nu]} = \mathcal{F}_H(e^{[\nu]}) + \Delta\mathcal{F}_H(e^{[\nu-1]}) + \Delta G(u^*). \quad (1.139)$$

Man sieht, dass man im Vergleich zur Fixpunktconvergenz wie bei IDeC (mit klassischem Defekt) einen Term (Quadraturfehler der exakten Lösung u^*) mehr auf seine Ordnung hin zu analysieren hätte. Auf weitere Ausführungen wird in dieser Arbeit allerdings verzichtet.

Bei sämtlichen in den folgenden Abschnitten 1.3.4 bis 1.3.8 vorgestellten Verfahren bleibt der Integralmitteldefekt (1.120) gleich, da es durch die Mittelwertbildung egal ist, an welcher (Zwischen-)Stelle (linksseitig, rechtsseitig oder in der Mitte zwischen zwei Knoten) der f -Term ausgewertet wird. Der Integralmitteldefekt ist im Gegensatz zum klassischen Defekt nicht punktweise, sondern bezüglich des Intervalls definiert. Sämtliche Basisverfahren wurden bereits in Abschnitt 1.2 vorgestellt.

1.3.4 IQDeC-Heun

Mit dem Basisverfahren (1.78) und dem Integralmitteldefekt (1.120) kann das Nachbarproblem wie folgt angeschrieben werden

$$Z_1^{[\nu]} = z_{j,l-1}^{[\nu]} + h_{j,l} \left[f(t_{j,l-1}, z_{j,l-1}^{[\nu]}) + \bar{d}_{j,l}^{[\nu]} \right], \quad (1.140a)$$

$$z_{j,l}^{[\nu]} = z_{j,l-1}^{[\nu]} + \frac{h_{j,l}}{2} \left[f(t_{j,l-1}, z_{j,l-1}^{[\nu]}) + f(t_{j,l}, Z_1^{[\nu]}) + 2\bar{d}_{j,l}^{[\nu]} \right]. \quad (1.140b)$$

Beispiel 1.40. Wie Beispiel 1.36, jedoch mit dem Verfahren von Heun als Basisverfahren. Tabelle 1.23 zeigt die Resultate.

Bemerkung 1.41. Das IQDeC-Heun-Verfahren konvergiert zwar wie erwartet gegen den Fixpunkt, allerdings so wie alle IQDeC*-Verfahren mit einem Basisverfahren der Ordnung 2 am *nicht* äquidistanten Gitter nur mit der Ordnungsfolge 2,3,... Die Ordnung des Basisverfahrens überträgt sich beim Integralmitteldefekt somit nicht auf das Defektkorrekturverfahren. Da Beispiel 1.40 nach der ersten Iteration mit den Ordnungen knapp unter 4 einen falschen Eindruck der Verfahrensgenauigkeit vortäuscht, wird auf Beispiel 1.42 verwiesen.

Beispiel 1.42. Wie Beispiel 1.36, jedoch mit dem Verfahren von Heun als Basisverfahren und Polynomgrad 6 mit $c_1 = 0, 2567$; $c_2 = 0, 3343$; $c_3 = 0, 6998$; $c_4 = 0, 8125$; $c_5 = 0, 9007$; $c_6 = 1$. Tabelle 1.24 zeigt die Resultate.

Beispiel 1.43. Wie Beispiel 1.36, jedoch mit dem Verfahren von Heun als Basisverfahren und auf einem äquidistanten Gitter. Tabelle 1.25 zeigt die Resultate.

Beispiel 1.44. Wie Beispiel 1.36, jedoch mit dem Verfahren von Heun als Basisverfahren auf einem äquidistanten Gitter und Polynomgrad 3. Tabelle 1.26 zeigt die Resultate.

Bemerkung 1.45. Man sieht, dass man bei ungeradem Polynomgrad im Gegensatz zur Defektkorrektur mit klassischem Defekt (siehe Abschnitt 1.2) *keine* Konvergenzordnung „geschenkt“ bekommt. Dies ist insofern zu erwarten, da das Verfahren mit Integralmitteldefekt gegen die Euler-Kollokation (genauer IEUL-Kollokation) konvergiert und diese im Gegensatz zur Heun-Kollokation bei ungeradem Polynomgrad ebenfalls keine zusätzliche Konvergenzordnung liefert (siehe dazu die Abschnitte 5.1 und 5.2).

h	Heun	IQDeC1	IQDeC2	IQDeC3	COLL
0,1	2,77E-04	4,16E-08	6,47E-09	5,55E-09	5,55E-09
0,05	6,93E-05	2,76E-09	4,04E-10	3,47E-10	3,47E-10
0,025	1,73E-05	1,93E-10	2,52E-11	2,17E-11	2,17E-11
0,0125	4,33E-06	1,46E-11	1,58E-12	1,35E-12	1,35E-12
0,1	2,00	3,91	4,00	4,00	4,00
0,05	2,00	3,84	4,00	4,00	4,00
0,025	2,00	3,73	4,00	4,00	4,00
0,0125	2,00	3,73	4,00	4,00	4,00

Tabelle 1.23: Resultate zu Beispiel 1.40.

h	Heun	IQDeC1	IQDeC2	IQDeC3	IQDeC4	IQDeC5	COLL
0,1	2,13E-04	3,41E-07	4,66E-08	2,60E-11	2,50E-12	5,45E-12	5,46E-12
0,05	5,32E-05	3,95E-08	2,91E-09	6,54E-13	4,08E-14	8,51E-14	8,51E-14
0,025	1,33E-05	4,75E-09	1,82E-10	1,79E-14	6,50E-16	1,33E-15	1,33E-15
0,0125	3,33E-06	5,82E-10	1,13E-11	5,72E-16	1,03E-17	2,08E-17	2,08E-17
0,1	2,00	3,11	4,00	5,32	5,94	6,00	6,00
0,05	2,00	3,06	4,00	5,19	5,97	6,00	6,00
0,025	2,00	3,03	4,00	5,10	5,99	6,00	6,00
0,0125	2,00	3,03	4,00	5,10	5,99	6,00	6,00

Tabelle 1.24: Resultate zu Beispiel 1.42.

h	Heun	IQDeC1	IQDeC2	COLL
0,1	1,92E-04	2,11E-08	3,81E-08	3,81E-08
0,05	4,80E-05	1,31E-09	2,38E-09	2,37E-09
0,025	1,20E-05	8,19E-11	1,49E-10	1,49E-10
0,0125	3,00E-06	5,12E-12	9,28E-12	9,28E-12
0,1	2,00	4,00	4,00	4,00
0,05	2,00	4,00	4,00	4,00
0,025	2,00	4,00	4,00	4,00
0,0125	2,00	4,00	4,00	4,00

Tabelle 1.25: Resultate zu Beispiel 1.43.

h	Heun	IQDeC1	IQDeC2	COLL
0,1	3,41E-04	7,16E-07	8,46E-07	8,46E-07
0,05	8,53E-05	8,71E-08	9,52E-08	9,52E-08
0,025	2,13E-05	1,07E-08	1,12E-08	1,12E-08
0,0125	5,34E-06	1,33E-09	1,36E-09	1,36E-09
0,1	2,00	3,04	3,15	3,15
0,05	2,00	3,02	3,08	3,08
0,025	2,00	3,01	3,04	3,04
0,0125	2,00	3,01	3,04	3,04

Tabelle 1.26: Resultate zu Beispiel 1.44.

h	EMR	IQDeC1	IQDeC2	IQDeC3	COLL
0,1	4,44E-05	2,63E-10	5,71E-09	5,55E-09	5,55E-09
0,05	1,11E-05	5,60E-11	3,57E-10	3,47E-10	3,47E-10
0,025	2,78E-06	1,26E-11	2,23E-11	2,17E-11	2,17E-11
0,0125	6,96E-07	1,93E-12	1,39E-12	1,35E-12	1,35E-12
0,1	2,00	2,23	4,00	4,00	4,00
0,05	2,00	2,15	4,00	4,00	4,00
0,025	2,00	2,70	4,00	4,00	4,00
0,0125	2,00	2,70	4,00	4,00	4,00

Tabelle 1.27: Resultate zu Beispiel 1.46.

1.3.5 IQDeC-EMR

Analog zu 1.3.4 gilt für das Nachbarproblem der expliziten Mittelpunktsregel

$$Z_{\frac{1}{2}}^{[\nu]} = z_{j,l-1}^{[0]} + \frac{h_{j,l}}{2} \left[f \left(t_{j,l-1}, z_{j,l-1}^{[0]} \right) + \bar{d}_{j,l}^{[\nu]} \right], \quad (1.141a)$$

$$z_{j,l}^{[0]} = z_{j,l-1}^{[0]} + h_{j,l} \left[f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, Z_{\frac{1}{2}}^{[\nu]} \right) + \bar{d}_{j,l}^{[\nu]} \right], \quad (1.141b)$$

bei unverändertem Basisverfahren (1.94) und Defekt (1.120).

Beispiel 1.46. Wie Beispiel 1.36, jedoch mit der expliziten Mittelpunktsregel als Basisverfahren. Tabelle 1.27 zeigt die Resultate.

Beispiel 1.47. Wie Beispiel 1.36, jedoch mit der expliziten Mittelpunktsregel als Basisverfahren auf einem äquidistanten Gitter. Tabelle 1.28 zeigt die Resultate.

1.3.6 IQDeC-IEUL

Mit dem impliziten Euler-Verfahren (1.102), dem Integralmitteldefekt (1.120) und dem Nachbarproblem

$$\frac{z_{j,l}^{[\nu]} - z_{j,l-1}^{[\nu]}}{h_{j,l}} = f \left(t_{j,l}, z_{j,l}^{[\nu]} \right) + \bar{d}_{j,l}^{[\nu]}, \quad (1.142)$$

h	EMR	IQDeC1	IQDeC2	COLL
0,1	3,08E-05	4,10E-08	3,81E-08	3,81E-08
0,05	7,71E-06	2,56E-09	2,38E-09	2,37E-09
0,025	1,93E-06	1,60E-10	1,49E-10	1,49E-10
0,0125	4,82E-07	1,00E-11	9,28E-12	9,28E-12
0,1	2,00	4,00	4,00	4,00
0,05	2,00	4,00	4,00	4,00
0,025	2,00	4,00	4,00	4,00
0,0125	2,00	4,00	4,00	4,00

Tabelle 1.28: Resultate zu Beispiel 1.47.

h	IEUL	IQDeC1	IQDeC2	IQDeC3	IQDeC4	COLL
0,1	1,95E-03	1,94E-04	2,91E-06	3,14E-08	5,73E-09	5,55E-09
0,05	1,01E-03	4,96E-05	3,53E-07	1,84E-09	3,51E-10	3,47E-10
0,025	5,15E-04	1,25E-05	4,33E-08	1,11E-11	2,18E-11	2,17E-11
0,0125	2,60E-04	3,15E-06	5,37E-09	6,80E-12	1,36E-12	1,35E-12
0,1	0,95	1,97	3,05	4,10	4,03	4,00
0,05	0,97	1,98	3,03	4,05	4,01	4,00
0,025	0,99	1,99	3,01	4,03	4,00	4,00
0,0125	0,99	1,99	3,01	4,03	4,00	4,00

Tabelle 1.29: Resultate zu Beispiel 1.48.

ist IQDeC-IEUL komplett.

Beispiel 1.48. Wie Beispiel 1.36, jedoch mit dem impliziten Euler-Verfahren als Basisverfahren. Tabelle 1.29 zeigt die Resultate.

Beispiel 1.49. Wie Beispiel 1.36, jedoch mit dem impliziten Euler-Verfahren als Basisverfahren auf einem äquidistanten Gitter. Tabelle 1.30 zeigt die Resultate.

1.3.7 IQDeC-ITR

Für das Verfahren mit der impliziten Trapezregel (1.105) als Basisverfahren und dem Defekt (1.120) wird das Nachbarproblem wie folgt definiert:

$$\frac{z_{j,l}^{[\nu]} - z_{j,l-1}^{[\nu]}}{h_{j,l}} = \frac{1}{2} \left[f \left(t_{j,l-1}, z_{j,l-1}^{[\nu]} \right) + f \left(t_{j,l}, z_{j,l}^{[\nu]} \right) \right] + \bar{d}_{j,l}^{[\nu]}. \quad (1.143)$$

Beispiel 1.50. Wie Beispiel 1.36, jedoch mit der impliziten Trapezregel als Basisverfahren. Tabelle 1.31 zeigt die Resultate.

Beispiel 1.51. Wie Beispiel 1.36, jedoch mit der impliziten Trapezregel als Basisverfahren auf einem äquidistanten Gitter. Tabelle 1.32 zeigt die Resultate.

h	IEUL	IQDeC1	IQDeC2	IQDeC3	IQDeC4	COLL
0,1	1,71E-03	1,63E-04	2,06E-06	2,17E-08	3,80E-08	3,81E-08
0,05	8,85E-03	4,16E-05	2,54E-07	1,42E-09	2,37E-09	2,38E-09
0,025	4,50E-04	1,05E-05	3,14E-08	9,06E-11	1,48E-10	1,49E-10
0,0125	2,27E-04	2,63E-06	3,91E-09	5,73E-12	9,28E-12	9,28E-12
0,1	0,95	1,97	3,02	3,94	4,00	4,00
0,05	0,98	1,99	3,01	3,97	4,00	4,00
0,025	0,99	1,99	3,01	3,98	4,00	4,00
0,0125						

Tabelle 1.30: Resultate zu Beispiel 1.49.

h	ITR	IQDeC1	IQDeC2	IQDeC3	COLL
0,1	9,40E-05	1,19E-08	5,84E-09	5,55E-09	5,55E-09
0,05	2,35E-05	7,51E-10	3,65E-10	3,47E-10	3,47E-10
0,025	5,88E-06	4,78E-11	2,28E-11	2,17E-11	2,17E-11
0,0125	1,47E-06	3,10E-12	1,42E-12	1,35E-12	1,35E-12
0,1	2,00	3,99	4,00	4,00	4,00
0,05	2,00	3,97	4,00	4,00	4,00
0,025	2,00	3,94	4,00	4,00	4,00
0,0125					

Tabelle 1.31: Resultate zu Beispiel 1.50.

h	ITR	IQDeC1	IQDeC2	COLL
0,1	6,52E-05	3,51E-08	3,81E-08	3,81E-08
0,05	1,63E-05	2,19E-09	2,38E-09	2,37E-09
0,025	4,07E-06	1,37E-10	1,49E-10	1,49E-10
0,0125	1,02E-06	8,57E-12	9,28E-12	9,28E-12
0,1	2,00	4,00	4,00	4,00
0,05	2,00	4,00	4,00	4,00
0,025	2,00	4,00	4,00	4,00
0,0125				

Tabelle 1.32: Resultate zu Beispiel 1.51.

h	IMR	IQDeC1	IQDeC2	IQDeC3	COLL
0,1	1,39E-04	1,09E-08	5,09E-09	5,55E-09	5,55E-09
0,05	3,46E-05	6,02E-10	3,18E-10	3,47E-10	3,47E-10
0,025	8,66E-06	2,75E-11	1,99E-11	2,17E-11	2,17E-11
0,0125	2,16E-06	4,66E-13	1,24E-12	1,35E-12	1,35E-12
0,1	2,00	4,18	4,00	4,00	4,00
0,05	2,00	4,45	4,00	4,00	4,00
0,025	2,00	5,88	4,00	4,00	4,00
0,0125	2,00	5,88	4,00	4,00	4,00

Tabelle 1.33: Resultate zu Beispiel 1.52.

h	IMR	IQDeC1	IQDeC2	COLL
0,1	9,60E-05	3,35E-08	3,81E-08	3,81E-08
0,05	2,40E-05	2,09E-09	2,38E-09	2,37E-09
0,025	6,00E-06	1,31E-10	1,49E-10	1,49E-10
0,0125	1,50E-06	8,18E-12	9,28E-12	9,28E-12
0,1	2,00	4,00	4,00	4,00
0,05	2,00	4,00	4,00	4,00
0,025	2,00	4,00	4,00	4,00
0,0125	2,00	4,00	4,00	4,00

Tabelle 1.34: Resultate zu Beispiel 1.53.

1.3.8 IQDeC-IMR

Analog zu IQDeC-ITR gilt für die Variante mit der impliziten Mittelpunktsregel (1.107) als Basisverfahren für das Nachbarproblem

$$\frac{z_{j,l}^{[\nu]} - z_{j,l-1}^{[\nu]}}{h_{j,l}} = f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, \frac{z_{j,l-1}^{[\nu]} + z_{j,l}^{[\nu]}}{2} \right) + \bar{d}_{j,l}^{[\nu]}, \quad (1.144)$$

wobei für den Defekt $\bar{d}_{j,l}^{[\nu]}$ Beziehung (1.120) gilt.

Beispiel 1.52. Wie Beispiel 1.36, jedoch mit der impliziten Mittelpunktsregel als Basisverfahren. Tabelle 1.33 zeigt die Resultate.

Beispiel 1.53. Wie Beispiel 1.36, jedoch mit der impliziten Mittelpunktsregel als Basisverfahren auf einem äquidistanten Gitter. Tabelle 1.34 zeigt die Resultate.

1.4 Interpolierter Defekt – IPDeC

Auch diese Defektvariante wurde erstmals in [2] vorgestellt. Als Vorbereitung auf die Behandlung impliziter Differentialgleichungen mittels iterierter Defektkorrektur mit interpoliertem

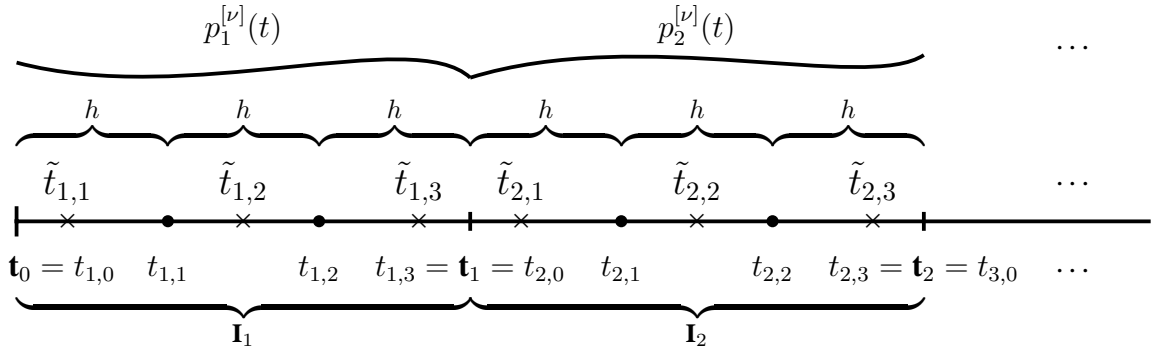


Abbildung 1.4: Schematische Darstellung des IPDeC-Gitters für 3 Gauß-Knoten ($= \tilde{t}_{j,\kappa}$) und 4 äquidistant liegende Knoten ($= t_{j,l}$).

Defekt, kurz IPDeC für *Iterated Defect Correction with Defect Interpolation*, wird der Algorithmus wiedergegeben. Diese Art des Defekts zielt darauf ab superkonvergente Verfahren mit maximaler Konvergenzordnung (Gauß-Gitter) zu konstruieren.

1.4.1 IPDeC-EEUL

Für diese Art der Defektkorrektur werden ein äquidistantes Gitter $\{t_{j,l}, l = 1, \dots, m\}$ und ein beliebiges nicht äquidistantes Gitter $\{\tilde{t}_{j,\kappa}, \kappa = 1, \dots, m\}$ (meist ein Gitter mit Gauß-Knoten) parallel verwendet (siehe dazu Abbildung 1.4). Für das Verfahren wird der in Abschnitt 1.2 beschriebene Algorithmus wie folgt modifiziert: Der Defekt (1.16) wird nicht an den Stellen $t_{j,l}$ ausgewertet, sondern an den Knoten $\tilde{t}_{j,\kappa}$ des Gauß-Gitters. D.h.

$$\tilde{d}^{[\nu]}(\tilde{t}_{j,\kappa}) = p^{[\nu]'}(\tilde{t}_{j,\kappa}) - f(\tilde{t}_{j,\kappa}, p^{[\nu]}(\tilde{t}_{j,\kappa})), \quad (1.145)$$

wobei die stückweise definierten Polynomfunktion $p^{[\nu]}(t)$ die Werte $x_{j,l}^{[\nu]}$ in den jeweiligen Intervallen \mathbf{I}_j interpoliert. Die so berechneten Werte $\tilde{d}^{[\nu]}(\tilde{t}_{j,\kappa})$ werden wiederum stückweise mittels der Polynome $\tilde{p}^{[\nu]}(t)$ (Grad $\leq m - 1$) interpoliert. Wertet man diese Polynome an den Knoten $t_{j,l}$ aus, gelangt man zum IPDeC-Defekt

$$\hat{d}_{j,l}^{[\nu]} = \tilde{p}^{[\nu]}(t_{j,l}), \quad (1.146)$$

welcher anstatt von (1.16) im ansonsten unveränderten Algorithmus aus Abschnitt 1.2.1 im Nachbarproblem (1.15a) zu berücksichtigen ist.

Aufgrund der Definition des interpolierten Defekts ist der Fixpunkt des Verfahrens jene Kollokationslösung $p^{[C]}(t)$, für die der Defekt an den Knoten $\tilde{t}_{j,\kappa}$ verschwindet. Wählt man als zugrunde liegendes Gitter ein Gauß-Gitter, erhält man ein Verfahren mit höchstmöglicher Konvergenzordnung (Superkonvergenz).

Bemerkung 1.54. Wie numerische Experimente belegen, muss die Anzahl der Gauß-Knoten exakt auf die Anzahl der Knoten des äquidistanten Gitters abgestimmt sein. Genauer gesagt, muss für Differentialgleichungen 1. Ordnung die Anzahl der Kollokationsknoten des äquidistanten Gitters (=Anzahl der Gitterpunkte -1) gleich groß sein wie die Anzahl der Gauß-Knoten.

h	EEUL	IPDeC1	IPDeC2	IPDeC3	IPDeC4	IPDeC5	G-COLL
0,1	3,31E-02	1,83E-03	1,16E-05	2,36E-06	3,08E-08	5,88E-09	3,34E-12
0,05	1,65E-02	4,55E-04	1,91E-06	1,44E-07	6,62E-10	8,81E-11	5,23E-14
0,025	8,21E-03	1,13E-04	2,66E-07	8,85E-09	1,60E-11	1,34E-12	8,18E-16
0,0125	4,10E-03	2,83E-05	3,50E-08	5,49E-10	4,25E-13	2,07E-14	1,28E-17
0,1	1,00	2,01	2,60	4,04	5,54	6,06	6,00
0,05	1,00	2,00	2,84	4,02	5,37	6,03	6,00
0,025	1,00	2,00	2,93	4,01	5,23	6,01	6,00
0,0125	1,00	2,00	2,93	4,01	5,23	6,01	6,00

Tabelle 1.35: Resultate zu Beispiel 1.55.

Beispiel 1.55. Gegeben sei das skalare nichtlineare Anfangswertproblem (1.19) mit der exakten Lösung $u^*(t) = \cos(t)$. Die IPDeC-EEUL-Lösung erfolgt auf einem äquidistanten Gitter mit 3 Kollokationsknoten und einem nicht äquidistanten Gitter mit 3 Gauß-Knoten gemäß Abbildung 1.4 ($c_1 = \frac{1}{3}$, $c_2 = \frac{2}{3}$, $c_3 = 1$ und $\tilde{c}_1 = \frac{1}{2} - \frac{\sqrt{15}}{10}$, $\tilde{c}_2 = \frac{1}{2}$, $\tilde{c}_3 = \frac{1}{2} + \frac{\sqrt{15}}{10}$ für $t_{j,l} = \mathbf{t}_{j-1} + \tilde{c}_l \mathbf{h}$) im Intervall $[t_0, t_{end}] = [\mathbf{t}_0, \mathbf{t}_N] = [0, 3]$. Die globalen Fehler bezüglich der exakten Lösung $e^{*,[\#]}(t_{end}) = p^{[\#]}(t_{end}) - u^*(t_{end})$, $\# = \nu, C$ und die daraus berechneten Konvergenzordnungen an der Stelle $t_{end} = 3$ für das Basisverfahren, IPDeC und das zugehörige Fixpunkt-Kollokationsverfahren (=Gauß-Kollokation), sind in Tabelle 1.35 aufgelistet.

1.4.2 IPDeC-EEUL – Fixpunktconvergenz

Satz 1.56. (*Fixpunktconvergenz von IPDeC-EEUL*) Löst man das Anfangswertproblem vom Typ (1.1) mittels der in Abschnitt 1.4.1 beschriebenen iterierten Defektkorrektur mit interpoliertem Defekt, so erfüllt der Fehler $e_{j,l}^{[\nu]}$ der Defektkorrektur-Lösungen $x_{j,l}^{[\nu]}$ der einzelnen Iterationen bezüglich des Fixpunktes $p_{j,l}^{[C]}$ (=Kollokationslösung auf dem $\tilde{t}_{j,\kappa}$ -Gitter) die Abschätzung

$$\left\| e_{j,l}^{[\nu]} \right\| = \left\| x_{j,l}^{[\nu]} - p_{j,l}^{[C]} \right\| = \mathcal{O}(h^{\nu+1}), \quad \nu = 0, 1, \dots \quad (1.147)$$

Der Beweis von Satz 1.56 ist, allerdings mit IEUL-Basisverfahren, im Anhang von [2] ausführlich dargestellt, auf eine Wiederholung wird verzichtet. Festgehalten sei nur, dass die Inhomogenität der Fehlergleichung in zwei Teile zerlegt werden kann: Der erste Teil ist ident mit der Inhomogenität in (1.42), der zweite Teil repräsentiert den Fehler, der bei der Defektinterpolation entsteht.

1.4.3 IPDeC-Heun

Mit dem Basisverfahren (1.78) und dem interpolierten Defekt (1.146) kann das Nachbarproblem wie folgt angeschrieben werden:

$$Z_1^{[\nu]} = z_{j,l-1}^{[\nu]} + h_{j,l} \left[f \left(t_{j,l-1}, z_{j,l-1}^{[\nu]} \right) + \hat{d}_{j,l-1}^{[\nu]} \right], \quad (1.148a)$$

$$z_{j,l}^{[\nu]} = z_{j,l-1}^{[\nu]} + \frac{h_{j,l}}{2} \left[f \left(t_{j,l-1}, z_{j,l-1}^{[\nu]} \right) + f \left(t_{j,l}, Z_1^{[\nu]} \right) + \hat{d}_{j,l-1}^{[\nu]} + \hat{d}_{j,l}^{[\nu]} \right]. \quad (1.148b)$$

h	Heun	IPDeC1	IPDeC2	IPDeC3	G-COLL
0,1	4,30E-04	1,39E-08	8,81E-12	3,42E-12	3,34E-12
0,05	1,06E-04	8,44E-10	1,29E-13	5,26E-14	5,23E-14
0,025	2,65E-05	5,17E-11	1,95E-15	8,19E-16	8,18E-16
0,0125	6,60E-06	3,20E-12	3,01E-17	1,28E-17	1,28E-17
0,1	2,01	4,05	6,10	6,02	6,00
0,05	2,01	4,03	6,04	6,01	6,00
0,025	2,00	4,02	6,02	6,00	6,00
0,0125	2,00	4,02	6,02	6,00	6,00

Tabelle 1.36: Resultate zu Beispiel 1.57.

h	Heun	IPDeC1	IPDeC2	IPDeC3	IPDeC4	IPDeC5	G-COLL
0,1	1,54E-04	9,33E-10	1,55E-12	5,46E-17	2,26E-18	1,24E-20	1,31E-20
0,05	3,82E-05	5,52E-11	2,48E-14	3,44E-19	2,16E-21	1,27E-23	1,29E-23
0,025	9,51E-06	3,35E-12	3,92E-16	1,51E-21	2,08E-24	1,26E-26	1,26E-26
0,0125	2,37E-06	2,06E-13	6,17E-18	6,14E-24	2,01E-27	1,23E-29	1,23E-29
0,1	2,01	4,08	5,96	7,31	10,03	9,93	9,99
0,05	2,00	4,04	5,98	7,83	10,02	9,98	10,00
0,025	2,00	4,02	5,99	7,94	10,01	10,00	10,00
0,0125	2,00	4,02	5,99	7,94	10,01	10,00	10,00

Tabelle 1.37: Resultate zu Beispiel 1.58.

Beispiel 1.57. Wie Beispiel 1.55, jedoch mit dem Verfahren von Heun als Basisverfahren. Tabelle 1.36 zeigt die Resultate.

Beispiel 1.58. Wie Beispiel 1.55, jedoch mit dem Verfahren von Heun als Basisverfahren und 5 äquidistanten Kollokations- bzw. 5 Gauß-Knoten: $\tilde{c}_1 = \frac{1}{2} - \frac{\sqrt{245+14\sqrt{70}}}{42}$; $\tilde{c}_2 = \frac{1}{2} - \frac{\sqrt{245-14\sqrt{70}}}{42}$; $\tilde{c}_3 = \frac{1}{2}$; $\tilde{c}_4 = \frac{1}{2} + \frac{\sqrt{245-14\sqrt{70}}}{42}$; $\tilde{c}_5 = \frac{1}{2} + \frac{\sqrt{245+14\sqrt{70}}}{42}$. Tabelle 1.37 zeigt die Resultate.

1.4.4 IPDeC-EMR

Analog zu Abschnitt 1.4.3 gilt für das Nachbarproblem der expliziten Mittelpunktsregel

$$Z_{\frac{1}{2}}^{[\nu]} = z_{j,l-1}^{[0]} + \frac{h_{j,l}}{2} \left[f \left(t_{j,l-1}, z_{j,l-1}^{[0]} \right) + \hat{d}_{j,l-1}^{[\nu]} \right], \quad (1.149a)$$

$$z_{j,l}^{[0]} = z_{j,l-1}^{[0]} + h_{j,l} \left[f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, Z_{\frac{1}{2}}^{[\nu]} \right) + \hat{d}_{\frac{1}{2}}^{[\nu]} \right], \quad (1.149b)$$

bei unverändertem Basisverfahren (1.94) und den Defekten (1.146) bzw.

$$\hat{d}_{\frac{1}{2}}^{[\nu]} := \hat{d}^{[\nu]} \left(t_{j,l-1} + \frac{h_{j,l}}{2} \right) = \tilde{p}^{[\nu]} \left(t_{j,l-1} + \frac{h_{j,l}}{2} \right). \quad (1.150)$$

h	EMR	IPDeC1	IPDeC2	IPDeC3	G-COLL
0,1	1,34E-04	4,94E-08	2,66E-11	3,36E-12	3,34E-12
0,05	3,39E-05	3,18E-09	4,30E-13	5,24E-14	5,23E-14
0,025	8,54E-06	2,01E-10	6,83E-15	8,18E-16	8,18E-16
0,0125	2,14E-06	1,27E-11	1,08E-16	1,28E-17	1,28E-17
0,1	1,98	3,96	5,95	6,00	6,00
0,05	1,99	3,98	5,98	6,00	6,00
0,025	1,99	3,99	5,99	6,00	6,00
0,0125					

Tabelle 1.38: Resultate zu Beispiel 1.59.

h	IEUL	IPDeC1	IPDeC2	IPDeC3	IPDeC4	IPDeC5	G-COLL
0,1	3,23E-02	1,77E-03	2,51E-05	2,08E-06	8,05E-09	4,73E-09	3,34E-12
0,05	1,63E-02	4,47E-04	2,75E-06	1,35E-07	5,56E-11	7,90E-11	5,23E-14
0,025	8,15E-03	1,12E-04	3,19E-07	8,58E-09	6,50E-12	1,27E-12	8,18E-16
0,0125	4,08E-03	2,82E-05	3,82E-08	5,40E-10	2,77E-13	2,02E-14	1,28E-17
0,1	0,99	1,99	3,19	3,95	7,18	5,91	6,00
0,05	1,00	1,99	3,11	3,97	3,09	5,96	6,00
0,025	1,00	2,00	3,06	3,99	4,55	5,98	6,00
0,0125							

Tabelle 1.39: Resultate zu Beispiel 1.60.

Beispiel 1.59. Wie Beispiel 1.55, jedoch mit der expliziten Mittelpunktsregel als Basisverfahren. Tabelle 1.38 zeigt die Resultate.

1.4.5 IPDeC-IEUL

Mit dem impliziten Euler-Verfahren (1.102), dem interpolierten Defekt (1.146) und dem Nachbarproblem

$$\frac{z_{j,l}^{[\nu]} - z_{j,l-1}^{[\nu]}}{h_{j,l}} = f\left(t_{j,l}, z_{j,l}^{[\nu]}\right) + \hat{d}_{j,l}^{[\nu]} \quad (1.151)$$

ist IPDeC-IEUL komplett.

Beispiel 1.60. Wie Beispiel 1.55, jedoch mit dem impliziten Euler-Verfahren als Basisverfahren. Tabelle 1.39 zeigt die Resultate.

1.4.6 IPDeC-ITR

Für das IPDeC-Verfahren mit der impliziten Trapezregel (1.105) als Basisverfahren und dem Defekt (1.146) wird das Nachbarproblem wie folgt definiert:

$$\frac{z_{j,l}^{[\nu]} - z_{j,l-1}^{[\nu]}}{h_{j,l}} = \frac{1}{2} \left[f\left(t_{j,l-1}, z_{j,l-1}^{[\nu]}\right) + f\left(t_{j,l}, z_{j,l}^{[\nu]}\right) + \hat{d}_{j,l-1}^{[\nu]} + \hat{d}_{j,l}^{[\nu]} \right]. \quad (1.152)$$

h	ITR	IPDeC1	IPDeC2	IPDeC3	G-COLL
0,1	3,48E-04	7,96E-08	6,77E-12	3,30E-12	3,34E-12
0,05	8,71E-05	4,97E-09	1,07E-13	5,22E-14	5,23E-14
0,025	2,18E-05	3,11E-10	1,68E-15	8,17E-16	8,18E-16
0,0125	5,44E-06	1,94E-11	2,62E-17	1,28E-17	1,28E-17
0,1	2,00	4,00	5,98	5,98	6,00
0,05	2,00	4,00	6,00	6,00	6,00
0,025	2,00	4,00	6,00	6,00	6,00
0,0125	2,00	4,00	6,00	6,00	6,00

Tabelle 1.40: Resultate zu Beispiel 1.61.

h	IMR	IPDeC1	IPDeC2	IPDeC3	G-COLL
0,1	2,11E-04	1,69E-09	1,73E-12	3,35E-12	3,34E-12
0,05	5,27E-05	1,04E-10	2,71E-14	5,24E-14	5,23E-14
0,025	1,32E-05	6,50E-12	4,24E-16	8,18E-16	8,18E-16
0,0125	3,29E-06	4,06E-13	6,63E-18	1,28E-17	1,28E-17
0,1	2,00	4,02	6,00	6,00	6,00
0,05	2,00	4,00	6,00	6,00	6,00
0,025	2,00	4,00	6,00	6,00	6,00
0,0125	2,00	4,00	6,00	6,00	6,00

Tabelle 1.41: Resultate zu Beispiel 1.62.

Beispiel 1.61. Wie Beispiel 1.55, jedoch mit der impliziten Trapezregel als Basisverfahren. Tabelle 1.40 zeigt die Resultate.

1.4.7 IPDeC-IMR

Analog zu IPDeC-ITR gilt für die Variante mit der impliziten Mittelpunktsregel (1.107) als Basisverfahren für das Nachbarproblem

$$\frac{z_{j,l}^{[\nu]} - z_{j,l-1}^{[\nu]}}{h_{j,l}} = f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, \frac{z_{j,l-1}^{[\nu]} + z_{j,l}^{[\nu]}}{2} \right) + \bar{d}_{\frac{1}{2}}^{[\nu]}, \quad (1.153)$$

wobei für den Defekt $\bar{d}_{\frac{1}{2}}^{[\nu]}$ Beziehung (1.150) gilt.

Beispiel 1.62. Wie Beispiel 1.55, jedoch mit der impliziten Mittelpunktsregel als Basisverfahren. Tabelle 1.41 zeigt die Resultate.

Kapitel 2

Implizite Differentialgleichungen erster Ordnung

Dieses Kapitel behandelt die direkte Lösung von impliziten Anfangswertproblemen erste Ordnung.

2.1 Klassischer Defekt – IIDeC

2.1.1 IIDeC-IEUL

Die im Folgenden vorgestellten Verfahren wurden für Anfangswertprobleme vom Typ

$$F(t, u(t), u'(t)) = 0, \quad (2.1a)$$

$$u(t_0) = u_0 \quad (2.1b)$$

entwickelt. Um mittels klassischer iterierter Defektkorrektur implizite Anfangswertprobleme erster Ordnung vom Typ (2.1) auf einem Gitter gemäß Definition 1.1 numerisch lösen zu können, wird der in Abschnitt 1.2 vorgestellte Algorithmus wie folgt modifiziert. Mit dem impliziten Euler-Verfahren als Basisverfahren erhält man an Stelle von (1.4a)

$$F\left(t_{j,l}, x_{j,l}^{[0]}, \frac{x_{j,l}^{[0]} - x_{j,l-1}^{[0]}}{h_{j,l}}\right) = 0, \quad j = 1, \dots, N, \quad l = 1, \dots, m \quad (2.2a)$$

$$x_{1,0}^{[0]} = u_0. \quad (2.2b)$$

Der Defekt (1.16) wird nun unter Beibehaltung sämtlicher anderer algorithmischer Details ebenfalls implizit definiert:

$$D_{j,l}^{[\nu]} := F\left(t_{j,l}, p_{j,l}^{[\nu]}, p_{j,l}^{[\nu]'}\right). \quad (2.3)$$

Die implizite Form des Nachbarproblems (1.15a) lautet

$$F\left(t_{j,l}, z_{j,l}^{[\nu]}, \frac{z_{j,l}^{[\nu]} - z_{j,l-1}^{[\nu]}}{h_{j,l}}\right) - D_{j,l}^{[\nu]} = 0, \quad (2.4a)$$

$$z_{1,0}^{[\nu]} = u_0. \quad (2.4b)$$

Bemerkung 2.1. Für linear implizite Differentialgleichungen gilt

$$F(t, u(t), u'(t)) = B(t) u'(t) + C(t) u(t) + g(t) = 0. \quad (2.5)$$

Ist $B(t)$ regulär, so sind bei Verwendung des *klassischen* Defekts sowohl das Defektkorrekturverfahren als auch das zugehörige Fixpunkt-kollokationsverfahren explizit und implizit angeschrieben zueinander äquivalent. Mit dem impliziten Euler-Verfahren als Basisverfahren¹ sowie für den Defekt gilt

$$B_{j,l} \frac{x_{j,l}^{[0]} - x_{j,l-1}^{[0]}}{h_{j,l}} + C_{j,l} x_{j,l} + g_{j,l} = 0 \Leftrightarrow \frac{x_{j,l}^{[0]} - x_{j,l-1}^{[0]}}{h_{j,l}} + B_{j,l}^{-1} C_{j,l} x_{j,l} + B_{j,l}^{-1} g_{j,l} = 0 \quad (2.6)$$

und

$$D_{j,l}^{[\nu]} = B_{j,l} p_{j,l}^{[\nu]'} + C_{j,l} p_{j,l}^{[\nu]} + g_{j,l} \Leftrightarrow d_{j,l}^{[\nu]} = p_{j,l}^{[\nu]'} + B_{j,l}^{-1} C_{j,l} p_{j,l}^{[\nu]} + B_{j,l}^{-1} g_{j,l}, \quad (2.7)$$

bzw.

$$d_{j,l}^{[\nu]} = B_{j,l}^{-1} D_{j,l}^{[\nu]}. \quad (2.8)$$

Analog gilt für die Kollokation

$$\begin{aligned} B_{j,l} p_{j,l}^{[C]'} + C_{j,l} p_{j,l}^{[C]} + g_{j,l} = 0 &\Leftrightarrow p_{j,l}^{[C]'} + B_{j,l}^{-1} C_{j,l} p_{j,l}^{[C]} + B_{j,l}^{-1} g_{j,l} = 0 \\ &\Leftrightarrow p_{j,l}^{[C]'} = -B_{j,l}^{-1} C_{j,l} p_{j,l}^{[C]} - B_{j,l}^{-1} g_{j,l}. \end{aligned} \quad (2.9)$$

Einen echten Mehrwert stellen die Verfahren mit klassischem Defekt somit nur bei nichtlinear impliziten Differentialgleichungen oder linear impliziten mit nicht regulärem $B(t)$ dar. Aufgrund von (2.9) kann man zumindest für den linear impliziten Fall ($B(t)$ regulär) mit den Resultaten von Kapitel 5 die Genauigkeit der Kollokationslösungen begründen.

Beispiel 2.2. Gegeben sei das skalare nichtlineare implizite Anfangswertproblem

$$e^{u'(t)} + u'(t) + u(t) - e^{-\sin(t)} - \cos(t) + \sin(t) = 0, \quad (2.10a)$$

$$u(t_0 = 0) = 1, \quad (2.10b)$$

mit der exakten Lösung $u^*(t) = \cos(t)$. Die IIDeC-IEUL-Lösung erfolgt auf einem äquidistanten Gitter gemäß Definition 1.1 mit Polynomgrad 4 im Intervall $[t_0, t_{end}] = [\mathbf{t}_0, \mathbf{t}_N] = [0, 3]$. Die globalen Fehler bezüglich der exakten Lösung $e^{*,[\#]}(t_{end}) = p^{[\#]}(t_{end}) - u^*(t_{end})$, $\# = \nu, C$ und die daraus berechneten Konvergenzordnungen an der Stelle $t_{end} = 3$ für das Basisverfahren, IIDeC und das zugehörige Fixpunkt-Kollokationsverfahren (=IEUL-Kollokation), sind in Tabelle 2.1 aufgelistet.

2.1.2 IIDeC-IEUL – Fixpunkt-konvergenz

Für IIDeC-IEUL gilt eine zu Satz 1.10 analoge Konvergenzaussage:

¹Die Lösung der Differenzgleichungen erfolgt in jedem Schritt mit dem Newton-Verfahren.

h	IIEUL	IIDeC1	IIDeC2	IIDeC3	IIDeC4	COLL
0,1	5,51E-03	8,15E-05	7,57E-07	4,31E-09	2,43E-08	2,49E-08
0,05	2,76E-03	2,04E-05	9,65E-08	2,72E-10	1,57E-09	1,59E-09
0,025	1,38E-03	5,10E-06	1,22E-08	1,70E-11	9,98E-11	1,00E-10
0,0125	6,90E-04	1,27E-06	1,53E-09	1,07E-12	6,29E-12	6,31E-12
0,1	1,00	2,00	2,97	3,99	3,95	3,97
0,05	1,00	2,00	2,99	3,99	3,98	3,98
0,025	1,00	2,00	2,99	4,00	3,99	3,99
0,0125	1,00	2,00	2,99	4,00	3,99	3,99

Tabelle 2.1: Resultate zu Beispiel 2.2.

Satz 2.3. (*Fixpunktkonvergenz von IIDeC*) Löst man das Anfangswertproblem vom Typ (2.1) auf einem äquidistanten Gitter (d.h. $h \equiv h_{j,l}$) gemäß Definition 1.1 mittels iterierter Defektkorrektur mit klassischem Defekt, so erfüllt der Fehler $e_{j,l}^{[\nu]}$ der Defektkorrektur-Lösungen $x_{j,l}^{[\nu]}$ der einzelnen Iterationen ν bezüglich des Fixpunktes $p_{j,l}^{[C]}$ (=Kollokationslösung auf demselben Gitter) die Abschätzung

$$\left\| e_{j,l}^{[\nu]} \right\| = \left\| x_{j,l}^{[\nu]} - p_{j,l}^{[C]} \right\| = \mathcal{O}(h^{\nu+1}), \quad \nu = 0, 1, \dots \quad (2.11)$$

Beweis von Satz 2.3 – Herleitung der Fehlergleichungen (lineare Analysis):

Die folgenden Ausführungen beschränken sich auf n -dimensionale linear implizite Anfangswertprobleme vom Typ

$$B(t) u'(u) + C(t) u(t) + g(t) = 0 \quad (2.12a)$$

$$u(t_0) = u_0, \quad (2.12b)$$

wobei $B(t)$ und $C(t)$ hinreichend glatte $n \times n$ -Matrizen und $g(t)$ ein hinreichend glatter n -dimensionaler Vektor sind. Wie bereits angemerkt, sind die Ausführungen (im Falle von $B(t)$ regulär) äquivalent zur explizit umgeschriebenen Variante (siehe Bemerkung 2.1). Der Beweis dient jedoch als (didaktische) Vorbereitung auf den Konvergenzbeweis von IIQDeC (siehe Abschnitt 2.2.2). Die numerische Behandlung bezieht sich direkt auf die implizite Gestalt (2.12) ohne Umformulierung auf eine explizite Differentialgleichung.

- $\nu = 0$: Subtrahiert man vom Basisverfahren (2.2) (gemäß (2.12), mit $B_{j,l} \equiv B(t_{j,l})$, $C_{j,l} \equiv C(t_{j,l})$ und $g_{j,l} \equiv g(t_{j,l})$)

$$B_{j,l} \frac{x_{j,l}^{[0]} - x_{j,l-1}^{[0]}}{h} + C_{j,l} x_{j,l}^{[0]} + g_{j,l} = 0, \quad (2.13)$$

den ebenfalls linearen Defekt (2.3) mit eingesetzter Kollokationslösung ($p_{j,l}^{[C]} \equiv p^{[C]}(t_{j,l})$)

$$B_{j,l} p_{j,l}^{[C]'} + C_{j,l} p_{j,l}^{[C]} + g_{j,l} = 0 \quad (2.14)$$

und die sich zu Null summierenden Terme

$$B_{j,l} \frac{p_{j,l}^{[C]} - p_{j,l-1}^{[C]}}{h} - B_{j,l} \frac{p_{j,l}^{[C]} - p_{j,l-1}^{[C]}}{h} = 0, \quad (2.15)$$

erhält man durch Zusammenfassen der Terme und Berücksichtigung von $e^{[\nu]}(t) = p^{[\nu]}(t) - p^{[C]}(t)$, wobei $p^{[\nu]}(t)$ die Interpolante der $x_{j,l}^{[\nu]}$ ist, die Fehlergleichung für $\nu = 0$

$$B_{j,l} \frac{e_{j,l}^{[0]} - e_{j,l-1}^{[0]}}{h} + C_{j,l} e_{j,l}^{[0]} + B_{j,l} \underbrace{\left(\frac{p_{j,l}^{[C]} - p_{j,l-1}^{[C]} - p_{j,l}^{[C]'}}{h} \right)}_{=: r_{j,l}^{I,[0]}}. \quad (2.16)$$

Für die Ordnung des Restgliedes $r_{j,l}^{I,[0]}$ gilt (Taylor-Reihenentwicklung)

$$\begin{aligned} r_{j,l}^{I,[0]} &= \frac{p_{j,l}^{[C]} - \left(p_{j,l}^{[C]} - h p_{j,l}^{[C]'} + \frac{h^2}{2} p_{j,l}^{[C]''} + \mathcal{O}(h^3) \right)}{h} - p_{j,l}^{[C]'} \\ &= -\frac{h}{2} p_{j,l-1}^{[C]''} + \mathcal{O}(h^2) = \mathcal{O}(h). \end{aligned} \quad (2.17)$$

Setzt man die Ordnung des Restgliedes in die Fehlergleichung (2.16) ein, erhält man zusammen mit einem Stabilitätsargument und dem Lemma von Gronwall (siehe z.B. [4]) die Ordnung auch global.

- $\nu - 1 \rightarrow \nu$: Ausgehend von der Verbesserung der Lösung (1.17)

$$x_{j,l}^{[\nu]} - x_{j,l}^{[0]} + z_{j,l}^{[\nu-1]} - x_{j,l}^{[\nu-1]} = 0 \quad (2.18)$$

gelangt man zu

$$\underbrace{x_{j,l}^{[\nu]} - p_{j,l}^{[C]} - x_{j,l}^{[0]}}_{=: e_{j,l}^{[\nu]}} + z_{j,l}^{[\nu-1]} - \underbrace{x_{j,l}^{[\nu-1]} + p_{j,l}^{[C]}}_{=: e_{j,l}^{[\nu-1]}} = 0. \quad (2.19)$$

Anwendung des Differenzenoperators L und Multiplikation mit $B_{j,l}$ ergibt

$$B_{j,l} \frac{e_{j,l}^{[\nu]} - e_{j,l-1}^{[\nu]}}{h} - B_{j,l} \frac{x_{j,l}^{[0]} - x_{j,l-1}^{[0]}}{h} + B_{j,l} \frac{z_{j,l}^{[\nu-1]} - z_{j,l-1}^{[\nu-1]}}{h} - B_{j,l} \frac{e_{j,l}^{[\nu-1]} - e_{j,l-1}^{[\nu-1]}}{h} = 0. \quad (2.20)$$

Setzt man die entsprechenden linearen Identitäten (2.2), (2.3) und (2.4) in (2.20) ein, folgt

$$\begin{aligned} & B_{j,l} \frac{e_{j,l}^{[\nu]} - e_{j,l-1}^{[\nu]}}{h} + C_{j,l} x_{j,l}^{[0]} + g_{j,l} \\ & - C_{j,l} z_{j,l}^{[\nu-1]} - g_{j,l} + B_{j,l} p_{j,l}^{[\nu-1]'} + C_{j,l} p_{j,l}^{[\nu-1]} + g_{j,l} - B_{j,l} \frac{e_{j,l}^{[\nu-1]} - e_{j,l-1}^{[\nu-1]}}{h} = 0 \end{aligned} \quad (2.21)$$

Subtrahiert man von (2.21) die Identität $B_{j,l} p_{j,l}^{[C]'} + C_{j,l} p_{j,l}^{[C]} + g_{j,l} = 0$ und fasst man die Terme zusammen, erhält man die Fehlergleichung

$$\begin{aligned} & B_{j,l} \frac{e_{j,l}^{[\nu]} - e_{j,l-1}^{[\nu]}}{h} + C_{j,l} \underbrace{\left(x_{j,l}^{[0]} - \left(z_{j,l}^{[\nu-1]} - x_{j,l}^{[\nu-1]} \right) \right)}_{=: e_{j,l}^{[\nu]}} \\ & + B_{j,l} \underbrace{\left(e_{j,l}^{[\nu-1]'} - \frac{e_{j,l}^{[\nu-1]} - e_{j,l-1}^{[\nu-1]}}{h} \right)}_{=: r_{j,l}^{I,[\nu]}} = 0. \end{aligned} \quad (2.22)$$

h	ITR	IIDeC1	IIDeC2	ITR-COLL
0,1	6,44E-05	1,12E-08	8,82E-09	8,82E-09
0,05	1,61E-05	6,99E-10	5,51E-10	5,51E-10
0,025	4,03E-06	4,37E-11	3,44E-11	3,44E-11
0,0125	1,01E-06	2,73E-12	2,15E-12	2,15E-12
0,1	2,00	4,00	4,00	4,00
0,05	2,00	4,00	4,00	4,00
0,025	2,00	4,00	4,00	4,00
0,0125	2,00	4,00	4,00	4,00

Tabelle 2.2: Resultate zu Beispiel 2.4.

Kompakt geschrieben erhält man

$$B_{j,l} \frac{e_{j,l}^{[\nu]} - e_{j,l-1}^{[\nu]}}{h} + C_{j,l} e_{j,l}^{[\nu]} + B_{j,l} r_{j,l}^{I,[\nu]} = 0, \quad (2.23)$$

wobei das Restglied $r_{j,l}^{I,[\nu]}$ die gleiche Bauform wie in (1.42) besitzt². Die Ordnungen des Restgliedes wurden (allerdings für das explizite Euler-Verfahren) in Abschnitt 1.2 ausführlich bewiesen. Für den vollständigen Beweis von Satz 2.3 ist analog vorzugehen.

2.1.3 IIDeC-ITR

Wendet man die implizite Trapezregel anstatt (2.2) als Basisverfahren zur Diskretisierung des impliziten Anfangswertproblems (2.1) an, erhält man

$$\frac{1}{2} \left[F \left(t_{j,l-1}, x_{j,l-1}^{[0]}, \frac{x_{j,l}^{[0]} - x_{j,l-1}^{[0]}}{h_{j,l}} \right) + F \left(t_{j,l}, x_{j,l}^{[0]}, \frac{x_{j,l}^{[0]} - x_{j,l-1}^{[0]}}{h_{j,l}} \right) \right] = 0. \quad (2.24)$$

Mit dem klassischen Defekt für implizite Differentialgleichungen (2.3) gelangt man zum ITR-Nachbarproblem

$$\frac{1}{2} \left[F \left(t_{j,l-1}, z_{j,l-1}^{[\nu]}, \frac{z_{j,l}^{[\nu]} - z_{j,l-1}^{[\nu]}}{h_{j,l}} \right) + F \left(t_{j,l}, z_{j,l}^{[\nu]}, \frac{z_{j,l}^{[\nu]} - z_{j,l-1}^{[\nu]}}{h_{j,l}} \right) - D_{j,l-1}^{[\nu]} - D_{j,l}^{[\nu]} \right] = 0. \quad (2.25)$$

Alle anderen algorithmischen Details bleiben unberührt.

Beispiel 2.4. Wie Beispiel 2.2, jedoch mit der impliziten Trapezregel als Basisverfahren. Tabelle 2.2 zeigt die Resultate.

²Lediglich die Ableitung wird aufgrund des impliziten Basisverfahrens an einer anderen Stelle ausgewertet.

h	IEUL	IIQDeC1	IIQDeC2	IIQDeC3	IIQDeC4	COLL
0,1	6,31E-03	1,14E-04	1,02E-06	3,83E-09	3,95E-09	3,98E-09
0,05	3,16E-03	2,90E-05	1,31E-07	2,69E-10	2,42E-10	2,43E-10
0,025	1,58E-03	7,30E-05	1,66E-08	1,77E-11	1,49E-11	1,50E-11
0,0125	7,91E-04	1,83E-06	2,09E-09	1,14E-12	9,29E-12	9,31E-12
0,1	1,00	1,98	2,96	3,83	4,03	4,04
0,05	1,00	1,99	2,98	3,92	4,01	4,02
0,025	1,00	1,99	2,99	3,96	4,00	4,01
0,0125	1,00	1,99	2,99	3,96	4,00	4,01

Tabelle 2.3: Resultate zu Beispiel 2.6.

2.2 Integralmitteldefekt – IIQDeC

2.2.1 IIQDeC-IEUL

Zur Lösung von (2.1) mittels Integralmitteldefekt wird (2.3) aus Abschnitt 2.1 durch

$$\begin{aligned} \bar{D}_{j,l}^{[\nu]} &:= \sum_{\mu=1}^m \alpha_{l,\mu} D^{[\nu]}(t_{j,\mu}) = \sum_{\mu=1}^m \alpha_{l,\mu} F\left(t_{j,\mu}, p^{[\nu]}(t_{j,\mu}), p^{[\nu]'}(t_{j,\mu})\right) \\ &\approx \frac{1}{h_{j,l}} \int_{t_{j,l-1}}^{t_{j,l}=t_{j,l-1}+h_{j,l}} F\left(s, p^{[\nu]}(s), p^{[\nu]'}(s)\right) ds \end{aligned} \quad (2.26)$$

ersetzt, der Rest des beschriebenen Algorithmus bleibt unverändert.

Bemerkung 2.5. Wie folgende Ausführungen zeigen, sind bei linear impliziten Differentialgleichungen (2.12) mit regulärem $B(t)$ durch die Bildung des Integralmittels die explizite und implizite Variante der Defektkorrektur zueinander nicht mehr äquivalent: Integriert man (2.8) numerisch, so erhält man

$$\bar{d}_{j,l}^{[\nu]} = \sum_{\mu=1}^m \alpha_{l,\mu} B_{j,\mu}^{-1} D_{j,\mu}^{[\nu]} \neq B_{j,l}^{-1} \bar{D}_{j,l}^{[\nu]}. \quad (2.27)$$

Beispiel 2.6. Gegeben sei das skalare nichtlineare Anfangswertproblem (2.10) mit der exakten Lösung $u^*(t) = \cos(t)$. Die IIQDeC-IEUL-Lösung erfolgt auf einem *nicht* äquidistanten Gitter mit 4 Kollokationsknoten gemäß Definition 1.1 ($c_1 = 0,1234$, $c_2 = 0,5054$, $c_3 = 0,7134$, $c_4 = 1$) im Intervall $[t_0, t_{end}] = [\mathbf{t}_0, \mathbf{t}_N] = [0, 3]$. Die Ergebnisse (globale Fehler bezüglich der exakten Lösung $e^{*,[\#]}(t_{end}) = p^{[\#]}(t_{end}) - u^*(t_{end})$, $\# = \nu, C$ und zugehörige Konvergenzordnungen) an der Stelle $t_{end} = 3$ für das Basisverfahren, IIQDeC und das zugehörige Fixpunkt-Kollokationsverfahren (=IEUL-Kollokation), sind in Tabelle 2.3 aufgelistet.

2.2.2 IIQDeC-IEUL – Fixpunkt-Konvergenz

Satz 2.7. (*Fixpunkt-Konvergenz von IIQDeC-IEUL*) Löst man das Anfangswertproblem vom Typ (2.1) auf einem beliebigen Gitter gemäß Definition 1.1 mittels iterierter Defektkorrektur

mit Integralmittelfekt, so erfüllt der Fehler $e_{j,l}^{[\nu]}$ der Lösungen $x_{j,l}^{[\nu]}$ der einzelnen Iterationen bezüglich des Fixpunktes $p_{j,l}^{[C]}$ (=Kollokationslösung auf demselben Gitter) folgende Bedingung

$$\left\| e_{j,l}^{[\nu]} \right\| = \left\| x_{j,l}^{[\nu]} - p_{j,l}^{[C]} \right\| = \mathcal{O}(\mathbf{h}^{\nu+1}), \quad \nu = 0, 1, \dots \quad (2.28)$$

Beweis von Satz 2.7 (lineare Analysis):

- $\nu = 0$: Subtrahiert man vom Basisverfahren (2.2) (gemäß (2.12))

$$B_{j,l} \frac{x_{j,l}^{[0]} - x_{j,l-1}^{[0]}}{h_{j,l}} + C_{j,l} x_{j,l}^{[0]} - g_{j,l} = 0 \quad (2.29)$$

den ebenfalls (linearen) Defekt (2.26)

$$\sum_{\mu=1}^m \alpha_{l,\mu} \left(B_{j,\mu} p_{j,\mu}^{[C]'} + C_{j,\mu} p_{j,\mu}^{[C]} + g_{j,\mu} \right) = 0, \quad (2.30)$$

sowie die sich zu Null summierenden Terme

$$B_{j,l} \frac{p_{j,l}^{[C]} - p_{j,l-1}^{[C]}}{h_{j,l}} - B_{j,l} \frac{p_{j,l}^{[C]} - p_{j,l-1}^{[C]}}{h_{j,l}} + C_{j,l} p_{j,l}^{[C]} - C_{j,l} p_{j,l}^{[C]} = 0, \quad (2.31)$$

erhält man unter Ausnutzung von $e_{j,l}^{[0]} = p_{j,l}^{[0]} - p_{j,l}^{[C]}$ ($p^{[\nu]}(t)$ interpoliert stückweise die $x_{j,l}^{[\nu]}$, $\nu = 0, 1, \dots$) und $\sum_{\mu=1}^m \alpha_{l,\mu} = 1$ sowie durch Zusammenfassen der Terme die Fehlergleichung für $\nu = 0$,

$$\begin{aligned} & B_{j,l} \frac{e_{j,l}^{[0]} - e_{j,l-1}^{[0]}}{h_{j,l}} + C_{j,l} e_{j,l}^{[0]} \\ + & \underbrace{B_{j,l} \frac{p_{j,l}^{[C]} - p_{j,l-1}^{[C]}}{h_{j,l}} - \sum_{\mu=1}^m \alpha_{l,\mu} B_{j,\mu} p_{j,\mu}^{[C]'}}_{=\bar{r}_{j,l}^{[0]}} + \underbrace{C_{j,l} p_{j,l}^{[C]} - \sum_{\mu=1}^m \alpha_{l,\mu} C_{j,\mu} p_{j,\mu}^{[C]}}_{=\bar{r}_{j,l}^{[0]}} = 0. \end{aligned} \quad (2.32)$$

Für die Ordnungen der Restglieder gilt mittels Taylor-Reihenentwicklung und unter Ausnutzung, dass die Kollokationspolynome vom Grad m numerisch exakt integriert werden,

$$\begin{aligned} \bar{r}_{j,l}^{[0]} &= B_{j,l} \frac{p_{j,l}^{[C]} - p_{j,l-1}^{[C]}}{h_{j,l}} - \sum_{\mu=1}^m \alpha_{l,\mu} \left(B_{j,\mu} + (t_{j,\mu} - t_{j,l}) B'_{j,\mu} + \mathcal{O}(\mathbf{h}^2) \right) p_{j,\mu}^{[C]'} \\ &= B_{j,l} \left(\underbrace{\frac{p_{j,l}^{[C]} - p_{j,l-1}^{[C]}}{h_{j,l}} - \sum_{\mu=1}^m \alpha_{l,\mu} p_{j,\mu}^{[C]'}}_{=0} \right) - \underbrace{\sum_{\mu=1}^m \alpha_{l,\mu} (t_{j,\mu} - t_{j,l}) B_{j,\mu} p_{j,\mu}^{[C]'}}_{=\mathcal{O}(\mathbf{h})} + \mathcal{O}(\mathbf{h}^2) \\ &= \mathcal{O}(\mathbf{h}), \end{aligned} \quad (2.33)$$

bzw. wie schon in Anhang A von [2] beschrieben (Taylor-Reihenentwicklung und Anwendung von $\sum_{\mu=1}^m \alpha_{l,\mu} = 1$)

$$\begin{aligned}
\hat{r}_{j,l}^{[0]} &= C_{j,l} p_{j,l}^{[C]} - \sum_{\mu=1}^m \alpha_{l,\mu} \left(C_{j,l} p_{j,l}^{[C]} + (t_{j,\mu} - t_{j,l}) \left(C_{j,\mu} p_{j,\mu}^{[C]} \right)' + \mathcal{O}(\mathbf{h}^2) \right) \\
&= C_{j,l} p_{j,l}^{[C]} - \underbrace{\sum_{\mu=1}^m \alpha_{l,\mu} C_{j,l} p_{j,l}^{[C]}}_{=C_{j,l} p_{j,l}^{[C]}} - \underbrace{\sum_{\mu=1}^m \alpha_{l,\mu} (t_{j,\mu} - t_{j,l}) \left(C_{j,\mu} p_{j,\mu}^{[C]} \right)'}_{=\mathcal{O}(\mathbf{h})} + \mathcal{O}(\mathbf{h}^2) \\
&= \mathcal{O}(\mathbf{h}). \tag{2.34}
\end{aligned}$$

Setzt man die Ordnungen der Restglieder in die Fehlergleichung (2.32) ein, erhält man zusammen mit einem Stabilitätsargument und dem Lemma von Gronwall (siehe z.B. [4]) die Ordnung auch global. Somit ist Satz 2.7 für $\nu = 0$ bewiesen.

- $\nu - 1 \rightarrow \nu$: Betrachtet man das linear implizite Anfangswertproblem (2.12), so gelangt man, ausgehend von (2.20), unter Ausnutzung der linearen Form von (2.2), (2.4) und (2.26), zu (vgl. (2.21), jedoch mit Integralmittel- statt punktwissem Defekt)

$$\begin{aligned}
&B_{j,l} \frac{e_{j,l}^{[\nu]} - e_{j,l-1}^{[\nu]}}{h_{j,l}} + C_{j,l} x_{j,l}^{[0]} + g_{j,l} - C_{j,l} z_{j,l}^{[\nu-1]} - g_{j,l} \\
&+ \sum_{\mu=1}^m \alpha_{l,\mu} \left(B_{j,\mu} p_{j,\mu}^{[\nu-1]'} + C_{j,\mu} p_{j,\mu}^{[\nu-1]} + g_{j,\mu} \right) - B_{j,l} \frac{e_{j,l}^{[\nu-1]} - e_{j,l-1}^{[\nu-1]}}{h_{j,l}} = 0. \tag{2.35}
\end{aligned}$$

Auflösen der Klammer, Subtraktion von (2.30) und der sich zu Null summierenden Terme

$$\underbrace{C_{j,l} x_{j,l}^{[\nu-1]} - C_{j,l} p_{j,l}^{[C]}}_{=C_{j,l} e_{j,l}^{[\nu-1]}} - C_{j,l} x_{j,l}^{[\nu-1]} + C_{j,l} p_{j,l}^{[C]} = 0, \tag{2.36}$$

Ausnutzung von $\sum_{\mu=1}^m \alpha_{l,\mu} = 1$ sowie Zusammenfassen der Terme ergibt (vgl. (2.22))

$$\begin{aligned}
&B_{j,l} \frac{e_{j,l}^{[\nu]} - e_{j,l-1}^{[\nu]}}{h_{j,l}} + C_{j,l} \left(\underbrace{x_{j,l}^{[0]} - \left(z_{j,l}^{[\nu-1]} - x_{j,l}^{[\nu-1]} \right)}_{=e_{j,l}^{[\nu]}} - p_{j,l}^{[C]} \right) \\
&+ \underbrace{\sum_{\mu=1}^m \alpha_{l,\mu} B_{j,\mu} e_{j,\mu}^{[\nu-1]'}}_{=\bar{r}_{j,l}^{[\nu]}} - B_{j,l} \frac{e_{j,l}^{[\nu-1]} - e_{j,l-1}^{[\nu-1]}}{h_{j,l}} + \underbrace{\sum_{\mu=1}^m \alpha_{l,\mu} C_{j,\mu} e_{j,\mu}^{[\nu-1]} - C_{j,l} e_{j,l}^{[\nu-1]}}_{=\hat{r}_{j,l}^{[\nu]}} = 0. \tag{2.37}
\end{aligned}$$

Kompakt angeschrieben erhält man die allgemeine Fehlergleichung für IIQDeC-IEUL

$$B_{j,l} \frac{e_{j,l}^{[\nu]} - e_{j,l-1}^{[\nu]}}{h_{j,l}} + C_{j,l} e_{j,l}^{[\nu]} + \bar{r}_{j,l}^{[\nu]} + \hat{r}_{j,l}^{[\nu]} = 0. \tag{2.38}$$

Die Berechnung der Ordnung für $\hat{r}_{j,l}^{[\nu]}$ ist für $\nu = 1$ in Anhang A von [2] beschrieben.

Angenommen Satz 2.7 wäre für $\nu - 1$ bewiesen, d.h.

$$\left\| e_{j,l}^{[\nu-1]} \right\| = \mathcal{O}(\mathbf{h}^\nu) \quad \text{und} \quad \left\| \frac{e_{j,l}^{[\nu-1]} - e_{j,l-1}^{[\nu-1]}}{h_{j,l}} \right\| = \mathcal{O}(\mathbf{h}^\nu) \quad (2.39)$$

gelten, so können die Ordnungen der Restglieder $\bar{r}_{j,l}^{[\nu]}$ und $\hat{r}_{j,l}^{[\nu]}$ aus (2.38) mittels Taylor-Reihenentwicklung wie folgt bestimmt werden

$$\begin{aligned} \bar{r}_{j,l}^{[\nu]} &= \sum_{\mu=1}^m \alpha_{l,\mu} \left(B_{j,l} + (t_{j,\mu} - t_{j,l}) B'_{j,\mu} + \mathcal{O}(\mathbf{h}^2) \right) e_{j,\mu}^{[\nu-1]'} - B_{j,l} \frac{e_{j,l}^{[\nu-1]} - e_{j,l-1}^{[\nu-1]}}{h_{j,l}} \\ &= \underbrace{\sum_{\mu=1}^m \alpha_{l,\mu} (t_{j,\mu} - t_{j,l}) B_{j,\mu} e_{j,\mu}^{[\nu-1]'}}_{=\mathcal{O}(\mathbf{h}^{\nu+1})} + \mathcal{O}(\mathbf{h}^{\nu+2}) \\ &\quad - B_{j,l} \underbrace{\left(\sum_{\mu=1}^m \alpha_{l,\mu} e_{j,\mu}^{[\nu-1]'} - \frac{e_{j,l}^{[\nu-1]} - e_{j,l-1}^{[\nu-1]}}{h_{j,l}} \right)}_{=0} \\ &= \mathcal{O}(\mathbf{h}^{\nu+1}), \end{aligned} \quad (2.40)$$

bzw.

$$\begin{aligned} \hat{r}_{j,l}^{[\nu]} &= \sum_{\mu=1}^m \alpha_{l,\mu} \left(C_{j,l} e_{j,l}^{[\nu-1]} + (t_{j,\mu} - t_{j,l}) \left(C_{j,\mu} e_{j,\mu}^{[\nu-1]} \right)' + \mathcal{O}(\mathbf{h}^{\nu+2}) \right) - C_{j,l} e_{j,l}^{[\nu-1]} \\ &= \underbrace{\sum_{\mu=1}^m \alpha_{l,\mu} (t_{j,\mu} - t_{j,l}) \left(C_{j,\mu} e_{j,\mu}^{[\nu-1]} \right)'}_{=\mathcal{O}(\mathbf{h}^{\nu+1})} + \mathcal{O}(\mathbf{h}^{\nu+2}) + \underbrace{\sum_{\mu=1}^m \alpha_{l,\mu} C_{j,l} e_{j,l}^{[\nu-1]} - C_{j,l} e_{j,l}^{[\nu-1]}}_{=C_{j,l} e_{j,l}^{[\nu-1]} - C_{j,l} e_{j,l}^{[\nu-1]} = 0} \\ &= \mathcal{O}(\mathbf{h}^{\nu+1}). \end{aligned} \quad (2.41)$$

Zu zeigen bleibt, dass für die ersten Ableitungen des Fehlerpolynoms $e_{j,l}^{[\nu-1]'} = \mathcal{O}(\mathbf{h}^\nu)$ gilt, wenn für das Fehlerpolynom selbst $e_{j,l}^{[\nu-1]} = \mathcal{O}(\mathbf{h}^\nu)$ gilt. Die Vererbung der Ordnung des Fehlerpolynoms auf die erste Ableitung wird in Abschnitt 5.5.5 gezeigt.

Setzt man die Ordnungen der Restglieder (2.40) und (2.41) in die Fehlergleichung (2.38) ein, erhält man zusammen mit einem Stabilitätsargument und dem Lemma von Gronwall (siehe z.B. [4]) die Ordnung auch global und Satz 2.7 ist bewiesen.

Bemerkung 2.8. Ersetzt man im impliziten Anfangswertproblem (2.12) die Matrix $B(t)$ durch die Einheitsmatrix I , so vereinfacht sich das Restglied (2.40) aus Fehlergleichung (2.38) wegen der Exaktheit der angewendeten Quadraturformel zu

$$\bar{r}_{j,l}^{[\nu]} = \sum_{\mu=1}^m \alpha_{l,\mu} e_{j,\mu}^{[\nu-1]'} - \frac{e_{j,l}^{[\nu-1]} - e_{j,l-1}^{[\nu-1]}}{h_{j,l}} = 0, \quad (2.42)$$

h	ITR	IIQDeC1	IIQDeC2	IIQDeC3	COLL
0,1	9,29E-05	6,31E-08	2,05E-09	4,02E-09	3,99E-09
0,05	2,32E-05	6,81E-09	1,28E-10	2,44E-10	2,43E-10
0,025	5,81E-06	7,84E-10	8,04E-12	1,50E-11	1,50E-11
0,0125	1,45E-06	9,39E-11	5,03E-13	9,32E-13	9,31E-13
0,1	2,00	3,21	3,99	4,04	4,04
0,05	2,00	3,12	4,00	4,02	4,02
0,025	2,00	3,06	4,00	4,01	4,01
0,0125	2,00	3,06	4,00	4,01	4,01

Tabelle 2.4: Resultate zu Beispiel 2.9.

h	ITR	IIQDeC1	IIQDeC2	COLL
0,1	6,44E-05	2,50E-08	2,49E-08	2,49E-08
0,05	1,61E-05	1,58E-09	1,59E-09	1,59E-09
0,025	4,03E-06	9,96E-11	1,00E-10	1,00E-10
0,0125	1,01E-06	6,25E-12	6,31E-12	6,31E-12
0,1	2,00	3,98	3,97	3,97
0,05	2,00	3,99	3,98	3,98
0,025	2,00	4,00	3,99	3,99
0,0125	2,00	4,00	3,99	3,99

Tabelle 2.5: Resultate zu Beispiel 2.10.

während das Restglied (2.41) unverändert bleibt. Durch den Beweis von Satz 2.7 ist somit auch Satz 1.38 bewiesen.

2.2.3 IIQDeC-ITR

Kombiniert man das Basisverfahren (2.24) mit dem Defekt (2.26) und dem Nachbarproblem

$$\frac{1}{2} \left[F \left(t_{j,l-1}, z_{j,l-1}^{[\nu]}, \frac{z_{j,l}^{[\nu]} - z_{j,l-1}^{[\nu]}}{h_{j,l}} \right) + F \left(t_{j,l}, z_{j,l}^{[\nu]}, \frac{z_{j,l}^{[\nu]} - z_{j,l-1}^{[\nu]}}{h_{j,l}} \right) \right] - \bar{D}_{j,l}^{[\nu]} = 0, \quad (2.43)$$

erhält man IIQDeC-ITR.

Beispiel 2.9. Wie Beispiel 2.6, jedoch mit der impliziten Trapezregel als Basisverfahren. Tabelle 2.4 zeigt die Resultate.

Beispiel 2.10. Wie Beispiel 1.48, jedoch mit der impliziten Trapezregel als Basisverfahren auf einem äquidistanten Gitter. Tabelle 2.5 zeigt die Resultate.

h	IEUL	IIPDeC1	IIPDeC2	IIPDeC3	IIPDeC4	IIPDeC5	G-COLL
0,1	7,34E-03	1,45E-04	1,84E-06	3,27E-08	1,49E-09	5,64E-11	1,90E-12
0,05	3,68E-03	3,62E-05	2,32E-07	2,10E-09	4,86E-11	9,19E-13	2,97E-14
0,025	1,84E-03	9,06E-06	2,91E-08	1,33E-10	1,55E-12	1,47E-14	4,64E-16
0,0125	9,20E-04	2,26E-06	3,64E-09	8,37E-12	4,89E-14	2,31E-16	7,26E-18
0,1	1,00	2,01	2,99	3,96	4,94	5,94	6,00
0,05	1,00	2,00	3,00	3,98	4,97	5,97	6,00
0,025	1,00	2,00	3,00	3,99	4,99	5,99	6,00
0,0125	1,00	2,00	3,00	3,99	4,99	5,99	6,00

Tabelle 2.6: Resultate zu Beispiel 2.11.

2.3 Interpolierter Defekt – IIPDeC

2.3.1 IIPDeC-IEUL

Der IIPDeC-Algorithmus ist das implizite Analogon zum IPDeC-Verfahren³. Man kann damit implizite Anfangswertprobleme vom Typ (2.1) auf einem Gitter gemäß Abbildung 1.4 lösen. Dazu muss der Defekt (2.3) in Beziehung (2.4) durch die implizite Variante von (1.146) ersetzt werden. D.h., dass der Defekt wie im expliziten Fall an den Knoten $\tilde{t}_{j,\kappa}$ des Gauß-Gitters ausgewertet wird

$$\tilde{D}^{[\nu]}(\tilde{t}_{j,\kappa}) := F\left(\tilde{t}_{j,\kappa}, p^{[\nu]}(\tilde{t}_{j,\kappa}), p^{[\nu]'}(\tilde{t}_{j,\kappa})\right), \quad (2.44)$$

wobei die stückweise definierten Polynome $p^{[\nu]}(t)$ die Werte $x_{j,l}^{[\nu]}$ interpolieren. Die Defektwerte $\tilde{D}^{[\nu]}(\tilde{t}_{j,\kappa})$ werden anschließend mittels der Polynome $\tilde{p}^{[\nu]}(t)$ vom Grad $\leq m - 1$ interpoliert. Wertet man diese Polynome an den Knoten $t_{j,l}$ aus, gelangt man zum IIPDeC-Defekt

$$\hat{D}_{j,l}^{[\nu]} := \hat{D}^{[\nu]}(t_{j,l}) = \tilde{p}^{[\nu]}(t_{j,l}), \quad j = 1, \dots, m, \quad (2.45)$$

der den Defekt (2.3) in (2.4) ersetzt. Der Rest des Algorithmus bleibt unberührt, es gilt eine zu Satz 1.56 analoge Konvergenzaussage, wie folgendes Beispiel zeigt.

Beispiel 2.11. Gegeben sei das skalare nichtlineare implizite Anfangswertproblem (2.10) mit der exakten Lösung $u^*(t) = \cos(t)$. Die IIPDeC-IEUL-Lösung erfolgt auf einem äquidistanten Gitter mit 3 Kollokationsknoten und einem nicht äquidistanten Gitter mit 3 Gauß-Knoten gemäß Abbildung 1.4 ($c_1 = \frac{1}{3}$, $c_2 = \frac{2}{3}$, $c_3 = 1$ und $\tilde{c}_1 = \frac{1}{2} - \frac{\sqrt{15}}{10}$, $\tilde{c}_2 = \frac{1}{2}$, $\tilde{c}_3 = \frac{1}{2} + \frac{\sqrt{15}}{10}$ für $\tilde{t}_{j,l} = \mathbf{t}_{j-1} + \tilde{c}_l \mathbf{h}$) im Intervall $[t_0, t_{end}] = [\mathbf{t}_0, \mathbf{t}_N] = [0, 3]$. Die globalen Fehler bezüglich der exakten Lösung $e^{*,[\#]}(t_{end}) = p^{[\#]}(t_{end}) - u^*(t_{end})$, $\# = \nu, C$ und die daraus berechneten Konvergenzordnungen an der Stelle $t_{end} = 3$ für das Basisverfahren, IIPDeC und das zugehörige Fixpunkt-Kollokationsverfahren (=Gauß-Kollokation), sind in Tabelle 2.6 aufgelistet.

³Die Notation ist ident mit jener aus Abschnitt 1.4.

h	ITR	IIPDeC1	IIPDeC2	IIPDeC3	G-COLL
0,1	7,03E-06	6,83E-10	7,66E-13	6,47E-13	6,47E-13
0,05	1,76E-06	4,26E-11	1,20E-14	1,01E-14	1,01E-14
0,025	4,39E-07	2,66E-12	1,87E-16	1,58E-16	1,58E-16
0,0125	1,10E-07	1,67E-13	2,92E-18	2,47E-17	2,47E-17
0,1	2,00	4,00	6,00	6,00	6,00
0,05	2,00	4,00	6,00	6,00	6,00
0,025	2,00	4,00	6,00	6,00	6,00
0,0125	2,00	4,00	6,00	6,00	6,00

Tabelle 2.7: Resultate zu Beispiel 2.12.

2.3.2 IIPDeC-ITR

Ausgehend vom Basisverfahren (2.24) und dem Defekt (2.45) mit dem Nachbarproblem

$$\frac{1}{2} \left[F \left(t_{j,l-1}, z_{j,l-1}^{[\nu]}, \frac{z_{j,l}^{[\nu]} - z_{j,l-1}^{[\nu]}}{h_{j,l}} \right) + F \left(t_{j,l}, z_{j,l}^{[\nu]}, \frac{z_{j,l}^{[\nu]} - z_{j,l-1}^{[\nu]}}{h_{j,l}} \right) - \hat{D}_{j,l-1}^{[\nu]} - \hat{D}_{j,l}^{[\nu]} \right] = 0, \quad (2.46)$$

gelangt man zu einem superkonvergenten Verfahren mit der Ordnungsfolge 2,4,...

Beispiel 2.12. Wie Beispiel 2.11, jedoch mit der impliziten Trapezregel als Basisverfahren. Tabelle 2.7 zeigt die Resultate.

Bemerkung 2.13. Zusammenfassend wird festgehalten, dass alle Defektkorrekturvarianten für implizite Differentialgleichungen erster Ordnung (II*DeC) das gleiche Konvergenzverhalten zeigen wie jene für explizite Differentialgleichungen (I*DeC).

Kapitel 3

Explizite Differentialgleichungen zweiter Ordnung

In diesem Kapitel werden die in Kapitel 1 vorgestellten Verfahren (IDeC, IQDeC und IPDeC) so erweitert, dass explizite Anfangswertprobleme 2. Ordnung ohne und mit Rückführung auf ein System erster Ordnung mittels iterierter Defektkorrektur gelöst werden können.

3.1 Klassischer Defekt – IDeC2-DQ2

Dieses Verfahren wurde im Kontext von Randwertproblemen in [12] für Probleme vom Typ $u''(t) = f(t, u(t))$ ausführlich beschrieben. Da es als Ausgangspunkt für die vorgestellten darauf aufbauenden Verfahren dient, wird es an dieser Stelle kurz zusammengefasst wiederholt und auf Anfangswertprobleme vom Typ $u''(t) = f(t, u(t), u'(t))$ erweitert. Bei der Anwendung auf Anfangswertprobleme ergibt sich im Gegensatz zu Randwertproblemen eine zusätzliche Fragestellung: Wie sieht die Ausführung des Startschritts mit Defektdefinition inkl. Berücksichtigung im Nachbarproblem aus? Eine Antwort darauf liefern die folgenden Abschnitte.

3.1.1 IDeC2-DQ2

Man betrachtet ein n -dimensionales explizites Anfangswertproblem 2. Ordnung

$$u''(t) = f(t, u(t), u'(t)), \quad (3.1a)$$

$$u(t_0) = u_0 \quad \text{und} \quad u'(t_0) = v_0, \quad (3.1b)$$

wobei u_0 und v_0 die Startwerte des Anfangswertproblems bezeichnen. Legt man der numerischen Lösung von (3.1) ein Gitter gemäß Definition 1.1 zugrunde, so bietet sich folgende Diskretisierung als Basisverfahren an ($h \equiv h_{j,l}$)

$$\begin{cases} x_{1,1}^{[0]} = u_0 + h v_0 + \frac{h^2}{2} f(t_{1,0}, u_0, v_0), & j = 1, l = 1 \\ \frac{x_{j,l+1}^{[0]} - 2x_{j,l}^{[0]} + x_{j,l-1}^{[0]}}{h^2} = f\left(t_{j,l}, x_{j,l}^{[0]}, \frac{x_{j,l+1}^{[0]} - x_{j,l-1}^{[0]}}{2h}\right), & j \geq 1, l > 1 \\ \frac{x_{j,1}^{[0]} - 2x_{j-1,m}^{[0]} + x_{j-1,m-1}^{[0]}}{h^2} = f\left(t_{j-1,m}, x_{j-1,m}^{[0]}, \frac{x_{j,1}^{[0]} - x_{j-1,m-1}^{[0]}}{2h}\right), & j > 1, l = 1. \end{cases} \quad (3.2)$$

Die Lösungen der impliziten Differenzengleichungen liefern eine Folge von $x_{j,l}^{[0]}$, $j = 1, \dots, N$, $l = 0, \dots, m$ mit $x_{j,0}^{[0]} = x_{j-1,m}^{[0]}$ für $j > 1$. Interpoliert man diese Lösungen bzw. in weiterer Folge die verbesserten Lösungen $x_{j,l}^{[\nu]}$, $\nu = 0, 1, \dots$ stückweise durch die Polynome $p^{[\nu]}(t)$ ($p_{j,l}^{[\nu]} \equiv p^{[\nu]}(t_{j,l})$) vom Grad $\leq m - 1$, so berechnen sich die Defekte wie folgt

$$\begin{cases} d_{1,0}^{A,[\nu]} = p_{1,0}^{[\nu]''} - f(t_{1,0}, u_0, v_0), & d_{1,0}^{B,[\nu]} = p_{1,0}^{[\nu]'} - v_0, & j = 1, l = 0 \\ d_{j,l}^{[\nu]} = p_{j,l}^{[\nu]''} - f(t_{j,l}, p_{j,l}^{[\nu]}, p_{j,l}^{[\nu]'}), & & j \geq 1, l > 0 \\ d_{j,0}^{[\nu]} = \frac{p_{j,0}^{[\nu]''} + p_{j-1,m}^{[\nu]''}}{2} + \frac{p_{j,0}^{[\nu]'} - p_{j-1,m}^{[\nu]'}}{h} - f(t_{j,l}, p_{j,l}^{[\nu]}, p_{j,l}^{[\nu]'}), & & j > 1, l = 0, \end{cases} \quad (3.3)$$

wobei für die Konstruktion des so genannten Sprungdefekts $d_{j,0}^{[\nu]}$, $j > 1$ auf [12] verwiesen wird. Die Notwendigkeit dieser „Korrektur“ liegt in der unstetigen 1. Ableitung $p^{[\nu]'}(t)$ der Polynome $p^{[\nu]}(t)$ an den Intervallenden.

Das zugehörige Nachbarproblem wird wie folgt definiert:

$$\begin{cases} z_{1,1}^{[\nu]} = u_0 + h \left(v_0 + d_{1,0}^{B,[\nu]} \right) + \frac{h^2}{2} \left[f \left(t_{1,0}, u_0, v_0 + d_{1,0}^{B,[\nu]} \right) + d_{1,0}^{A,[\nu]} \right], & j = 1, l = 1 \\ \frac{z_{j,l+1}^{[\nu]} - 2z_{j,l}^{[\nu]} + z_{j,l-1}^{[\nu]}}{h^2} = f \left(t_{j,l}, z_{j,l}^{[\nu]}, \frac{z_{j,l+1}^{[\nu]} - z_{j,l-1}^{[\nu]}}{2h} \right) + d_{j,l}^{[\nu]}, & j \geq 1, l > 1 \\ \frac{z_{j,1}^{[\nu]} - 2z_{j-1,m}^{[\nu]} + z_{j-1,m-1}^{[\nu]}}{h^2} = f \left(t_{j-1,m}, z_{j-1,m}^{[\nu]}, \frac{z_{j,1}^{[\nu]} - z_{j-1,m-1}^{[\nu]}}{2h} \right) + d_{j,0}^{[\nu]}, & j > 1, l = 1. \end{cases} \quad (3.4)$$

Der Rest des Algorithmus, die Berechnung der verbesserten Lösung usw., ist ident mit jenem für Differentialgleichungen 1. Ordnung.

Bemerkung 3.1. Setzt man für den $d_{1,0}^{B,[\nu]}$ -Defekt (=Defekt bezüglich der Ableitungsanfangsbedingung) aus (3.3) in die oberste Zeile von (3.4) ein, vereinfacht sich diese zu

$$z_{1,1}^{[\nu]} = u_0 + h p_{1,0}^{[\nu]'} + \frac{h^2}{2} \left[f \left(t_{1,0}, u_0, p_{1,0}^{[\nu]'} \right) + d_{1,0}^{A,[\nu]} \right], \quad j = 1, l = 1, \quad (3.5)$$

allerdings „verschleiert“ diese Darstellung den Zugang zur bzw. mittels Defektkorrektur.

Bemerkung 3.2. Da sowohl der 2. Differenzenquotient auf der linken Seite von (3.2), als auch die Approximation der 2. Ableitung in f auf der rechten Seite Ordnung 2 haben, hat auch das Basisverfahren Ordnung 2. Man erhält für die Folge der Lösungen $x_{j,l}^{[\nu]}$, $\nu = 0, 1, \dots$ die Ordnungsfolgen 2, 4, \dots

Beispiel 3.3. Gegeben sei das skalare nichtlineare Anfangswertproblem

$$u''(t) = \frac{u'(t)}{1 + u^2(t)} - u(t) - \frac{\cos(t)}{1 + (1 + \sin(t))^2} + 1, \quad (3.6a)$$

$$u(t_0 = 0) = 1, u'(t_0 = 0) = v(t_0 = 0) = 1 \quad (3.6b)$$

mit der exakten Lösung $u^*(t) = 1 + \sin(t)$. Die IDEC2-DQ2-Lösung erfolgt auf einem äquidistanten Gitter gemäß Definition 1.1 mit Polynomgrad 6 im Intervall $[t_0, t_{end}] = [\mathbf{t}_0, \mathbf{t}_N] = [0, 3]$. Die globalen Fehler bezüglich der exakten Lösung $e^{*,[\#]}(t_{end}) = p^{[\#]}(t_{end}) - u^*(t_{end})$, $\# = \nu, C$ und die daraus berechneten Konvergenzordnungen an der Stelle $t_{end} = 3$ für das Basisverfahren, IDEC2 und für ein Standard-Kollokationsverfahren für Differentialgleichungen 2. Ordnung (siehe Abschnitt 5.4) sind in Tabelle 3.1 aufgelistet.

h	DQ2	IDeC21	IDeC22	IDeC23	COLL
0,1	4,17E-05	2,55E-09	6,35E-12	6,27E-12	2,10E-12
0,05	1,04E-05	1,60E-10	9,90E-14	9,77E-14	3,26E-14
0,025	2,61E-06	9,99E-12	1,54E-15	1,52E-15	5,08E-16
0,0125	6,52E-07	6,24E-13	2,41E-17	2,38E-17	7,92E-18
0,1	2,00	4,00	6,00	6,00	6,01
0,05	2,00	4,00	6,00	6,00	6,00
0,025	2,00	4,00	6,00	6,00	6,00
0,0125	2,00	4,00	6,00	6,00	6,00

Tabelle 3.1: Resultate zu Beispiel 3.3.

h	DQ2	IDeC21	IDeC22	IDeC23	COLL
0,1	6,00E-05	1,41E-08	2,17E-08	2,17E-08	7,16E-11
0,05	1,50E-05	8,81E-10	1,35E-09	1,35E-09	2,49E-12
0,025	3,75E-06	5,50E-11	8,44E-11	8,44E-11	8,17E-14
0,0125	9,38E-07	3,44E-12	5,27E-12	5,27E-12	2,61E-15
0,1	2,00	4,00	4,00	4,00	4,85
0,05	2,00	4,00	4,00	4,00	4,93
0,025	2,00	4,00	4,00	4,00	4,93
0,0125	2,00	4,00	4,00	4,00	4,97

Tabelle 3.2: Resultate zu Beispiel 3.5.

Bemerkung 3.4. Dieses Standard-Kollokationsverfahren (siehe dazu Abschnitt 5.4) ist für keine der Defektkorrekturvarianten mit 2. Differenzenquotienten $I^*DeC2-DQ2$ gleich dem Fixpunktverfahren, da in diesem Verfahren der Sprungdefekt und der Startschrittdefekt des Defektkorrekturverfahrens *nicht* berücksichtigt werden. Es dient lediglich als Vergleichsverfahren zur Überprüfung der erzielbaren Ordnungen. Für die IDeC2-Verfahren mit Lösung durch Rückführung auf ein System 1. Ordnung (siehe Abschnitt 3.4) ist das Standard-Kollokationsverfahren nur dann der Fixpunkt, wenn als Basisverfahren das implizite Euler-Verfahren verwendet wird.

Beispiel 3.5. Wie Beispiel 3.3, jedoch mit 5 äquidistant liegenden Kollokationsknoten. Tabelle 3.2 zeigt die Resultate.

Bemerkung 3.6. Man sieht in Tabelle 3.2, dass die Ordnungen für IDeC2-DQ2 und das Vergleichskollokationsverfahren bei einer ungeraden Anzahl von Kollokationsknoten nicht übereinstimmen. Während das Kollokationsverfahren bei 5 Knoten erwartungsgemäß Ordnung 5 liefert, stagniert die IDeC2-DQ2-Ordnung bei 4. Ohne diese Diskrepanz analysiert zu haben, liegt die Vermutung nahe, dass diese Ordnungsreduktion durch den beim DQ2-Verfahren zwingend erforderlichen Sprungdefekt verursacht wird.

3.1.2 IDeC2-DQ2-PC

Bei der Diskretisierung der 1. Ableitung in f von (3.1) durch den symmetrischen 1. Differenzenquotienten 2. Ordnung ist in jedem Schritt ein im Allgemeinen nichtlineares Gleichungssystem zu lösen. Um den Rechenaufwand zu minimieren, bietet sich (zumindest für nicht steife Probleme) ein Predictor-Corrector-Verfahren (PC-Verfahren) an. Der Startschritt des PC-Basisverfahrens ist ident mit dem in Abschnitt 3.1.1 vorgestellten Verfahren. Für den Predictor- und den Corrector-Schritt des Basisverfahrens der Ordnung 2 gilt¹

$$\begin{cases} x_{j,l+1}^{P,[0]} = 2x_{j,l}^{[0]} - x_{j,l-1}^{[0]} + h^2 f \left(t_{j,l}, x_{j,l}^{[0]}, \frac{x_{j,l}^{[0]} - x_{j,l-1}^{[0]}}{h} \right), & j \geq 1, l > 1 \\ x_{j,1}^{P,[0]} = 2x_{j-1,m}^{[0]} - x_{j-1,m-1}^{[0]} + h^2 f \left(t_{j-1,m}, x_{j-1,m}^{[0]}, \frac{x_{j-1,m}^{[0]} - x_{j-1,m-1}^{[0]}}{h} \right), & j > 1, l = 1 \end{cases} \quad (3.7)$$

bzw.

$$\begin{cases} x_{j,l+1}^{[0]} = 2x_{j,l}^{[0]} - x_{j,l-1}^{[0]} + h^2 f \left(t_{j,l}, x_{j,l}^{[0]}, \frac{x_{j,l+1}^{P,[0]} - x_{j,l-1}^{[0]}}{2h} \right), & j \geq 1, l > 1 \\ x_{j,1}^{[0]} = 2x_{j-1,m}^{[0]} - x_{j-1,m-1}^{[0]} + h^2 f \left(t_{j-1,m}, x_{j-1,m}^{[0]}, \frac{x_{j,1}^{P,[0]} - x_{j-1,m-1}^{[0]}}{2h} \right), & j > 1, l = 1. \end{cases} \quad (3.8)$$

Mit Hilfe der Defektdefinition (3.3) ergibt sich für das Nachbarproblem (ohne Startschritt, weil ident mit jenem aus (3.4))

$$\begin{cases} z_{j,l+1}^{P,[\nu]} = 2z_{j,l}^{[\nu]} - z_{j,l-1}^{[\nu]} + h^2 \left[f \left(t_{j,l}, z_{j,l}^{[\nu]}, \frac{z_{j,l}^{[\nu]} - z_{j,l-1}^{[\nu]}}{h} \right) + d_{j,l}^{[\nu]} \right], & j \geq 1, l > 1 \\ z_{j,1}^{P,[\nu]} = 2z_{j-1,m}^{[\nu]} - z_{j-1,m-1}^{[\nu]} + h^2 \left[f \left(t_{j-1,m}, z_{j-1,m}^{[\nu]}, \frac{z_{j-1,m}^{[\nu]} - z_{j-1,m-1}^{[0]}}{h} \right) + d_{j,0}^{[\nu]} \right], & j > 1, l = 1 \end{cases} \quad (3.9)$$

bzw.

$$\begin{cases} z_{j,l+1}^{[\nu]} = 2z_{j,l}^{[\nu]} - z_{j,l-1}^{[\nu]} + h^2 \left[f \left(t_{j,l}, z_{j,l}^{[\nu]}, \frac{z_{j,l+1}^{P,[\nu]} - z_{j,l-1}^{[\nu]}}{2h} \right) + d_{j,l}^{[\nu]} \right], & j \geq 1, l > 1 \\ z_{j,1}^{[\nu]} = 2z_{j-1,m}^{[\nu]} - z_{j-1,m-1}^{[\nu]} + h^2 \left[f \left(t_{j-1,m}, z_{j-1,m}^{[\nu]}, \frac{z_{j,1}^{P,[\nu]} - z_{j-1,m-1}^{[\nu]}}{2h} \right) + d_{j,0}^{[\nu]} \right], & j > 1, l = 1. \end{cases} \quad (3.10)$$

Beispiel 3.7. Wie Beispiel 3.3, jedoch mit einem Predictor-Corrector-Verfahren als Basisverfahren. Tabelle 3.3 zeigt die Resultate.

3.2 Integralmitteldefekt – IQDeC2-DQ2

3.2.1 Diskretisierung der 2. Ableitung mittels lokaler Integration

Einleitend werden die Grundlagen, um den Integralmitteldefekt für Differentialgleichungen 2. Ordnung verstehen zu können, kurz zusammengefasst wiedergegeben. Eine ausführliche Darstellung findet man unter anderem in [19].

¹Zur Vereinfachung der Notation gilt $x_{j,l}^{[0]} := x_{j,l}^{C,[0]}$.

h	DQ2-PC	IDeC21	IDeC22	IDeC23	COLL
0,1	3,41E-05	2,37E-09	6,41E-12	6,27E-12	2,10E-12
0,05	8,54E-06	1,48E-10	9,98E-14	9,77E-14	3,26E-14
0,025	2,13E-06	9,28E-12	1,56E-15	1,52E-15	5,08E-16
0,0125	5,33E-07	5,80E-13	2,43E-17	2,38E-17	7,92E-18
0,1	2,00	4,00	6,00	6,00	6,01
0,05	2,00	4,00	6,00	6,00	6,00
0,025	2,00	4,00	6,00	6,00	6,00
0,0125	2,00	4,00	6,00	6,00	6,00

Tabelle 3.3: Resultate zu Beispiel 3.7.

Äquidistantes Gitter

Ist die Schrittweite des Gitters 1.1 konstant $h \equiv h_{j,l}$, so kann man die 2. Ableitung $u''(t)$ (einer zweimal stetig differenzierbaren Funktion $u(t)$) wie folgt mittels Integration diskretisieren: Wir betrachten

$$\int_{-1}^1 (1 - |\vartheta|) u''(t + \vartheta h) d\vartheta = \int_{-1}^0 (1 + \vartheta) u''(t + \vartheta h) d\vartheta + \int_0^1 (1 - \vartheta) u''(t + \vartheta h) d\vartheta. \quad (3.11)$$

Partielle Integration und Einsetzen der Integralgrenzen zeigen, dass (3.11) mit dem 2. Differenzenquotienten übereinstimmt:

$$\begin{aligned} (3.11) &= \left. \frac{(1 + \vartheta) u'(t + \vartheta h)}{h} \right|_{-1}^0 - \int_{-1}^0 \frac{u'(t + \vartheta h)}{h} d\vartheta \\ &+ \left. \frac{(1 - \vartheta) u'(t + \vartheta h)}{h} \right|_0^1 + \int_0^1 \frac{u'(t + \vartheta h)}{h} d\vartheta \\ &= \frac{u'(t)}{h} - \frac{u(t) - u(t-h)}{h^2} - \frac{u'(t)}{h} + \frac{u(t+h) - u(t)}{h^2} \\ &= \frac{u(t+h) - 2u(t) + u(t-h)}{h^2}. \end{aligned} \quad (3.12)$$

Bemerkung 3.8. Man sieht, dass (3.12) auch dann gilt, wenn u'' an der Stelle t unstetig ist. Diese Tatsache ist wichtig im Hinblick auf den IQDeC2-DQ2-Sprungdefekt.

Nicht äquidistantes Gitter

Wendet man obige Ausführungen auf ein nicht äquidistantes Gitter gemäß Definition 1.1 an, sind das linke und das rechte Integral mit je einem Faktor $(\gamma_{li}, \gamma_{re})$ zu gewichten. Der Ansatz

$$\gamma_{li} \int_{-1}^0 (1 + \vartheta) u''(t + \vartheta h_{j,l}) d\vartheta + \gamma_{re} \int_0^1 (1 - \vartheta) u''(t + \vartheta h_{j,l+1}) d\vartheta \quad (3.13)$$

mit $\gamma_{li} = \frac{2\delta_{j,l}}{1+\delta_{j,l}}$ bzw. $\gamma_{re} = \frac{2}{1+\delta_{j,l}}$, wobei $\delta_{j,l} = \frac{h_{j,l}}{h_{j,l+1}}$, führt zum gewünschten 2. Differenzenquotienten: Setzt man für γ_{li} und γ_{re} in (3.13) ein, wertet analog zu (3.12) die Integrale aus

und stellt die Terme um, erhält man

$$(3.13) = \frac{\frac{u(t+h_{j,l+1})-u(t+h_{j,l})}{h_{j,l+1}} - \frac{u(t+h_{j,l})-u(t+h_{j,l-1})}{h_{j,l-1}}}{\frac{h_{j,l}+h_{j,l+1}}{2}} \quad (3.14)$$

als Approximation der 2. Ableitung $u''(t)$ auf einem nicht äquidistanten Gitter. Diese hat auf einem nicht äquidistanten Gitter lokal nur die Ordnung 1, global jedoch die Ordnung 2, wie in [16, S.27] im Kontext von Randwertproblemen gezeigt.

Die konkrete Umsetzung der Diskretisierung der 2. Ableitung ist in den Maple-Programmen mittels Quadratur realisiert. Die Quadraturkoeffizienten sind dabei so gewählt, dass ein beliebiges Polynom $q(s)$ vom Grad $\leq m-1$ exakt integriert wird. Mit Hilfe der Lagrange-polynome (5.16) gilt für die Quadraturkoeffizienten

$$\alpha_{l,i}^{li} = \int_{-1}^0 (1+\vartheta) \Lambda_i(c_l + \vartheta(c_l - c_{l-1})) d\vartheta, \quad l=1, \dots, m, \quad i=0, \dots, m, \quad (3.15a)$$

$$\alpha_{l,i}^{re} = \int_0^1 (1-\vartheta) \Lambda_i(c_l + \vartheta(c_{l+1} - c_l)) d\vartheta, \quad l=1, \dots, m, \quad i=0, \dots, m. \quad (3.15b)$$

Die numerische Diskretisierung der 2. Ableitung eines Polynoms q auf einem nicht äquidistanten Gitter kann mittels Quadratur somit wie folgt angeschrieben werden

$$\frac{\frac{q(t_{j,l}+h_{j,l+1})-q(t_{j,l}+h_{j,l})}{h_{j,l+1}} - \frac{q(t_{j,l}+h_{j,l})-q(t_{j,l}+h_{j,l-1})}{h_{j,l-1}}}{\frac{h_{j,l}+h_{j,l+1}}{2}} = \gamma_{li} \sum_{\mu=0}^m \alpha_{l,\mu}^{li} q(c_\mu) + \gamma_{re} \sum_{\mu=0}^m \alpha_{l,\mu}^{re} q(c_\mu), \quad l=1, \dots, m. \quad (3.16)$$

3.2.2 IQDeC2-DQ2

Dieses Verfahren ist verwandt zu der in [21] für Randwertprobleme vorgestellten Konstruktion. Wie IQDeC für Differentialgleichungen 1. Ordnung funktioniert das Verfahren auch auf nicht äquidistanten Gittern, allerdings bei Anfangswertproblemen mit einer Einschränkung: Die Ordnungsfolge auf nicht äquidistanten Gittern lautet 2,3,4,... anstatt 2,4,6,..., da die erste Ableitung bei Anfangswertproblemen im Gegensatz zu Randwertproblemen an den Intervallenden nicht stetig ist (siehe dazu die numerischen Experimente am Ende des Abschnitts).

Das Basisverfahren (der Ordnung 2) zur Lösung des Anfangswertproblems (3.1) ist eine Erweiterung für nicht äquidistante Gitter gemäß Definition 1.1 von jenem des IDeC2-DQ2-Verfahrens aus Abschnitt 3.1.1. Es gilt somit

$$\begin{cases} x_{1,1}^{[0]} = u_0 + h_{1,1} v_0 + \frac{h_{1,1}^2}{2} f(t_{1,0}, u_0, v_0), & j=1, l=1 \\ \frac{\frac{x_{j,l+1}^{[0]} - x_{j,l}^{[0]}}{h_{j,l+1}} - \frac{x_{j,l}^{[0]} - x_{j,l-1}^{[0]}}{h_{j,l-1}}}{\frac{h_{j,l}+h_{j,l+1}}{2}} = f\left(t_{j,l}, x_{j,l}^{[0]}, \frac{x_{j,l+1}^{[0]} - x_{j,l-1}^{[0]}}{h_{j,l+1}+h_{j,l}}\right), & j \geq 1, l > 1 \\ \frac{\frac{x_{j,1}^{[0]} - x_{j-1,m}^{[0]}}{h_{j,1}} - \frac{x_{j-1,m}^{[0]} - x_{j-1,m-1}^{[0]}}{h_{j-1,m}}}{\frac{h_{j-1,m}+h_{j,1}}{2}} = f\left(t_{j,l}, x_{j-1,m}^{[0]}, \frac{x_{j,1}^{[0]} - x_{j-1,m-1}^{[0]}}{h_{j,1}+h_{j-1,m}}\right), & j > 1, l=1. \end{cases} \quad (3.17)$$

Für den Startschritt ($j = 1, l = 1$) kann bei Anfangswertproblemen der Defekt für Ableitungsrandbedingungen (siehe dazu [8]) herangezogen werden. Aufgrund der 2. Anfangsbedingung $u'(t_0 = 0) = v_0$ gilt das exakte Differenzenschema

$$\frac{u(t_{1,1}) - u_0}{h_{1,1}} = v_0 + h_{1,1} \int_0^1 (1 - \vartheta) \underbrace{f(\vartheta h_{1,1}, u(\vartheta h_{1,1}), u'(\vartheta h_{1,1}))}_{=u''(\vartheta h_{1,1})} d\vartheta. \quad (3.18)$$

Daraus folgt beim Ersetzen der exakten Lösung $u(t)$ durch die aus der Defektkorrektur erhaltenen Polynome $p^{[\nu]}(t)$ sowie des Integrals durch die entsprechende Quadraturformel mit der Abkürzung $p_{j,l}^{[\nu]} \equiv p^{[\nu]}(t_{j,l}) = x_{j,l}^{[\nu]}$ der Defekt für den Startschritt in der Form

$$\bar{d}_{1,0}^{[\nu]} = \frac{p_{1,1}^{[\nu]} - u_0}{h_{1,1}} - v_0 - h_{1,1} \sum_{\mu=0}^m \alpha_{1,\mu}^{re} f(t_{1,\mu}, p_{1,\mu}^{[\nu]}, p_{1,\mu}^{[\nu]'}) . \quad (3.19)$$

Der IQDeC2-DQ2-Defekt für innere Knoten wird analog zu (1.122) als Integralmittel des klassischen Defekts (3.3) unter Einbeziehung der Erkenntnisse aus Abschnitt 3.2.1 definiert:

$$\begin{aligned} \bar{d}_{j,l}^{[\nu]} &= \gamma_{li} \sum_{\mu=0}^m \alpha_{l,\mu}^{li} \left[p_{j,\mu}^{[\nu]''} - f(t_{j,\mu}, p_{j,\mu}^{[\nu]}, p_{j,\mu}^{[\nu]'}) \right] \\ &+ \gamma_{re} \sum_{\mu=0}^m \alpha_{l,\mu}^{re} \left[p_{j,\mu}^{[\nu]''} - f(t_{j,\mu}, p_{j,\mu}^{[\nu]}, p_{j,\mu}^{[\nu]'}) \right], \quad j > 1, l = 1, \dots, m. \end{aligned} \quad (3.20)$$

Aufgrund der unstetigen 1. Ableitung $p^{[\nu]'}(t)$ an den Intervallenden $t_{j-1,m} = \mathbf{t}_j = t_{j,0}$ muss an diesen Stellen ein „Sprungterm“ im Defekt berücksichtigt werden,

$$\begin{aligned} \bar{d}_{j,0}^{[\nu]} &= \gamma_{li} \left(\sum_{\mu=0}^m \alpha_{m,\mu}^{li} \left[p_{j-1,\mu}^{[\nu]''} - f(t_{j-1,\mu}, p_{j-1,\mu}^{[\nu]}, p_{j-1,\mu}^{[\nu]'}) \right] - \frac{p_{j-1,m}^{[\nu]'}}{h_{j-1,m}} \right) \\ &+ \gamma_{re} \left(\sum_{\mu=0}^m \alpha_{0,\mu}^{re} \left[p_{j,\mu}^{[\nu]''} - f(t_{j,\mu}, p_{j,\mu}^{[\nu]}, p_{j,\mu}^{[\nu]'}) \right] + \frac{p_{j,0}^{[\nu]'}}{h_{j,1}} \right), \quad j > 1, l = 0. \end{aligned} \quad (3.21)$$

Begründen kann man das Zustandekommen der Sprungdefekte wie folgt: Bei der Integration von $p_{j-1}^{[\nu]''}(t)$ und $p_j^{[\nu]''}(t)$ entstehen Terme, die sich zu Null summieren müssen, damit der 2. Differenzenquotient korrekt wiedergegeben wird. Da man an den Intervallgrenzen jedoch zwei verschiedene Polynome integriert, müssen die nicht verschwindenden Terme durch die zusätzlichen Sprungdefektterme $-\gamma_{li} \frac{p_{j-1,m}^{[\nu]'}}{h_{j-1,m}} + \gamma_{re} \frac{p_{j,0}^{[\nu]'}}{h_{j,1}}$ „künstlich“ eliminiert werden.

Mit der Definition des Nachbarproblems wird der Algorithmus komplettiert:

$$\left\{ \begin{aligned} z_{1,1}^{[\nu]} &= u_0 + h_{1,1} \left(v_0 + \bar{d}_{1,1}^{[\nu]} \right) + \frac{h_{1,1}^2}{2} f(t_{1,0}, u_0, v_0 + \bar{d}_{1,0}^{[\nu]}), & j = 1, l = 1 \\ \frac{z_{j,l+1}^{[\nu]} - z_{j,l}^{[\nu]}}{h_{j,l+1}} - \frac{z_{j,l}^{[\nu]} - z_{j,l-1}^{[\nu]}}{h_{j,l}} &= f\left(t_{j,l}, z_{j,l}^{[\nu]}, \frac{z_{j,l+1}^{[\nu]} - z_{j,l-1}^{[\nu]}}{h_{j,l+1} + h_{j,l}}\right) + \bar{d}_{j,l}^{[\nu]}, & j \geq 1, l > 1 \\ \frac{z_{j,1}^{[\nu]} - z_{j-1,m}^{[\nu]}}{h_{j,1}} - \frac{z_{j-1,m}^{[\nu]} - z_{j-1,m-1}^{[\nu]}}{h_{j-1,m}} &= f\left(t_{j,l}, z_{j-1,m}^{[\nu]}, \frac{z_{j,1}^{[\nu]} - z_{j-1,m-1}^{[\nu]}}{h_{j,1} + h_{j-1,m}}\right) + \bar{d}_{j,0}^{[\nu]}, & j > 1, l = 1. \end{aligned} \right. \quad (3.22)$$

h	DQ2	IQDeC21	IQDeC22	IQDeC23	IQDeC24	COLL
0,1	4,17E-05	1,81E-09	8,44E-14	5,96E-17	5,54E-17	7,61E-17
0,05	1,04E-05	1,13E-10	1,32E-15	2,33E-19	2,16E-19	2,97E-19
0,025	2,61E-06	7,09E-12	2,06E-17	9,09E-22	8,44E-22	1,16E-21
0,0125	6,52E-07	4,43E-13	3,22E-19	3,55E-24	3,30E-24	4,54E-24
0,1	2,00	4,00	6,00	8,00	8,00	8,00
0,05	2,00	4,00	6,00	8,00	8,00	8,00
0,025	2,00	4,00	6,00	8,00	8,00	8,00
0,0125	2,00	4,00	6,00	8,00	8,00	8,00

Tabelle 3.4: Resultate zu Beispiel 3.10.

h	DQ2	IQDeC21	IQDeC22	IQDeC23	COLL
0,1	6,00E-05	3,76E-09	2,35E-13	4,88E-13	5,76E-13
0,05	1,50E-05	2,35E-10	3,68E-15	7,62E-15	9,01E-15
0,025	3,75E-06	1,47E-11	5,76E-17	1,19E-16	1,41E-16
0,0125	9,38E-07	9,19E-13	9,00E-19	1,86E-18	2,20E-18
0,1	2,00	4,00	6,00	6,00	6,00
0,05	2,00	4,00	6,00	6,00	6,00
0,025	2,00	4,00	6,00	6,00	6,00
0,0125	2,00	4,00	6,00	6,00	6,00

Tabelle 3.5: Resultate zu Beispiel 3.12.

Bemerkung 3.9. Für die folgenden Beispiele werden zu Vergleichszwecken die globalen Fehler und die Ordnungen eines Standard-Kollokationsverfahrens (siehe Abschnitt 5.4) angegeben. Damit die Ordnungen mit jenen von IQDeC2-DQ2 übereinstimmen, muss der linke Randknoten ($c_0 = 0$) zusätzlich als Kollokationsknoten definiert werden.

Beispiel 3.10. Wie Beispiel 3.3, jedoch mit Integralmitteldefekt und inkl. linkem Randknoten 7 Kollokationsknoten. Tabelle 3.4 zeigt die Resultate.

Bemerkung 3.11. Man sieht durch Vergleich der Beispiele 3.3 und 3.10, dass der zusätzliche Kollokationsknoten (=linker Randknoten) bei IQDeC2-DQ2 eine Verbesserung der maximal erzielbaren Konvergenzordnung um 2 Ordnungen bringt.

Beispiel 3.12. Wie Beispiel 3.3, jedoch mit Integralmitteldefekt und inkl. linkem Randknoten 6 Kollokationsknoten. Tabelle 3.5 zeigt die Resultate.

Beispiel 3.13. Wie Beispiel 3.3, jedoch mit Integralmitteldefekt auf einem nicht äquidistanten Gitter mit 7 Kollokationsknoten ($c_0 = 0$; $c_1 = 0,0205$; $c_2 = 0,3004$; $c_3 = 0,4567$; $c_4 = 0,7779$; $c_5 = 0,9288$; $c_6 = 1$). Tabelle 3.6 zeigt die Resultate.

Beispiel 3.14. Wie Beispiel 3.3, jedoch mit Integralmitteldefekt auf einem nicht äquidistanten Gitter mit 6 Kollokationsknoten ($c_0 = 0$; $c_1 = 0,0205$; $c_2 = 0,3004$; $c_3 = 0,4567$; $c_4 = 0,8779$; $c_5 = 1$). Tabelle 3.7 zeigt die Resultate.

h	DQ2	IQDeC21	IQDeC22	IQDeC23	IQDeC24	IQDeC25	COLL
0,1	9,43E-05	2,28E-08	5,07E-11	2,58E-14	3,22E-15	3,13E-15	8,76E-16
0,05	2,36E-05	1,03E-09	3,37E-12	1,92E-16	2,58E-17	2,43E-17	7,01E-18
0,025	5,88E-06	1,46E-11	2,13E-13	3,56E-18	2,14E-19	1,89E-19	5,54E-20
0,0125	1,47E-06	5,28E-12	1,33E-14	2,59E-19	1,87E-21	1,48E-21	4,35E-22
0,1	2,00	4,47	3,91	7,07	6,96	7,01	6,97
0,05	2,00	6,14	3,99	5,75	6,91	7,00	6,98
0,025	2,00	1,47	4,00	3,78	6,84	7,00	6,99
0,0125	2,00	1,47	4,00	3,78	6,84	7,00	6,99

Tabelle 3.6: Resultate zu Beispiel 3.13.

h	DQ2	IQDeC21	IQDeC22	IQDeC23	IQDeC24	COLL
0,1	1,53E-04	2,75E-07	4,69E-10	1,44E-12	1,39E-12	1,65E-12
0,05	3,83E-05	2,87E-08	2,28E-11	2,28E-14	2,18E-14	2,58E-14
0,025	9,59E-06	3,24E-09	1,24E-12	3,56E-16	3,41E-16	4,03E-16
0,0125	2,40E-06	3,83E-10	7,20E-14	5,50E-18	5,33E-18	6,30E-18
0,1	2,00	3,26	4,36	5,99	6,00	6,00
0,05	2,00	3,15	4,20	6,00	6,00	6,00
0,025	2,00	3,08	4,11	6,01	6,00	6,00
0,0125	2,00	3,08	4,11	6,01	6,00	6,00

Tabelle 3.7: Resultate zu Beispiel 3.14.

Bemerkung 3.15. Man sieht, wie bei allen Verfahren mit Integralmitteldefekt (auch für Differentialgleichungen 1. Ordnung), dass am nicht äquidistanten Gitter Konvergenz gegen den Fixpunkt beobachtet werden kann, allerdings nicht mit der Ordnungsfolge $2, 4, \dots$, sondern mit Einser- bzw. Einser- und Zweierschritten (siehe Beispiele 3.13 und 3.14).

3.2.3 IQDeC2-DQ2-PC

Kombiniert man das auf nicht äquidistante Gitter erweiterte PC-Basisverfahren (3.7) bzw. (3.8) mit dem Startschritt aus (3.17)

$$\begin{cases} x_{j,l+1}^{P,[0]} = x_{j,l}^{[0]} + h_{j,l+1} \frac{x_{j,l}^{[0]} - x_{j,l-1}^{[0]}}{h_{j,l}} + h_{j,l+1} \frac{h_{j,l} + h_{j,l+1}}{2} f \left(t_{j,l}, x_{j,l}^{[0]}, \frac{x_{j,l}^{[0]} - x_{j,l-1}^{[0]}}{h_{j,l}} \right), & j \geq 1, l > 1 \\ x_{j,1}^{P,[0]} = x_{j-1,m}^{[0]} + h_{j,1} \frac{x_{j-1,m}^{[0]} - x_{j-1,m-1}^{[0]}}{h_{j-1,m}} \\ \quad + h_{j,1} \frac{h_{j-1,m} + h_{j,1}}{2} f \left(t_{j-1,m}, x_{j-1,m}^{[0]}, \frac{x_{j-1,m}^{[0]} - x_{j-1,m-1}^{[0]}}{h_{j-1,m}} \right), & j > 1, l = 1 \end{cases} \quad (3.23)$$

bzw.

$$\begin{cases} x_{j,l+1}^{[0]} = x_{j,l}^{[0]} + h_{j,l+1} \frac{x_{j,l}^{[0]} - x_{j,l-1}^{[0]}}{h_{j,l}} + h_{j,l+1} \frac{h_{j,l} + h_{j,l+1}}{2} f \left(t_{j,l}, x_{j,l}^{[0]}, \frac{x_{j,l+1}^{P,[0]} - x_{j,l-1}^{[0]}}{h_{j,l} + h_{j,l+1}} \right), & j \geq 1, l > 1 \\ x_{j,1}^{[0]} = x_{j-1,m}^{[0]} + h_{j,1} \frac{x_{j-1,m}^{[0]} - x_{j-1,m-1}^{[0]}}{h_{j-1,m}} \\ \quad + h_{j,1} \frac{h_{j-1,m} + h_{j,1}}{2} f \left(t_{j-1,m}, x_{j-1,m}^{[0]}, \frac{x_{j,1}^{P,[0]} - x_{j-1,m-1}^{[0]}}{h_{j-1,m} + h_{j,1}} \right), & j > 1, l = 1 \end{cases} \quad (3.24)$$

mit den Defekten (3.19) bis (3.21) und dem Nachbarproblem

$$\begin{cases} z_{j,l+1}^{P,[\nu]} = z_{j,l}^{[\nu]} + h_{j,l+1} \frac{z_{j,l}^{[\nu]} - z_{j,l-1}^{[\nu]}}{h_{j,l}} \\ \quad + h_{j,l+1} \frac{h_{j,l} + h_{j,l+1}}{2} \left[f \left(t_{j,l}, z_{j,l}^{[\nu]}, \frac{z_{j,l}^{[\nu]} - z_{j,l-1}^{[\nu]}}{h_{j,l}} \right) + \bar{d}_{j,l}^{[\nu]} \right], & j \geq 1, l > 1 \\ z_{j,1}^{P,[\nu]} = z_{j-1,m}^{[\nu]} + h_{j,1} \frac{z_{j-1,m}^{[\nu]} - z_{j-1,m-1}^{[\nu]}}{h_{j-1,m}} \\ \quad + h_{j,1} \frac{h_{j-1,m} + h_{j,1}}{2} \left[f \left(t_{j-1,m}, z_{j-1,m}^{[\nu]}, \frac{z_{j-1,m}^{[\nu]} - z_{j-1,m-1}^{[\nu]}}{h_{j-1,m}} \right) + \bar{d}_{j,0}^{[\nu]} \right], & j > 1, l = 1 \end{cases} \quad (3.25)$$

bzw.

$$\begin{cases} z_{j,l+1}^{[\nu]} = z_{j,l}^{[\nu]} + h_{j,l+1} \frac{z_{j,l}^{[\nu]} - z_{j,l-1}^{[\nu]}}{h_{j,l}} \\ \quad + h_{j,l+1} \frac{h_{j,l} + h_{j,l+1}}{2} \left[f \left(t_{j,l}, z_{j,l}^{[\nu]}, \frac{z_{j,l+1}^{P,[\nu]} - z_{j,l-1}^{[\nu]}}{h_{j,l} + h_{j,l+1}} \right) + \bar{d}_{j,l}^{[\nu]} \right], & j \geq 1, l > 1 \\ z_{j,1}^{[\nu]} = z_{j-1,m}^{[\nu]} + h_{j,1} \frac{z_{j-1,m}^{[\nu]} - z_{j-1,m-1}^{[\nu]}}{h_{j-1,m}} \\ \quad + h_{j,1} \frac{h_{j-1,m} + h_{j,1}}{2} \left[f \left(t_{j-1,m}, z_{j-1,m}^{[\nu]}, \frac{z_{j,1}^{P,[\nu]} - z_{j-1,m-1}^{[\nu]}}{h_{j-1,m} + h_{j,1}} \right) + \bar{d}_{j,0}^{[\nu]} \right], & j > 1, l = 1, \end{cases} \quad (3.26)$$

erhält man ein Verfahren, welches für nicht steife Differentialgleichungen ohne das aufwändige Lösen eines in der Regel nichtlinearen Gleichungssystems in jedem Schritt auskommt.

Beispiel 3.16. Wie Beispiel 3.3, jedoch mit einem PC-Verfahren als Basisverfahren, Integralmitteldefekt und inkl. linkem Randknoten 7 Kollokationsknoten. Tabelle 3.8 zeigt die Resultate.

Bemerkung 3.17. IQDeC2-DQ2-PC zeigt wie zu erwarten identes Verhalten wie IQDeC2-DQ2.

h	DQ2-PC	IQDeC21	IQDeC22	IQDeC23	IQDeC24	COLL
0,1	3,42E-05	1,67E-09	7,69E-14	5,96E-17	5,54E-17	7,61E-17
0,05	8,54E-06	1,04E-10	1,20E-15	2,33E-19	2,16E-19	2,97E-19
0,025	2,13E-06	6,53E-12	1,88E-17	9,09E-22	8,44E-22	1,16E-21
0,0125	5,33E-07	4,08E-13	2,93E-19	3,55E-24	3,30E-24	4,54E-24
0,1	2,00	4,00	6,00	8,00	8,00	8,00
0,05	2,00	4,00	6,00	8,00	8,00	8,00
0,025	2,00	4,00	6,00	8,00	8,00	8,00
0,0125	2,00	4,00	6,00	8,00	8,00	8,00

Tabelle 3.8: Resultate zu Beispiel 3.16.

3.3 Interpolierter Defekt – IPDeC2-DQ2

3.3.1 IPDeC2-DQ2

Dieses Verfahren stellt eine Erweiterung des in Abschnitt 1.4 vorgestellten Algorithmus auf die Lösung von Differentialgleichungen 2. Ordnung ohne Rückführung auf ein System 1. Ordnung dar. Die Erweiterungen betreffen den Einbau des erforderlichen Startschritts und die Definition des Sprungdefekts aufgrund der nicht stetigen 1. Ableitung $p^{[\nu]'}(t)$ der stückweise definierten Interpolationspolynome $p^{[\nu]}(t)$ an den Intervallenden.

Für dieses Verfahren werden (wie bei IPDeC, Abschnitt 1.4) ein äquidistantes $t_{j,l}, l = 1, \dots, m$ und ein Gauß-Gitter $\tilde{t}_{j,\kappa}, \kappa = 1, \dots, \tilde{m}$ parallel verwendet (siehe dazu Abbildung 1.4). Der einzige Unterschied ist, dass für die Anzahl der Gauß-Knoten $\tilde{m} = m - 1$ gilt.

Als Basisverfahren wird (3.2) verwendet. Die $x_{j,l}^{[\nu]}, j = 1, \dots, N, l = 1, \dots, m$ werden durch die stückweise definierten Polynomfunktion $p^{[\nu]}(t)$ interpoliert. Daraus wird der Defekt an den Gauß-Knoten $\tilde{t}_{j,\kappa}$

$$\tilde{d}^{[\nu]}(\tilde{t}_{j,\kappa}) = p^{[\nu]''}(\tilde{t}_{j,\kappa}) - f(\tilde{t}_{j,\kappa}, p^{[\nu]}(\tilde{t}_{j,\kappa}), p^{[\nu]'}(\tilde{t}_{j,\kappa})) \quad (3.27)$$

ermittelt. Die so berechneten Werte $\tilde{d}^{[\nu]}(\tilde{t}_{j,\kappa})$ werden mittels der stückweise definierten Polynome $\tilde{p}^{[\nu]}(t)$ (Grad $\leq \tilde{m} - 1$) interpoliert. Wertet man diese an den $t_{j,l}$ aus, gelangt man zum IPDeC-Defekt. Zusätzlich muss für den Startschritt der Defektanteil für die erste Ableitung von $p^{[\nu]'}(t_{j,l})$ aus (3.3) berücksichtigt werden, für die Intervallenden wurde ein adäquater Sprungdefekt konstruiert. Zusammengefasst erhält man für den interpolierten Defekt

$$\begin{cases} \hat{d}_{1,0}^{A,[\nu]} = \tilde{p}^{[\nu]}(t_{1,0}), d_j^{[B,[\nu]}(t_{1,0}) = p^{[\nu]'}(t_{1,0}) - v_0, & j = 1, l = 0 \\ \hat{d}_{j,l}^{[\nu]} = \tilde{p}^{[\nu]}(t_{j,l}), & j \geq 1, l > 0 \\ \hat{d}_{j,0}^{[\nu]} = \frac{p^{[\nu]'}(t_{j,0}) - p^{[\nu]'}(t_{j-1,m})}{h} + \frac{\tilde{p}^{[\nu]''}(t_{j,0}) + \tilde{p}^{[\nu]}(t_{j-1,m})}{2} & j > 1, l = 0, \end{cases} \quad (3.28)$$

welcher statt (3.3) im ansonsten unveränderten IDeC2-DQ2-Algorithmus im Nachbarproblem (3.4) zu berücksichtigen ist.

h	DQ2	IPDeC21	IPDeC22	IPDeC23	IPDeC24	IPDeC25	G-COLL
0,1	4,17E-05	2,56E-09	1,57E-13	5,50E-18	8,38E-20	5,01E-20	5,01E-20
0,05	1,04E-05	1,60E-10	2,45E-15	2,38E-20	8,26E-23	4,88E-23	4,88E-23
0,025	2,61E-06	9,99E-12	3,83E-17	9,66E-23	8,11E-26	4,76E-26	4,76E-26
0,0125	6,52E-07	6,24E-13	5,99E-19	3,83E-25	7,95E-29	4,65E-29	4,65E-29
0,1	2,00	4,00	6,00	7,85	9,99	10,00	10,00
0,05	2,00	4,00	6,00	7,94	9,99	10,00	10,00
0,025	2,00	4,00	6,00	7,98	9,99	10,00	10,00
0,0125	2,00	4,00	6,00	7,98	9,99	10,00	10,00

Tabelle 3.9: Resultate zu Beispiel 3.20.

Das Verfahren verhält sich analog zu IPDeC für Differentialgleichungen 1. Ordnung aus Abschnitt 1.4. Es konvergiert aufgrund des Interpolationsdefekts gegen jene Kollokationslösung $p^{[C]}(t)$, für die der Defekt an den Knoten $\tilde{t}_{j,\kappa}$ verschwindet.

Bemerkung 3.18. Im Sprungdefektterm von (3.28) fehlen beim ersten Bruch keine \sim -Zeichen, da der Sprung in den ersten Ableitungen von $p_j^{[\nu]}(t)$ und nicht von $\tilde{p}_j^{[\nu]}(t)$ korrigiert werden muss.

Bemerkung 3.19. Wie bei IPDeC für Differentialgleichungen 1. Ordnung muss, um die gewünschten Konvergenzordnungen zu erzielen, für Differentialgleichungen 2. Ordnung die Bedingung $\tilde{m} = m - 1$ für die Knotenzahl der beiden verwendeten Gitter strikt eingehalten werden (vgl. Bemerkung 1.54).

Beispiel 3.20. Wie Beispiel 3.3, jedoch mit interpoliertem Defekt, 7 äquidistanten Knoten und 5 Gauß-Knoten: $\tilde{c}_1 = \frac{1}{2} - \frac{\sqrt{245+14\sqrt{70}}}{42}$, $\tilde{c}_2 = \frac{1}{2} - \frac{\sqrt{245-14\sqrt{70}}}{42}$, $\tilde{c}_3 = \frac{1}{2}$, $\tilde{c}_4 = \frac{1}{2} + \frac{\sqrt{245-14\sqrt{70}}}{42}$, $\tilde{c}_5 = \frac{1}{2} + \frac{\sqrt{245+14\sqrt{70}}}{42}$. Tabelle 3.9 zeigt die Resultate.

3.3.2 IPDeC2-DQ2-PC

Kombiniert man das Basisverfahren (3.7) bzw. (3.8) sowie das Nachbarproblem (3.9) bzw. (3.10) von IDeC2-DQ2-PC mit dem Defekt (3.28), erhält man wiederum ein explizites Defektkorrekturverfahren für nicht steife Differentialgleichungen.

Beispiel 3.21. Wie Beispiel 3.20, jedoch mit einem PC-Verfahren als Basisverfahren. Tabelle 3.10 zeigt die Resultate.

3.4 Lösung durch Rückführung auf ein System 1. Ordnung

Die folgenden IQDeC2-Algorithmen wurden in der Frühphase der vorliegenden Arbeit erstellt, um für Anfangswertprobleme 2. Ordnung ein Verfahren mit Ordnungsfolge 2,4,... zu entwickeln, welche auf nicht äquidistantem Gitter konvergieren und somit die Konstruktion von superkonvergenten Verfahren erlauben. Die IDeC2-* -Algorithmen wurden dazu als

h	DQ2-PC	IPDeC21	IPDeC22	IPDeC23	IPDeC24	IPDeC25	G-COLL
0,1	3,42E-05	2,37E-09	2,11E-13	1,96E-17	7,57E-20	5,01E-20	5,01E-20
0,05	8,54E-06	1,48E-10	3,31E-15	7,98E-20	7,46E-23	4,88E-23	4,88E-23
0,025	2,13E-06	9,28E-12	5,17E-17	3,17E-22	7,33E-26	4,76E-26	4,76E-26
0,0125	5,33E-07	5,80E-13	8,07E-19	1,25E-24	7,19E-29	4,65E-29	4,65E-29
0,1	2,00	4,00	6,00	7,94	9,99	10,00	10,00
0,05	2,00	4,00	6,00	7,98	9,99	10,00	10,00
0,025	2,00	4,00	6,00	7,99	9,99	10,00	10,00
0,0125	2,00	4,00	6,00	7,99	9,99	10,00	10,00

Tabelle 3.10: Resultate zu Beispiel 3.21.

Vorstufe programmiert. Leider erwies sich der eingeschlagene Weg als Sackgasse. Die Verfahren funktionieren zwar, jedoch liefern die Versionen mit Integralmitteldefekt auf nicht äquidistanten Gittern nur die Ordnungsfolgen 2,3,... Eine Analyse, warum die Konvergenz auf nicht äquidistanten Gittern langsamer ist, wurde im Zuge der vorliegenden Arbeit nicht durchgeführt, es wird auf die experimentellen Resultate der folgenden Abschnitte verwiesen.

Die einzigen funktionierenden superkonvergenten Verfahren für explizite Anfangswertprobleme 2. Ordnung sind die später entwickelten und zuvor in den Abschnitten 3.3.1 und 3.3.2 vorgestellten Verfahren IPDeC2-DQ2 und IPDeC2-DQ2-PC.

Für die Lösung von (3.1) erfolgt eine Rückführung auf ein System von Differentialgleichungen 1. Ordnung

$$u'(t) = v(t), \quad (3.29a)$$

$$v'(t) = f(t, u(t), v(t)), \text{ mit} \quad (3.29b)$$

$$u(t_0) = u_0, u'(t_0) = v(t_0) = v_0. \quad (3.29c)$$

Die Diskretisierung von (3.29) erfolgt in den folgenden Abschnitten durch verschiedene Basisverfahren.

3.4.1 IDeC2-Heun

Wendet man das Verfahren von Heun angeschrieben als 1-stufiges Runge-Kutta-Verfahren auf (3.29) an, erhält man

$$W_1^{[0]} = w_{j,l-1}^{[0]} + h_{j,l} x_{j,l-1}^{[0]}, \quad (3.30a)$$

$$X_1^{[0]} = x_{j,l-1}^{[0]} + h_{j,l} f(t_{j,l-1}, w_{j,l-1}^{[0]}, x_{j,l-1}^{[0]}), \quad (3.30b)$$

$$w_{j,l}^{[0]} = w_{j,l-1}^{[0]} + \frac{h_{j,l}}{2} [x_{j,l-1}^{[0]} + X_1^{[0]}], \quad (3.30c)$$

$$x_{j,l}^{[0]} = x_{j,l-1}^{[0]} + \frac{h_{j,l}}{2} [f(t_{j,l-1}, w_{j,l-1}^{[0]}, x_{j,l-1}^{[0]}) + f(t_{j,l}, W_1^{[0]}, X_1^{[0]})], \quad (3.30d)$$

$$\text{mit } w_{1,0}^{[0]} = u_0, x_{1,0}^{[0]} = v_0. \quad (3.30e)$$

h	Heun	IDeC21	IDeC22	IDeC23	COLL
0,1	1,73E-04	1,87E-08	1,73E-13	8,77E-13	2,10E-12
0,05	4,32E-05	1,18E-09	2,99E-15	1,37E-14	3,26E-14
0,025	1,08E-05	7,37E-11	4,89E-17	2,14E-16	5,08E-16
0,0125	2,71E-06	4,61E-12	7,81E-19	3,35E-18	7,92E-18
0,1	2,00	3,99	5,85	6,00	6,01
0,05	2,00	4,00	5,93	6,00	6,00
0,025	2,00	4,00	5,97	6,00	6,00
0,0125	2,00	4,00	5,97	6,00	6,00

Tabelle 3.11: Resultate zu Beispiel 3.22.

Interpoliert man die $w_{j,l}^{[\nu]}$ stückweise mit $p^{w,[\nu]}(t)$ und die $x_{j,l}^{[\nu]}$ stückweise mit $p^{x,[\nu]}(t)$, kann man die Defekte mit $p_{j,l}^{\#, [\nu]} \equiv p^{\#, [\nu]}(t_{j,l})$, $\# = w, x$ wie folgt definieren

$$d_{j,l}^{w,[\nu]} = p_{j,l}^{w,[\nu]'} - p_{j,l}^{x,[\nu]} = p_{j,l}^{w,[\nu]'} - x_{j,l}^{[\nu]}, \quad (3.31a)$$

$$d_{j,l}^{x,[\nu]} = p_{j,l}^{x,[\nu]'} - f\left(t_{j,l}, p_{j,l}^{w,[\nu]}, p_{j,l}^{x,[\nu]}\right) = p_{j,l}^{x,[\nu]'} - f\left(t_{j,l}, w_{j,l}^{[\nu]}, x_{j,l}^{[\nu]}\right). \quad (3.31b)$$

Setzt man diese in das zugehörige Nachbarproblem ein, erhält man

$$Y_1^{[\nu]} = y_{j,l-1}^{[\nu]} + h_{j,l} \left[z_{j,l-1}^{[\nu]} + d_{j,l-1}^{w,[\nu]} \right], \quad (3.32a)$$

$$Z_1^{[\nu]} = z_{j,l-1}^{[\nu]} + h_{j,l} \left[f\left(t_{j,l-1}, y_{j,l-1}^{[\nu]}, z_{j,l-1}^{[\nu]}\right) + d_{j,l-1}^{x,[\nu]} \right], \quad (3.32b)$$

$$y_{j,l}^{[\nu]} = y_{j,l-1}^{[\nu]} + \frac{h_{j,l}}{2} \left[z_{j,l-1}^{[\nu]} + Z_1^{[\nu]} + d_{j,l-1}^{w,[\nu]} + d_{j,l}^{w,[\nu]} \right], \quad (3.32c)$$

$$z_{j,l}^{[\nu]} = z_{j,l-1}^{[\nu]} + \frac{h_{j,l}}{2} \left[f\left(t_{j,l-1}, y_{j,l-1}^{[\nu]}, z_{j,l-1}^{[\nu]}\right) + f\left(t_{j,l}, Y_1^{[\nu]}, Z_1^{[\nu]}\right) \right] \\ + \frac{h_{j,l}}{2} \left[d_{j,l-1}^{x,[\nu]} + d_{j,l}^{x,[\nu]} \right], \quad (3.32d)$$

$$\text{mit } y_{1,0}^{[\nu]} = u_0, \quad z_{1,0}^{[\nu]} = v_0, \quad (3.32e)$$

wobei die Lösungen durch Iteration analog zu (1.17) wie folgt verbessert werden

$$w_{j,l}^{[\nu]} = w_{j,l}^{[0]} - \left(y_{j,l}^{[\nu-1]} - w_{j,l}^{[\nu-1]} \right), \quad (3.33a)$$

$$x_{j,l}^{[\nu]} = x_{j,l}^{[0]} - \left(z_{j,l}^{[\nu-1]} - x_{j,l}^{[\nu-1]} \right). \quad (3.33b)$$

Beispiel 3.22. Wie Beispiel 3.3, jedoch mit dem Verfahren von Heun als Basisverfahren. Tabelle 3.11 zeigt die Resultate.

3.4.2 IDeC2-EMR

Analog zu IDeC2-Heun aus Abschnitt 3.4.1 folgt für die Anwendung der expliziten Mittelpunktsregel auf (3.29) unter Beibehaltung sämtlicher anderer algorithmischer Details

$$W_{\frac{1}{2}}^{[0]} = w_{j,l-1}^{[0]} + \frac{h_{j,l}}{2} x_{j,l-1}^{[0]}, \quad (3.34a)$$

$$X_{\frac{1}{2}}^{[0]} = x_{j,l-1}^{[0]} + \frac{h_{j,l}}{2} f\left(t_{j,l-1}, w_{j,l-1}^{[0]}, x_{j,l-1}^{[0]}\right), \quad (3.34b)$$

$$w_{j,l}^{[0]} = w_{j,l-1}^{[0]} + h_{j,l} X_{\frac{1}{2}}^{[0]}, \quad (3.34c)$$

$$x_{j,l}^{[0]} = x_{j,l-1}^{[0]} + h_{j,l} f\left(t_{j,l} + \frac{h_{j,l}}{2}, W_{\frac{1}{2}}^{[0]}, X_{\frac{1}{2}}^{[0]}\right). \quad (3.34d)$$

Mit den Defekten (3.31) und

$$\begin{aligned} d_{\frac{1}{2}}^{w,[\nu]} &\equiv d^{w,[\nu]}\left(t_{j,l-1} + \frac{h_{j,l}}{2}\right) \\ &= p^{w,[\nu]'}\left(t_{j,l-1} + \frac{h_{j,l}}{2}\right) - p^{x,[\nu]}\left(t_{j,l-1} + \frac{h_{j,l}}{2}\right), \end{aligned} \quad (3.35a)$$

$$\begin{aligned} d_{\frac{1}{2}}^{x,[\nu]} &\equiv d^{x,[\nu]}\left(t_{j,l-1} + \frac{h_{j,l}}{2}\right) = p^{x,[\nu]'}\left(t_{j,l-1} + \frac{h_{j,l}}{2}\right) \\ &\quad - f\left(t_{j,l-1} + \frac{h_{j,l}}{2}, p^{w,[\nu]}\left(t_{j,l-1} + \frac{h_{j,l}}{2}\right), p^{x,[\nu]}\left(t_{j,l-1} + \frac{h_{j,l}}{2}\right)\right), \end{aligned} \quad (3.35b)$$

gelangt man zum IDeC2-EMR-Nachbarproblem

$$Y_{\frac{1}{2}}^{[0]} = y_{j,l-1}^{[0]} + \frac{h_{j,l}}{2} \left[z_{j,l-1}^{[0]} + d_{j,l-1}^{w,[\nu]} \right], \quad (3.36a)$$

$$Z_{\frac{1}{2}}^{[0]} = z_{j,l-1}^{[0]} + \frac{h_{j,l}}{2} \left[f\left(t_{j,l-1}, y_{j,l-1}^{[0]}, z_{j,l-1}^{[0]}\right) + d_{j,l-1}^{x,[\nu]} \right], \quad (3.36b)$$

$$y_{j,l}^{[0]} = y_{j,l-1}^{[0]} + h_{j,l} \left[Z_{\frac{1}{2}}^{[0]} + d_{\frac{1}{2}}^{w,[\nu]} \right], \quad (3.36c)$$

$$z_{j,l}^{[0]} = z_{j,l-1}^{[0]} + h_{j,l} \left[f\left(t_{j,l} + \frac{h_{j,l}}{2}, Y_{\frac{1}{2}}^{[0]}, Z_{\frac{1}{2}}^{[0]}\right) + d_{\frac{1}{2}}^{x,[\nu]} \right]. \quad (3.36d)$$

Beispiel 3.23. Wie Beispiel 3.3, jedoch mit der expliziten Mittelpunktsregel als Basisverfahren. Tabelle 3.12 zeigt die Resultate.

3.4.3 IDeC2-Taylor2

Dass Taylor-Verfahren zur Lösung (großer) Differentialgleichungssysteme aufgrund des vergleichsweise großen Rechenaufwands nicht geeignet sind, wurde schon bei den Verfahren für Anfangswertprobleme 1. Ordnung angemerkt, trotzdem sei die Ausführung des Verfahrens für Anfangswertprobleme 2. Ordnung dokumentiert. Erweitert man den in Abschnitt 1.2.12 gezeigten IDeC-Taylor-Algorithmus 2. Ordnung auf Anfangswertprobleme 2. Ordnung bei Rückführung auf ein System 1. Ordnung, erhält man mit $u \equiv u(t)$, $v \equiv v(t)$ und $f \equiv f(t, u, v)$ aus (3.29)

$$u' = v, \quad (3.37a)$$

$$u'' = v' = f, \quad (3.37b)$$

$$u''' = v'' = f_t + f_u u' + f_v v' = f_t + f_u v + f_v f \quad (3.37c)$$

h	EMR	IDeC21	IDeC22	IDeC23	COLL
0,1	1,76E-04	2,31E-08	2,76E-12	8,50E-13	2,10E-12
0,05	4,40E-05	1,44E-09	4,25E-14	1,33E-14	3,26E-14
0,025	1,10E-05	9,02E-11	6,59E-16	2,09E-16	5,08E-16
0,0125	2,75E-06	5,64E-12	1,03E-17	3,27E-18	7,92E-18
0,1	2,00	4,00	6,02	5,99	6,01
0,05	2,00	4,00	6,01	6,00	6,00
0,025	2,00	4,00	6,01	6,00	6,00
0,0125	2,00	4,00	6,01	6,00	6,00

Tabelle 3.12: Resultate zu Beispiel 3.23.

folgendes Basisverfahren

$$w_{j,l}^{[0]} = w_{j,l-1}^{[0]} + h_{j,l} x_{j,l-1}^{[0]} + \frac{h_{j,l}^2}{2} f \left(t_{j,l-1}, w_{j,l-1}^{[0]}, x_{j,l-1}^{[0]} \right), \quad (3.38a)$$

$$\begin{aligned} x_{j,l}^{[0]} &= x_{j,l-1}^{[0]} + h_{j,l} f \left(t_{j,l-1}, w_{j,l-1}^{[0]}, x_{j,l-1}^{[0]} \right) \\ &+ \frac{h_{j,l}^2}{2} \left[f_t \left(t_{j,l-1}, w_{j,l-1}^{[0]}, x_{j,l-1}^{[0]} \right) + f_u \left(t_{j,l-1}, w_{j,l-1}^{[0]}, x_{j,l-1}^{[0]} \right) x_{j,l-1}^{[0]} \right. \\ &\left. + f_v \left(t_{j,l-1}, w_{j,l-1}^{[0]}, x_{j,l-1}^{[0]} \right) f \left(t_{j,l-1}, w_{j,l-1}^{[0]}, x_{j,l-1}^{[0]} \right) \right]. \end{aligned} \quad (3.38b)$$

Man benötigt insgesamt 4 Defektanteile

$$d_{j,l-1}^{w,A,[\nu]} = p_{j,l-1}^{w,[\nu]'} - p_{j,l-1}^{x,[\nu]} = p_{j,l-1}^{w,[\nu]'} - x_{j,l-1}^{[\nu]} \quad (3.39a)$$

$$\begin{aligned} d_{j,l-1}^{w,B,[\nu]} &= p_{j,l-1}^{w,[\nu]''} - f\left(t_{j,l-1}, p_{j,l-1}^{w,[\nu]}, p_{j,l-1}^{x,[\nu]}\right) \\ &= p_{j,l-1}^{w,[\nu]''} - f\left(t_{j,l-1}, w_{j,l-1}^{[\nu]}, x_{j,l-1}^{[\nu]}\right), \end{aligned} \quad (3.39b)$$

$$\begin{aligned} d_{j,l-1}^{x,A,[\nu]} &= p_{j,l-1}^{x,[\nu]'} - f\left(t_{j,l-1}, p_{j,l-1}^{w,[\nu]}, p_{j,l-1}^{x,[\nu]}\right) \\ &= p_{j,l-1}^{x,[\nu]'} - f\left(t_{j,l-1}, w_{j,l-1}^{[\nu]}, x_{j,l-1}^{[\nu]}\right), \end{aligned} \quad (3.39c)$$

$$\begin{aligned} d_{j,l-1}^{x,B,[\nu]} &= p_{j,l-1}^{x,[\nu]''} - \left[f_t\left(t_{j,l-1}, p_{j,l-1}^{w,[\nu]}, p_{j,l-1}^{x,[\nu]}\right) \right. \\ &\quad + f_u\left(t_{j,l-1}, p_{j,l-1}^{w,[\nu]}, p_{j,l-1}^{x,[\nu]}\right) p_{j,l-1}^{x,[\nu]} \\ &\quad \left. + f_v\left(t_{j,l-1}, p_{j,l-1}^{w,[\nu]}, p_{j,l-1}^{x,[\nu]}\right) f\left(t_{j,l-1}, p_{j,l-1}^{w,[\nu]}, p_{j,l-1}^{x,[\nu]}\right) \right] \\ &= p_{j,l-1}^{x,[\nu]''} - \left[f_t\left(t_{j,l-1}, w_{j,l-1}^{[\nu]}, x_{j,l-1}^{[\nu]}\right) \right. \\ &\quad + f_u\left(t_{j,l-1}, w_{j,l-1}^{[\nu]}, x_{j,l-1}^{[\nu]}\right) x_{j,l-1}^{[\nu]} \\ &\quad \left. + f_v\left(t_{j,l-1}, w_{j,l-1}^{[\nu]}, x_{j,l-1}^{[\nu]}\right) f\left(t_{j,l-1}, w_{j,l-1}^{[\nu]}, x_{j,l-1}^{[\nu]}\right) \right]. \end{aligned} \quad (3.39d)$$

Setzt man diese in das zugehörige Nachbarproblem ein, erhält man

$$\begin{aligned} y_{j,l}^{[\nu]} &= y_{j,l-1}^{[\nu]} + h_{j,l} z_{j,l-1}^{[\nu]} + \frac{h_{j,l}^2}{2} f\left(t_{j,l-1}, y_{j,l-1}^{[\nu]}, z_{j,l-1}^{[\nu]}\right) \\ &\quad + h_{j,l} d_{j,l-1}^{w,A,[\nu]} + \frac{h_{j,l}^2}{2} d_{j,l-1}^{w,B,[\nu]}, \end{aligned} \quad (3.40a)$$

$$\begin{aligned} z_{j,l}^{[\nu]} &= z_{j,l-1}^{[\nu]} + h_{j,l} f\left(t_{j,l-1}, y_{j,l-1}^{[\nu]}, z_{j,l-1}^{[\nu]}\right) \\ &\quad + \frac{h_{j,l}^2}{2} \left[f_t\left(t_{j,l-1}, y_{j,l-1}^{[\nu]}, z_{j,l-1}^{[\nu]}\right) + f_u\left(t_{j,l-1}, y_{j,l-1}^{[\nu]}, z_{j,l-1}^{[\nu]}\right) z_{j,l-1}^{[\nu]} \right. \\ &\quad \left. + f_v\left(t_{j,l-1}, y_{j,l-1}^{[\nu]}, z_{j,l-1}^{[\nu]}\right) f\left(t_{j,l-1}, y_{j,l-1}^{[\nu]}, z_{j,l-1}^{[\nu]}\right) \right] \\ &\quad + h_{j,l} d_{j,l-1}^{x,A,[\nu]} + \frac{h_{j,l}^2}{2} d_{j,l-1}^{x,B,[\nu]}. \end{aligned} \quad (3.40b)$$

Das Verfahren liefert die Ordnungsfolge $2, 4, \dots$

Beispiel 3.24. Wie Beispiel 3.3, jedoch mit dem Taylor-Verfahren 2. Ordnung als Basisverfahren. Tabelle 3.13 zeigt die Resultate.

h	Taylor2	IDeC21	IDeC22	IDeC23	COLL
0,1	1,79E-04	2,64E-08	9,14E-12	6,62E-12	2,10E-12
0,05	4,48E-05	1,65E-09	1,41E-13	1,03E-13	3,26E-14
0,025	1,12E-05	1,03E-10	2,18E-15	1,62E-15	5,08E-16
0,0125	2,80E-06	6,42E-12	3,39E-17	2,52E-17	7,92E-18
0,1	2,00	4,00	6,02	6,00	6,01
0,05	2,00	4,00	6,01	6,00	6,00
0,025	2,00	4,00	6,01	6,00	6,00
0,0125	2,00	4,00	6,01	6,00	6,00

Tabelle 3.13: Resultate zu Beispiel 3.24.

3.4.4 IDeC2-ITR

Wendet man die implizite Trapezregel auf (3.29) an, ergeben sich folgende Differenzgleichungen

$$\frac{w_{j,l}^{[0]} - w_{j,l-1}^{[0]}}{h_{j,l}} = \frac{1}{2} \left[x_{j,l-1}^{[0]} + x_{j,l}^{[0]} \right], \quad (3.41a)$$

$$\frac{x_{j,l}^{[0]} - x_{j,l-1}^{[0]}}{h_{j,l}} = \frac{1}{2} \left[f \left(t_{j,l-1}, w_{j,l-1}^{[0]}, x_{j,l-1}^{[0]} \right) + f \left(t_{j,l}, w_{j,l}^{[0]}, x_{j,l}^{[0]} \right) \right]. \quad (3.41b)$$

Mit den Defekten

$$d_{j,l}^{w,[\nu]} = p_{j,l}^{w,[\nu]'} - p_{j,l}^{x,[\nu]} = p_{j,l}^{w,[\nu]'} - x_{j,l}^{[\nu]}, \quad (3.42a)$$

$$d_{j,l}^{x,[\nu]} = p_{j,l}^{x,[\nu]'} - f \left(t_{j,l}, p_{j,l}^{w,[\nu]}, p_{j,l}^{x,[\nu]} \right) = p_{j,l}^{x,[\nu]'} - f \left(t_{j,l}, w_{j,l}^{[\nu]}, x_{j,l}^{[\nu]} \right) \quad (3.42b)$$

gelangt man zum Nachbarproblem

$$\frac{y_{j,l}^{[\nu]} - y_{j,l-1}^{[\nu]}}{h_{j,l}} = \frac{1}{2} \left[z_{j,l-1}^{[\nu]} + z_{j,l}^{[\nu]} + d_{j,l-1}^{w,[\nu]} + d_{j,l}^{w,[\nu]} \right], \quad (3.43a)$$

$$\frac{z_{j,l}^{[\nu]} - z_{j,l-1}^{[\nu]}}{h_{j,l}} = \frac{1}{2} \left[f \left(t_{j,l-1}, y_{j,l-1}^{[\nu]}, z_{j,l-1}^{[\nu]} \right) + f \left(t_{j,l}, y_{j,l}^{[\nu]}, z_{j,l}^{[\nu]} \right) + d_{j,l-1}^{x,[\nu]} + d_{j,l}^{x,[\nu]} \right]. \quad (3.43b)$$

Beispiel 3.25. Wie Beispiel 3.3, jedoch mit der impliziten Trapezregel als Basisverfahren. Tabelle 3.14 zeigt die Resultate.

3.4.5 IQDeC2-Heun

Kombiniert man das Basisverfahren (3.30) mit den Defekten

$$\bar{d}_{j,l}^{w,[\nu]} = \frac{w_{j,l}^{[\nu]} - w_{j,l-1}^{[\nu]}}{h_{j,l}} - \sum_{\mu=1}^m \alpha_{l,\mu} x_{j,\mu}^{[\nu]}, \quad (3.44a)$$

$$\bar{d}_{j,l}^{x,[\nu]} = \frac{x_{j,l}^{[\nu]} - x_{j,l-1}^{[\nu]}}{h_{j,l}} - \sum_{\mu=1}^m \alpha_{l,\mu} f \left(t_{j,\mu}, w_{j,\mu}^{[\nu]}, x_{j,\mu}^{[\nu]} \right) \quad (3.44b)$$

h	ITR	IDeC21	IDeC22	IDeC23	COLL
0,1	8,97E-05	6,57E-09	1,15E-12	8,76E-13	2,10E-12
0,05	2,24E-05	4,11E-10	1,81E-14	1,37E-14	3,26E-14
0,025	5,60E-06	2,57E-11	2,82E-16	2,14E-16	5,08E-16
0,0125	1,40E-06	1,60E-12	4,41E-18	3,35E-18	7,92E-18
0,1	2,00	4,00	6,00	6,00	6,01
0,05	2,00	4,00	6,00	6,00	6,00
0,025	2,00	4,00	6,00	6,00	6,00
0,0125	2,00	4,00	6,00	6,00	6,00

Tabelle 3.14: Resultate zu Beispiel 3.25.

h	Heun	IQDeC21	IQDeC22	IQDeC23	COLL
0,1	1,73E-04	1,06E-08	5,20E-12	2,10E-12	2,10E-12
0,05	4,32E-05	6,65E-10	8,07E-14	3,26E-14	3,26E-14
0,025	1,08E-05	4,15E-11	1,25E-15	5,08E-16	5,08E-16
0,0125	2,71E-06	2,59E-12	1,96E-17	7,92E-18	7,92E-18
0,1	2,00	4,00	6,00	6,00	6,01
0,05	2,00	4,00	6,00	6,00	6,00
0,025	2,00	4,00	6,00	6,00	6,00
0,0125	2,00	4,00	6,00	6,00	6,00

Tabelle 3.15: Resultate zu Beispiel 3.26.

unter Berücksichtigung von (1.117) für die Quadraturkoeffizienten $\alpha_{l,\mu}$, so gelangt man zum zugehörigen IQDeC2-Heun-Nachbarproblem

$$Y_1^{[\nu]} = y_{j,l-1}^{[\nu]} + h_{j,l} \left[z_{j,l-1}^{[\nu]} + \bar{d}_{j,l}^{w,[\nu]} \right], \quad (3.45a)$$

$$Z_1^{[\nu]} = z_{j,l-1}^{[\nu]} + h_{j,l} \left[f \left(t_{j,l-1}, y_{j,l-1}^{[\nu]}, z_{j,l-1}^{[\nu]} \right) + \bar{d}_{j,l}^{x,[\nu]} \right], \quad (3.45b)$$

$$y_{j,l}^{[\nu]} = y_{j,l-1}^{[\nu]} + \frac{h_{j,l}}{2} \left[z_{j,l-1}^{[\nu]} + Z_1^{[\nu]} + 2\bar{d}_{j,l}^{w,[\nu]} \right], \quad (3.45c)$$

$$z_{j,l}^{[\nu]} = z_{j,l-1}^{[\nu]} + \frac{h_{j,l}}{2} \left[f \left(t_{j,l-1}, y_{j,l-1}^{[\nu]}, z_{j,l-1}^{[\nu]} \right) + f \left(t_{j,l}, Y_1^{[\nu]}, Z_1^{[\nu]} \right) + 2\bar{d}_{j,l}^{x,[\nu]} \right]. \quad (3.45d)$$

Beispiel 3.26. Wie Beispiel 3.3, jedoch mit dem Verfahren von Heun als Basisverfahren und Integralmitteldefekt. Tabelle 3.15 zeigt die Resultate.

Beispiel 3.27. Wie Beispiel 3.3, jedoch mit dem Verfahren von Heun als Basisverfahren, Integralmitteldefekt und 4 nicht äquidistanten Kollokationsknoten ($c_1 = 0,0877$; $c_2 = 0,3339$; $c_3 = 0,7899$; $c_4 = 1$). Tabelle 3.16 zeigt die Resultate.

h	Heun	IQDeC21	IQDeC22	IQDeC23	COLL
0,1	7,37E-04	1,88E-07	3,05E-08	3,46E-08	1,77E-08
0,05	1,86E-04	3,60E-09	1,90E-09	2,16E-09	1,11E-09
0,025	4,66E-05	8,14E-10	1,18E-10	1,35E-10	6,91E-11
0,0125	1,17E-05	1,81E-10	7,39E-12	8,43E-12	4,31E-12
0,1	1,99	5,71	4,00	4,00	4,00
0,05	1,99	2,15	4,00	4,00	4,00
0,025	2,00	2,16	4,00	4,00	4,00
0,0125					

Tabelle 3.16: Resultate zu Beispiel 3.27.

h	EMR	IQDeC21	IQDeC22	IQDeC23	COLL
0,1	1,76E-04	8,32E-09	4,68E-12	4,11E-12	2,10E-12
0,05	4,40E-05	5,16E-10	7,25E-14	6,38E-14	3,26E-14
0,025	1,10E-05	3,22E-11	1,13E-15	9,93E-16	5,08E-16
0,0125	2,75E-06	2,01E-12	1,76E-17	1,55E-17	7,92E-18
0,1	2,00	4,01	6,01	6,01	6,01
0,05	2,00	4,00	6,01	6,00	6,00
0,025	2,00	4,00	6,00	6,00	6,00
0,0125					

Tabelle 3.17: Resultate zu Beispiel 3.28.

3.4.6 IQDeC2-EMR

Das EMR-Basisverfahren (3.34), die Defekte (3.44) und das zugehörige Nachbarproblem

$$Y_{\frac{1}{2}}^{[\nu]} = y_{j,l-1}^{[\nu]} + \frac{h_{j,l}}{2} \left[z_{j,l-1}^{[\nu]} + \bar{d}_{j,l}^{w,[\nu]} \right], \quad (3.46a)$$

$$Z_{\frac{1}{2}}^{[\nu]} = z_{j,l-1}^{[\nu]} + \frac{h_{j,l}}{2} \left[f \left(t_{j,l-1}, y_{j,l-1}^{[\nu]}, z_{j,l-1}^{[\nu]} \right) + \bar{d}_{j,l}^{x,[\nu]} \right], \quad (3.46b)$$

$$y_{j,l}^{[\nu]} = y_{j,l-1}^{[\nu]} + h_{j,l} \left[Z_{\frac{1}{2}}^{[\nu]} + \bar{d}_{j,l}^{w,[\nu]} \right], \quad (3.46c)$$

$$z_{j,l}^{[\nu]} = z_{j,l-1}^{[\nu]} + h_{j,l} \left[f \left(t_{j,l} + \frac{h_{j,l}}{2}, Y_{\frac{1}{2}}^{[\nu]}, Z_{\frac{1}{2}}^{[\nu]} \right) + \bar{d}_{j,l}^{x,[\nu]} \right] \quad (3.46d)$$

bilden das IQDeC2-EMR-Verfahren.

Beispiel 3.28. Wie Beispiel 3.3, jedoch mit der expliziten Mittelpunktsregel als Basisverfahren und Integralmitteldefekt. Tabelle 3.17 zeigt die Resultate.

Beispiel 3.29. Wie Beispiel 3.3, jedoch mit der expliziten Mittelpunktsregel als Basisverfahren, Integralmitteldefekt und 4 nicht äquidistanten Kollokationsknoten ($c_1 = 0,0877$; $c_2 = 0,3339$; $c_3 = 0,7899$; $c_4 = 1$). Tabelle 3.18 zeigt die Resultate.

h	EMR	IQDeC21	IQDeC22	IQDeC23	COLL
0,1	7,52E-04	1,79E-07	3,01E-08	3,46E-08	1,77E-08
0,05	1,89E-04	4,25E-09	1,88E-09	2,16E-09	1,11E-09
0,025	4,74E-05	5,99E-10	1,17E-10	1,35E-10	6,91E-11
0,0125	1,19E-05	1,45E-10	7,32E-12	8,43E-12	4,31E-12
0,1	1,99	5,40	4,00	4,00	4,00
0,05	2,00	2,82	4,00	4,00	4,00
0,025	2,00	2,04	4,00	4,00	4,00
0,0125	2,00	2,04	4,00	4,00	4,00

Tabelle 3.18: Resultate zu Beispiel 3.29.

h	ITR	IQDeC21	IQDeC22	IQDeC23	COLL
0,1	8,97E-05	4,49E-09	4,02E-12	4,11E-12	2,10E-12
0,05	2,24E-05	2,80E-10	6,23E-14	6,38E-14	3,26E-14
0,025	5,60E-06	1,76E-11	9,70E-16	9,93E-16	5,08E-16
0,0125	1,40E-06	1,10E-12	1,51E-17	1,55E-17	7,92E-18
0,1	2,00	4,00	6,01	6,01	6,01
0,05	2,00	4,00	6,01	6,01	6,00
0,025	2,00	4,00	6,00	6,00	6,00
0,0125	2,00	4,00	6,00	6,00	6,00

Tabelle 3.19: Resultate zu Beispiel 3.30.

3.4.7 IQDeC2-ITR

Die implizite Trapezregel (3.41), die Defekte (3.44) und das Nachbarproblem

$$\frac{y_{j,l}^{[\nu]} - y_{j,l-1}^{[\nu]}}{h_{j,l}} = \frac{1}{2} \left[z_{j,l-1}^{[\nu]} + z_{j,l}^{[\nu]} \right] + \bar{d}_{j,l}^{w,[\nu]}, \quad (3.47a)$$

$$\frac{z_{j,l}^{[\nu]} - z_{j,l-1}^{[\nu]}}{h_{j,l}} = \frac{1}{2} \left[f \left(t_{j,l-1}, y_{j,l-1}^{[\nu]}, z_{j,l-1}^{[\nu]} \right) + f \left(t_{j,l}, y_{j,l}^{[\nu]}, z_{j,l}^{[\nu]} \right) \right] + \bar{d}_{j,l}^{x,[\nu]} \quad (3.47b)$$

ergeben den IQDeC2-ITR-Algorithmus.

Beispiel 3.30. Wie Beispiel 3.3, jedoch mit der impliziten Trapezregel als Basisverfahren und Integralmitteldefekt. Tabelle 3.19 zeigt die Resultate.

Beispiel 3.31. Wie Beispiel 3.3, jedoch mit der impliziten Trapezregel als Basisverfahren, Integralmitteldefekt und 4 nicht äquidistanten Kollokationsknoten ($c_1 = 0,0877$; $c_2 = 0,3339$; $c_3 = 0,7899$; $c_4 = 1$). Tabelle 3.20 zeigt die Resultate.

h	ITR	IQDeC21	IQDeC22	IQDeC23	COLL
0,1	3,86E-04	1,59E-07	3,70E-08	3,46E-08	1,77E-08
0,05	9,65E-05	1,27E-08	2,31E-09	2,16E-09	1,11E-09
0,025	2,41E-05	1,14E-09	1,44E-10	1,35E-10	6,91E-11
0,0125	6,04E-06	1,14E-10	9,02E-12	8,43E-12	4,31E-12
0,1	2,00	3,65	4,00	4,00	4,00
0,05	2,00	3,48	4,00	4,00	4,00
0,025	2,00	3,32	4,00	4,00	4,00
0,0125	2,00	3,32	4,00	4,00	4,00

Tabelle 3.20: Resultate zu Beispiel 3.31.

Kapitel 4

Implizite Differentialgleichungen zweiter Ordnung

Im Folgenden werden Defektkorrekturverfahren zur direkten Lösung von n -dimensionalen impliziten Anfangswertproblemen 2. Ordnung der Bauart

$$F(t, u(t), u'(t), u''(t)) = 0, \quad (4.1a)$$

$$u(t_0) = u_0, \quad u'(t_0) = v_0, \quad (4.1b)$$

vorgestellt. Den Anfang bilden zwei Verfahren (eines davon superkonvergent) ohne Rückführung auf ein System 1. Ordnung, den Abschluss bildet ein Verfahren mit Rückführung auf ein System 1. Ordnung und Integralmitteldefekt.

4.1 Klassischer Defekt – IIDeC2-DQ2

Der Algorithmus ist eine Umformulierung des IIDeC2-DQ2-Verfahrens aus Abschnitt 3.1.1. Für das Basisverfahren gilt auf einem äquidistanten Gitter ($h \equiv h_{j,l}$) gemäß Definition 1.1

$$\begin{cases} F \left(t_{1,0}, u_0, v_0, \frac{x_{1,1}^{[0]} - u_0 - v_0}{\frac{h}{2}} \right) = 0, & j = 1, l = 1 \\ F \left(t_{j,l}, x_{j,l}^{[0]}, \frac{x_{j,l+1}^{[0]} - x_{j,l-1}^{[0]}}{2h}, \frac{x_{j,l+1}^{[0]} - 2x_{j,l}^{[0]} + x_{j,l-1}^{[0]}}{h^2} \right) = 0, & j \geq 1, l > 1 \\ F \left(t_{j-1,m}, x_{j-1,m}^{[0]}, \frac{x_{j,1}^{[0]} - x_{j-1,m-1}^{[0]}}{2h}, \frac{x_{j,1}^{[0]} - 2x_{j-1,m}^{[0]} + x_{j-1,m-1}^{[0]}}{h^2} \right) = 0, & j > 1, l = 1. \end{cases} \quad (4.2)$$

Die Lösungen der Differenzgleichungen des Basisverfahrens liefern eine Folge von $x_{j,l}^{[0]}$, $j = 1, \dots, N$, $l = 0, \dots, m$ mit $x_{j,0}^{[0]} = x_{j-1,m}^{[0]}$ für $j > 1$. Interpoliert man diese Lösungen bzw. in weiterer Folge die verbesserten Lösungen $x_{j,l}^{[\nu]}$, $\nu = 0, 1, \dots$ durch die stückweise definierten Polynome $p_{j,l}^{[\nu]} \equiv p^{[\nu]}(t_{j,l})$ vom Grad $\leq m - 1$, so berechnen sich die Defekte wie folgt

$$\begin{cases} D_{1,0}^{A,[\nu]} = F \left(t_{1,0}, u_0, v_0, p_{1,0}^{[\nu]''} \right), \quad D_{1,0}^{B,[\nu]} = p_{1,0}^{[\nu]'} - v_0, & j = 1, l = 0 \\ D_{j,l}^{[\nu]} = F \left(t_{j,l}, p_{j,l}^{[\nu]}, p_{j,l}^{[\nu]'}, p_{j,l}^{[\nu]''} \right), & j \geq 1, l > 0 \\ D_{j,0}^{[\nu]} = F \left(t_{j,l}, p_{j,0}^{[\nu]}, \frac{p_{j,0}^{[\nu]'} + p_{j-1,m}^{[\nu]'}}{2}, \frac{p_{j,0}^{[\nu]''} + p_{j-1,m}^{[\nu]''}}{2} + \frac{p_{j,0}^{[\nu]'} - p_{j-1,m}^{[\nu]'}}{h} \right), & j > 1, l = 0. \end{cases} \quad (4.3)$$

h	DQ2	IIDeC21	IIDeC22	IIDeC23	COLL
0,1	2,30E-05	1,93E-09	1,11E-12	1,01E-12	3,53E-13
0,05	5,75E-05	1,21E-10	1,74E-14	1,58E-14	5,64E-15
0,025	1,44E-06	7,54E-12	2,72E-16	2,47E-16	8,92E-17
0,0125	3,59E-07	4,71E-13	4,25E-18	3,86E-18	1,40E-18
0,1	2,00	4,00	6,00	6,00	5,97
0,05	2,00	4,00	6,00	6,00	5,98
0,025	2,00	4,00	6,00	6,00	5,99
0,0125	2,00	4,00	6,00	6,00	5,99

Tabelle 4.1: Resultate zu Beispiel 4.3.

Bemerkung 4.1. Der Defektanteil $D_{1,0}^{B,[\nu]}$ beschreibt analog zu $d_{1,0}^{B,[\nu]}$ bei IIDeC2-DQ2 (Abschnitt 3.1.1) den Defekt bezüglich der Ableitungsanfangsbedingung.

Bemerkung 4.2. Man beachte beim Sprungdefekt $D_{j,0}^{[\nu]}$ die zwingend erforderliche Mittelung der 1. Ableitungen an den Intervallenden, welche beim expliziten Verfahren IIDeC2-DQ2 nicht erforderlich ist (siehe (3.4)).

Das zugehörige Nachbarproblem lautet

$$\begin{cases}
F\left(t_{1,0}, u_0, v_0 + D_{1,0}^{B,[\nu]}, \frac{z_{1,1}^{[\nu]} - u_0}{h} - \left(v_0 + D_{1,0}^{B,[\nu]}\right) \frac{1}{h}\right) - D_{1,0}^{A,[\nu]} = 0, & j = 1, l = 1 \\
F\left(t_{j,l}, z_{j,l}^{[\nu]}, \frac{z_{j,l+1}^{[\nu]} - z_{j,l}^{[\nu]}}{2h}, \frac{z_{j,l+1}^{[\nu]} - 2z_{j,l}^{[\nu]} + z_{j,l-1}^{[\nu]}}{h^2}\right) - D_{j,l}^{[\nu]} = 0, & j \geq 1, l > 1 \\
F\left(t_{j-1,m}, z_{j-1,m}^{[\nu]}, \frac{z_{j,1}^{[\nu]} - z_{j-1,m-1}^{[\nu]}}{2h}, \frac{z_{j,1}^{[\nu]} - 2z_{j-1,m}^{[\nu]} + z_{j-1,m-1}^{[\nu]}}{h^2}\right) - D_{j,0}^{[\nu]} = 0, & j > 1, l = 1,
\end{cases} \quad (4.4)$$

unter Beibehaltung sämtlicher anderer algorithmischer Details.

Beispiel 4.3. Gegeben sei das skalare nichtlineare Anfangswertproblem

$$e^{u''(t)} + u'(t) + u(t) - e^{-\sin(t)} - 1 - \sin(t) - \cos(t) = 0, \quad (4.5a)$$

$$u(t_0) = 1, u'(t_0) = 0 = v(t_0) = 1 \quad (4.5b)$$

mit der exakten Lösung $u^*(t) = 1 + \sin(t)$. Die IIDeC2-DQ2-Lösung erfolgt auf einem äquidistanten Gitter gemäß Definition 1.1 mit 6 Kollokationsknoten im Intervall $[t_0, t_{end}] = [t_0, t_N] = [0, 3]$. Die globalen Fehler bezüglich der exakten Lösung $e^{*,[\#]}(t_{end}) = p^{[\#]}(t_{end}) - u^*(t_{end})$, $\# = \nu, C$ und die daraus berechneten Konvergenzordnungen an der Stelle $t_{end} = 3$ für das Basisverfahren, IIDeC2-DQ2 und für ein Standard-Kollokationsverfahren 2. Ordnung sind in Tabelle 4.1 aufgelistet.

Beispiel 4.4. Wie Beispiel 4.3, jedoch mit 5 äquidistant liegenden Kollokationsknoten. Tabelle 4.2 zeigt die Resultate.

Bemerkung 4.5. Man sieht in Tabelle 4.2, dass die Ordnungen für IIDeC2-DQ2 und für das Vergleichskollokationsverfahren bei einer ungeraden Anzahl von Kollokationsknoten wie im expliziten Fall nicht übereinstimmen (siehe Bemerkung 3.6).

h	DQ2	IIDeC21	IIDeC22	IIDeC23	COLL
0,1	3,31E-05	5,55E-10	3,84E-09	3,84E-09	3,86E-11
0,05	8,28E-06	3,44E-11	2,40E-10	2,40E-10	1,17E-12
0,025	2,07E-06	2,15E-12	1,50E-11	1,50E-11	3,60E-14
0,0125	5,17E-07	1,34E-13	9,39E-13	9,39E-13	1,12E-15
0,1	2,00	4,01	4,00	4,00	5,04
0,05	2,00	4,00	4,00	4,00	5,02
0,025	2,00	4,00	4,00	4,00	5,01
0,0125	2,00	4,00	4,00	4,00	5,01

Tabelle 4.2: Resultate zu Beispiel 4.4.

4.2 Interpolierter Defekt – IIPDeC2-DQ2-Lobatto

Das vorgestellte Verfahren ist eine implizite Variante des in Abschnitt 3.3.1 gezeigten Algorithmus. Der einzige Unterschied besteht darin, dass das nicht äquidistante Gitter ein Lobatto-Gitter sein *muss*. Die Änderung ist erforderlich, da der Algorithmus an den Intervallenden Kollokationsknoten benötigt, damit der implizite Sprungdefekt korrekt ausgewertet werden kann. Das Verfahren ist dennoch superkonvergent, hat allerdings, verglichen mit einem Verfahren auf einem Gauß-Gitter, 2 Ordnungen weniger, also $\mathcal{O}(h^{2\tilde{m}-2})$ statt $\mathcal{O}(h^{2\tilde{m}})$ bei \tilde{m} Lobatto- bzw. Gauß-Knoten pro Intervall.

Basisverfahren und Nachbarproblem sind ident mit jenem des IIDeC2-DQ2-Verfahrens (siehe (4.2) bzw. (4.4)). Der Defekt (4.3) muss jedoch durch folgenden Defekt ersetzt werden, wobei der Sprungdefekt ident bleibt, da an den Intervallenden die Knoten des äquidistanten Gitters und des Lobatto-Gitters zusammenfallen (d.h. $t_{j,m} = \tilde{t}_{j,\tilde{m}}$). Zusammengefasst erhält man

$$\begin{cases} \hat{D}_{1,0}^{A,[\nu]} = \tilde{p}^{[\nu]}(t_{1,0}), \quad \hat{D}_{1,0}^{B,[\nu]} = p_{1,0}^{[\nu]'} - v_0, & j = 1, l = 0 \\ \hat{D}_{j,l}^{[\nu]} = \tilde{p}^{[\nu]}(t_{j,l}), & j \geq 1, l > 0 \\ \hat{D}_{j,0}^{[\nu]} = F \left(t_{j,l}, p_{j,0}^{[\nu]}, \frac{p_{j,0}^{[\nu]'} + p_{j-1,m}^{[\nu]'}}{2}, \frac{p_{j,0}^{[\nu]''} + p_{j-1,m}^{[\nu]''}}{2} + \frac{p_{j,0}^{[\nu]'} - p_{j-1,m}^{[\nu]'}}{h} \right), & j > 1, l = 0, \end{cases} \quad (4.6)$$

wobei für $p^{[\nu]}(t)$ und $\tilde{p}^{[\nu]}(t)$ die Angaben aus Abschnitt 3.3.1 gelten.

Bemerkung 4.6. Man beachte, dass der Sprungdefekt $\hat{D}_{j,0}^{[\nu]}$ ausschließlich aus den Polynomen $p^{[\nu]}(t)$ gebildet wird. Bei der Entwicklung des Algorithmus wurden eine Reihe anderer „Sprungdefektkonstruktionen“ (auch unter Zuhilfenahme der $\tilde{p}^{[\nu]}(t)$) getestet, von denen aber keine das gewünschte Ergebnis bzw. Verhalten zeigte.

Beispiel 4.7. Wie Beispiel 4.3, jedoch mit interpoliertem Defekt, 7 äquidistanten Knoten und 5 Lobatto-Knoten: $\tilde{c}_1 = 0; \tilde{c}_2 = \frac{1}{2} - \frac{\sqrt{21}}{14}; \tilde{c}_3 = \frac{1}{2}; \tilde{c}_4 = \frac{1}{2} + \frac{\sqrt{21}}{14}; \tilde{c}_5 = 1$. Tabelle 4.3 zeigt die Resultate.

Bemerkung 4.8. In Anhang A befindet sich eine kommentierte Version des IIPDeC2-DQ2-Lobatto-Maple-Codes.

h	DQ2	IIPDeC21	IIPDeC22	IIPDeC23	IIPDeC24	L-COLL
0,1	2,30E-05	1,93E-09	9,62E-14	6,07E-16	6,32E-16	6,32E-16
0,05	5,75E-06	1,21E-10	1,51E-15	2,37E-18	2,47E-18	2,47E-18
0,025	1,44E-06	7,54E-12	2,36E-17	9,24E-21	9,63E-21	9,63E-21
0,0125	3,59E-07	4,71E-13	3,69E-19	3,61E-23	3,76E-23	3,76E-23
0,1	2,00	4,00	5,99	8,00	8,00	8,00
0,05	2,00	4,00	6,00	8,00	8,00	8,00
0,025	2,00	4,00	6,00	8,00	8,00	8,00
0,0125	2,00	4,00	6,00	8,00	8,00	8,00

Tabelle 4.3: Resultate zu Beispiel 4.7.

4.3 Lösung durch Rückführung auf ein System 1. Ordnung

Da schon bei expliziten Differentialgleichungen 2. Ordnung kein superkonvergentes Verfahren mit Integralmitteldefekt durch Rückführung auf ein System 1. Ordnung und Ordnungsfolge 2,4,... gefunden werden konnte, war dies auch für implizite Differentialgleichungen nicht zu erwarten. Um trotzdem die grundsätzliche Machbarkeit eines IIQDeC2-* -Verfahrens zu zeigen, wurde IIQDeC2-ITR entwickelt.

4.3.1 IIQDeC2-ITR

Der Vollständigkeit halber sei dieses Verfahren angeführt, obwohl es, wie auch sein explizites Pendant auf einem nicht äquidistanten Gitter, lediglich die Ordnungsfolgen 2,3,... anstatt 2,4,... liefert.

Für die Lösung des impliziten Anfangswertproblems 2. Ordnung vom Typ (4.1) wird dieses auf ein System von Differentialgleichungen 1. Ordnung umgeschrieben

$$u'(t) = v(t), \tag{4.7a}$$

$$F(t, u(t), v(t), v'(t)) = 0, \text{ mit} \tag{4.7b}$$

$$u(t_0) = u_0, u'(t_0) = v(t_0) = v_0. \tag{4.7c}$$

Für die Diskretisierung wird eine halb-implizite Form von (3.41) verwendet

$$\frac{w_{j,l}^{[0]} - w_{j,l-1}^{[0]}}{h_{j,l}} = \frac{1}{2} [x_{j,l-1}^{[0]} + x_{j,l}^{[0]}], \tag{4.8a}$$

$$\frac{1}{2} \left[F \left(t_{j,l-1}, w_{j,l-1}^{[0]}, x_{j,l-1}^{[0]}, \frac{x_{j,l}^{[0]} - x_{j,l-1}^{[0]}}{h_{j,l}} \right) + F \left(t_{j,l}, w_{j,l}^{[0]}, x_{j,l}^{[0]}, \frac{x_{j,l}^{[0]} - x_{j,l-1}^{[0]}}{h_{j,l}} \right) \right] = 0. \tag{4.8b}$$

h	ITR	IIQDeC21	IIQDeC22	IIQDeC23	COLL
0,1	2,99E-05	1,37E-10	1,02E-12	1,04E-12	3,53E-13
0,05	7,47E-06	8,61E-12	1,62E-14	1,65E-14	5,65E-15
0,025	1,87E-06	5,69E-13	2,54E-16	2,60E-16	8,92E-17
0,0125	4,67E-07	3,37E-14	3,99E-18	4,08E-18	1,40E-18
0,1	2,00	3,99	5,98	5,98	5,97
0,05	2,00	4,00	5,99	5,99	5,98
0,025	2,00	4,00	5,99	5,99	5,99
0,0125	2,00	4,00	5,99	5,99	5,99

Tabelle 4.4: Resultate zu Beispiel 4.9.

Formuliert man den Integralmitteldefekt (3.44) ebenfalls teilweise implizit, erhält man

$$\bar{d}_{j,l}^{w,[\nu]} = \frac{w_{j,l}^{[\nu]} - w_{j,l-1}^{[\nu]}}{h_{j,l}} - \sum_{\mu=1}^m \alpha_{l,\mu} x_{j,\mu}^{[\nu]}, \quad (4.9a)$$

$$\bar{D}_{j,l}^{x,[\nu]} = \sum_{\mu=1}^m \alpha_{l,\mu} F \left(t_{j,\mu}, w_{j,\mu}^{[\nu]}, x_{j,\mu}^{[\nu]}, p_{j,\mu}^{x,[\nu]'} \right), \quad (4.9b)$$

wobei für die Quadraturkoeffizienten $\alpha_{l,\mu}$ die Beziehung (1.117) gilt. Mit diesen Defekten gelangt man zum zugehörigen Nachbarproblem

$$\frac{y_{j,l}^{[\nu]} - y_{j,l-1}^{[\nu]}}{h_{j,l}} = \frac{1}{2} \left[z_{j,l-1}^{[\nu]} + z_{j,l}^{[\nu]} \right] + \bar{d}_{j,l}^{w,[\nu]}, \quad (4.10a)$$

$$\frac{1}{2} \left[F \left(t_{j,l-1}, y_{j,l-1}^{[\nu]}, z_{j,l-1}^{[\nu]}, \frac{z_{j,l}^{[\nu]} - z_{j,l-1}^{[\nu]}}{h_{j,l}} \right) + F \left(t_{j,l}, y_{j,l}^{[\nu]}, z_{j,l}^{[\nu]}, \frac{z_{j,l}^{[\nu]} - z_{j,l-1}^{[\nu]}}{h_{j,l}} \right) \right] - \bar{D}_{j,l}^{x,[\nu]} = 0. \quad (4.10b)$$

Beispiel 4.9. Wie Beispiel 4.3, jedoch mit der impliziten Trapezregel als Basisverfahren und Integralmitteldefekt. Tabelle 4.4 zeigt die Resultate.

Beispiel 4.10. Wie Beispiel 4.3, jedoch mit der impliziten Trapezregel als Basisverfahren, Integralmitteldefekt und 4 nicht äquidistanten Kollokationsknoten ($c_1 = 0,0877$; $c_2 = 0,3339$; $c_3 = 0,7899$; $c_4 = 1$). Tabelle 4.5 zeigt die Resultate.

h	ITR	IIQDeC21	IIQDeC22	IIQDeC23	COLL
0,1	1,29E-04	1,05E-07	1,92E-08	8,73E-09	3,11E-09
0,05	3,22E-05	1,43E-08	1,19E-09	5,59E-10	1,96E-10
0,025	8,05E-06	1,69E-09	7,40E-11	3,53E-11	1,23E-11
0,0125	2,01E-06	2,37E-10	4,61E-12	2,22E-12	7,67E-13
0,1	2,00	2,88	4,01	3,96	3,99
0,05	2,00	2,94	4,01	3,98	4,00
0,025	2,00	2,97	4,00	3,99	4,00
0,0125	2,00	2,97	4,00	3,99	4,00

Tabelle 4.5: Resultate zu Beispiel 4.10.

Kapitel 5

Fixpunkt-Kollokationsverfahren

Alle in Abschnitt 1.2 vorgestellten Verfahren konvergieren gegen verschiedene Fixpunkte. Diese (verallgemeinerten) Fixpunkt-Kollokationsverfahren werden mit Ausnahme der Taylor-Fixpunktverfahren in Abschnitt 5.1 hergeleitet. Sämtliche IQDeC- und IPDeC-Verfahren aus den Abschnitten 1.3 und 1.4 haben aufgrund der Konstruktion der Defekte (gemittelt bzw. interpoliert) jeweils dieselben Fixpunkte (siehe dazu die Abschnitte 5.2 und 5.3).

5.1 IDeC-Fixpunkte

Die Idee zur Herleitung der Fixpunkt-Kollokationsbedingungen ist im Prinzip für alle Basisverfahren gleich, jedoch unterschiedlich aufwändig. Die Tatsache, dass die IDeC-Verfahren unterschiedliche Fixpunkte haben, wurde schon in [14] mit der Angabe der ITR-Kollokationsbedingung belegt. Da am Fixpunkt das Basisverfahren und das Nachbarproblem dieselbe Lösung haben, muss die Differenz aus diesen beiden Beziehungen (jeweils mit eingesetzter Kollokationslösung) Null sein.

Am einfachsten ist die Herleitung für die Euler-Verfahren, gefolgt von den impliziten Verfahren der Ordnung 2 (IMR und ITR). Komplizierter sind die Kollokationsbedingungen für die expliziten Verfahren der Ordnung 2 (EMR und Heun) sowie für das klassische Runge-Kutta-Verfahren (RK4).

Obwohl sämtliche IDeC-Verfahren nur auf äquidistanten Gittern konvergieren, werden die in den folgenden Abschnitten beschriebenen Fixpunkt-Kollokationsverfahren für nicht äquidistante Gitter angeschrieben, da sie auch auf diesen Gittern die erwarteten Ordnungen zeigen.

5.1.1 Euler-Kollokation

Die Fixpunktbedingung für die Euler-Basisverfahren kann man unter anderem in [2] nachlesen, zur Einführung in die Konstruktion für andere Basisverfahren werden sie an dieser Stelle kurz wiederholt. Als „Euler-Kollokation“ werden zur Unterscheidung von weiteren *verallgemeinerten* Kollokationsverfahren in dieser Arbeit jene Kollokationsverfahren bezeichnet, welche in der Literatur gemeinhin (nur) als Kollokationsverfahren bezeichnet werden.

EEUL-Kollokation

Wendet man das in den einleitenden Ausführungen beschriebene Konstruktionsprinzip auf die IDeC-EEUL an, erhält man mit $p^{[C]}(t)$ eingesetzt in (1.4a)

$$p_{j,l}^{[C]} = p_{j,l-1}^{[C]} + h_{j,l} f \left(t_{j,l-1}, p_{j,l-1}^{[C]} \right) \quad (5.1)$$

und (1.15a)

$$p_{j,l}^{[C]} = p_{j,l-1}^{[C]} + h_{j,l} f \left(t_{j,l-1}, p_{j,l-1}^{[C]} \right) + \underbrace{p_{j,l-1}^{[C]'} - f \left(t_{j,l-1}, p_{j,l-1}^{[C]} \right)}_{=d_{j,l-1}^{[C]}}. \quad (5.2)$$

Subtrahiert man (5.2) von (5.1), erhält man als Fixpunkt-Kollokationsbedingung für IDeC-EEUL

$$d_{j,l-1}^{[C]} = p_{j,l-1}^{[C]'} - f \left(t_{j,l-1}, p_{j,l-1}^{[C]} \right) = 0, \quad l = 1, \dots, m \quad (5.3)$$

bzw.

$$d_{j,l}^{[C]} = p_{j,l}^{[C]'} - f \left(t_{j,l}, p_{j,l}^{[C]} \right) = 0, \quad l = 0, \dots, m-1. \quad (5.4)$$

D.h. der IDeC-EEUL-Fixpunkt entspricht einer Standard-Kollokation, wobei das Kollokationspolynom $p^{[C]}(t)$ jeweils an den linksseitigen Knoten auszuwerten ist.

IEUL-Kollokation

Geht man analog für die IDeC-IEUL vor, gelangt man zu

$$d_{j,l}^{[C]} = p_{j,l}^{[C]'} - f \left(t_{j,l}, p_{j,l}^{[C]} \right) = 0, \quad l = 1, \dots, m. \quad (5.5)$$

Dieser Fixpunkt ist *die* (Standard-)Kollokation, das Kollokationspolynom $p^{[C]}(t)$ wird jeweils an den rechtsseitigen Knoten ausgewertet.

5.1.2 IMR-Kollokation

Setzt man das Fixpunktpolynom in das IMR-Basisverfahren (1.107)

$$\frac{p_{j,l}^{[C]} - p_{j,l-1}^{[C]}}{h_{j,l}} = f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, p^{[C]} \left(t_{j,l-1} + \frac{h_{j,l}}{2} \right) \right) \quad (5.6)$$

und in das zugehörige Nachbarproblem (1.108)

$$\begin{aligned} \frac{p_{j,l}^{[C]} - p_{j,l-1}^{[C]}}{h_{j,l}} &= f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, p^{[C]} \left(t_{j,l-1} + \frac{h_{j,l}}{2} \right) \right) \\ &+ \underbrace{p^{[C]'} \left(t_{j,l-1} + \frac{h_{j,l}}{2} \right) - f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, p^{[C]} \left(t_{j,l-1} + \frac{h_{j,l}}{2} \right) \right)}_{=d_{\frac{1}{2}}^{[C]}} \end{aligned} \quad (5.7)$$

ein, erhält man durch Subtraktion die IMR-Fixpunkt-Kollokationsbedingung

$$d_{\frac{1}{2}}^{[C]} = p^{[C]'} \left(t_{j,l-1} + \frac{h_{j,l}}{2} \right) - f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, p^{[C]} \left(t_{j,l-1} + \frac{h_{j,l}}{2} \right) \right) = 0, \quad l = 1, \dots, m. \quad (5.8)$$

ν	$e^{[\nu/C]}(t_{end})$
0	2,40561E-04
1	4,66993E-08
2	3,87967E-08
3	3,87943E-08
Heun-COLL	3,87943E-08

Tabelle 5.1: Resultate zu Beispiel 5.2.

Bemerkung 5.1. Die IMR-Kollokation ist somit nichts anderes als „gewöhnliche“ Kollokation ausgewertet an den Mittelpunkten.

5.1.3 ITR-Kollokation

Analog erhält man, wie schon auf S.14 in [14] festgehalten, für die ITR-Kollokationsbedingung

$$\frac{1}{2} \left[d_{j,l-1}^{[C]} + d_{j,l}^{[C]} \right] = \frac{1}{2} \left[p_{j,l-1}^{[C]'} - f \left(t_{j,l-1}, p_{j,l-1}^{[C]} \right) + p_{j,l}^{[C]'} - f \left(t_{j,l}, p_{j,l}^{[C]} \right) \right] = 0, \quad l = 1, \dots, m. \quad (5.9)$$

5.1.4 EMR-Kollokation

Setzt man $p^{[C]}(t)$ in (1.94) und (1.96) ein, erhält man nach anschließender Subtraktion

$$\underbrace{p_{j,l-1}^{[C]'} \left(t_{j,l-1} + \frac{h_{j,l}}{2} \right) - f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, p_{j,l-1}^{[C]} \left(t_{j,l-1} + \frac{h_{j,l}}{2} \right) \right)}_{=d_{j,l-1}^{[C]}} + f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, p_{j,l-1}^{[C]} + \frac{h_{j,l}}{2} p_{j,l-1}^{[C]'} \right) - f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, p_{j,l-1}^{[C]} + \frac{h_{j,l}}{2} f \left(t_{j,l-1}, p_{j,l-1}^{[C]} \right) \right) = 0, \quad l = 1, \dots, m. \quad (5.10)$$

5.1.5 Heun-Kollokation

Analog zum EMR-Fixpunkt erhält man unter Zuhilfenahme von (1.78) und (1.79) für den Heun-Fixpunkt

$$\underbrace{p_{j,l-1}^{[C]'} - f \left(t_{j,l-1}, p_{j,l-1}^{[C]} \right)}_{=d_{j,l-1}^{[C]}} + \underbrace{p_{j,l}^{[C]'} - f \left(t_{j,l}, p_{j,l}^{[C]} \right)}_{=d_{j,l}^{[C]}} + f \left(t_{j,l}, p_{j,l-1}^{[C]} + h_{j,l} p_{j,l-1}^{[C]'} \right) - f \left(t_{j,l}, p_{j,l-1}^{[C]} + h_{j,l} f \left(t_{j,l-1}, p_{j,l-1}^{[C]} \right) \right) = 0, \quad l = 1, \dots, m. \quad (5.11)$$

Beispiel 5.2. Als Beleg für die erfolgreiche Implementierung der Heun-Fixpunkt-Kollokation zeigt Tabelle 5.1 die globalen Fehler der einzelnen IDEC-Iterationen und der Fixpunkt-Kollokation für Beispiel 1.17 am größten Gitter $\mathbf{h} = 0,1$ auf 6 Stellen genau. Man sieht, dass nach nur 3 Iterationen alle 6 Stellen übereinstimmen.

5.1.6 RK4-Kollokation

Gleiche Vorgangsweise wie bei EMR- und Heun-Kollokation liefert mit Hilfe von (1.97) bzw. (1.98) und (1.99) bzw. (1.100) sowie folgenden Abkürzungen

$$d_{j,l-1}^{[C]} \equiv p_{j,l-1}^{[C]'} - f\left(t_{j,l-1}, p_{j,l-1}^{[C]}\right), \quad (5.12a)$$

$$d_{j,l}^{[C]} \equiv p_{j,l}^{[C]'} - f\left(t_{j,l-1}, p_{j,l}^{[C]}\right), \quad (5.12b)$$

$$t_{\frac{1}{2}} \equiv t_{j,l-1} + \frac{h_{j,l}}{2}, \quad (5.12c)$$

$$p_{\frac{1}{2}}^{[C]} \equiv p^{[C]}\left(t_{\frac{1}{2}}\right), \quad (5.12d)$$

$$d_{\frac{1}{2}}^{[C]} \equiv p_{\frac{1}{2}}^{[C]'} - f\left(t_{\frac{1}{2}}, p_{\frac{1}{2}}^{[C]}\right), \quad (5.12e)$$

$$T_{1,p} \equiv f\left(t_{\frac{1}{2}}, p_{j,l-1}^{[C]} + \frac{h_{j,l}}{2} p_{j,l-1}^{[C]'}\right), \quad (5.12f)$$

$$T_{1,f} \equiv f\left(t_{\frac{1}{2}}, p_{j,l-1}^{[C]} + \frac{h_{j,l}}{2} f\left(t_{j,l-1}, p_{j,l-1}^{[C]}\right)\right), \quad (5.12g)$$

$$T_{2,p} \equiv f\left(t_{\frac{1}{2}}, p_{j,l-1}^{[C]} + \frac{h_{j,l}}{2} \left(T_{1,p} + d_{\frac{1}{2}}^{[C]}\right)\right), \quad (5.12h)$$

$$T_{2,f} \equiv f\left(t_{\frac{1}{2}}, p_{j,l-1}^{[C]} + \frac{h_{j,l}}{2} T_{1,f}\right), \quad (5.12i)$$

$$T_{3,p} \equiv f\left(t_{j,l}, p_{j,l-1}^{[C]} + h_{j,l} \left(T_{2,p} + d_{\frac{1}{2}}^{[C]}\right)\right), \quad (5.12j)$$

$$T_{3,f} \equiv f\left(t_{j,l}, p_{j,l-1}^{[C]} + h_{j,l} T_{2,f}\right) \quad (5.12k)$$

die RK4-Fixpunkt-Kollokationsbedingung

$$\begin{aligned} & \frac{1}{6} [2(T_{1,p} - T_{1,f} + T_{2,p} - T_{2,f}) + T_{3,p} - T_{3,f}] \\ & + \frac{1}{6} [d_{j,l-1}^{[C]} + 4d_{\frac{1}{2}}^{[C]} + d_{j,l}^{[C]}] = 0, \quad l = 1, \dots, m. \end{aligned} \quad (5.13)$$

Bemerkung 5.3. Durch die in f steckenden „Schachteldefekte“ (=innere Defekte) sind die Fixpunkt-Kollokationsbedingungen für explizite Basisverfahren komplizierter und somit bei der Lösung eines konkreten Beispiels deutlich rechenaufwändiger in der Auswertung. Aus diesem Grund wurde für das RK4-Beispiel 1.24 eine lineare Differentialgleichung gewählt.

Beispiel 5.4. Als Beleg für die erfolgreiche Implementierung der RK4-Fixpunkt-Kollokation zeigt Tabelle 5.2 die globalen Fehler der einzelnen IDeC-Iterationen und der Fixpunkt-Kollokation für Beispiel 1.24 am größten Gitter $\mathbf{h} = 0,1$ auf 6 Stellen genau. Man sieht, dass nach nur 2 Iterationen alle 6 Stellen übereinstimmen.

5.2 IQDeC-Fixpunkte

Bei sämtlichen IQDeC*-Verfahren ist der Fixpunkt jener der Standard-Kollokation für das implizite Euler-Verfahren aus Abschnitt 5.1.1. Die Begründung liegt darin, dass der Defekt auf dem exakten Differenzenschema (1.116) aufgebaut ist. Formt man diese Beziehung

ν	$e^{\nu/[C]}(t_{end})$
0	2,37609E-11
1	6,40007E-23
2	1,12947E-21
RK4-COLL	1,12947E-21

Tabelle 5.2: Resultate zu Beispiel 5.4.

entsprechend um, erhält man

$$\bar{d}_{j,l}^{[C]} = \frac{p_{j,l}^{[C]} - p_{j,l-1}^{[C]}}{h_{j,l}} - \sum_{\mu=1}^m \alpha_{l,\mu} f(t_{j,\mu}, p_{j,\mu}^{[C]}) = 0, \quad (5.14)$$

was für alle IQDeC-* -Kollokationslösungen gelten muss.

Bemerkung 5.5. Der Zusammenhang mit der Runge-Kutta-Formulierung wird in [2] ausführlich diskutiert.

5.3 IPDeC-Fixpunkte

Durch die Definition des interpolierten Defekts (siehe Abschnitt 1.4) ist der IPDeC-Fixpunkt immer eine (Standard-)Kollokation an den Gauß- bzw. Lobatto-Knoten.

5.4 Kollokation für Differentialgleichungen 2. Ordnung

Die für Vergleichszwecke programmierte Standard-Kollokation für Anfangswertprobleme 2. Ordnung entspricht dem in [17] auf S.283 bzw. S.301, Beispiel 6, ausführlich beschriebenen (Nyström-)Verfahren. Auf eine Wiederholung wird verzichtet.

5.5 Konvergenz von (verallgemeinerten) Kollokationsverfahren

In den Abschnitten 1.2.3 und 1.2.6 wurden für IDeC-EEUL und IDeC-Heun Fixpunkt-Konvergenzbeweise geführt. In den folgenden Abschnitten werden Konvergenzbeweise für das IEUL-, EEUL- und das verallgemeinerte Heun-Kollokationsverfahren geführt. Der Beweis für die IEUL-Kollokation (=„gewöhnliche“ Kollokation) stellt kein „mathematisches Neuland“ dar (siehe [18]), er dient als (didaktische) Vorbereitung für den EEUL- und den Heun-Kollokationsbeweis. Den Abschluss bildet ein Abschnitt zur Superkonvergenz.

5.5.1 IEUL-Kollokation – Konvergenz

Satz 5.6. Für die Fehlerfunktion $e^{*,[C]}(t) = p^{[C]}(t) - u^*(t)$ zwischen der Lösung $p^{[C]}(t)$ des auf dem Gitter 1.1 definierten IEUL-Kollokationsverfahrens (5.5) und der exakten Lösung

$u^*(t)$ des Anfangswertproblems (1.1) gilt

$$e^{*,[C]}(\mathbf{t}_j) = \mathcal{O}(\mathbf{h}^m). \quad (5.15)$$

Bemerkung 5.7. Die Konvergenzaussage von Satz 5.6 gilt sogar gleichmäßig in t . Da jedoch die Konvergenz der IDeC-EEUL/Heun-Lösungen gegen den Fixpunkt (=Euler/Heun-Kollokation) lediglich punktweise bewiesen wurde, beschränkt sich der Beweis von Satz 5.6 vorerst ebenfalls auf punktweise Konvergenz. Die gleichmäßige Konvergenz wird im Anschluss für die Sätze 5.6, 5.8 und 5.10 in Abschnitt 5.5.5 im Zuge der Behandlung der gleichmäßigen Beschränktheit des (verallgemeinerten) Kollokationspolynoms $p^{[C]}(t)$ und dessen Ableitungen gemeinsam gezeigt.

Beweis von Satz 5.6 (lineare Analysis):

Die grundsätzliche Vorgangsweise ist kurz zusammengefasst folgende: Das (Euler-)Kollokationsverfahren lässt sich in ein implizites Runge-Kutta-Verfahren umschreiben (siehe dazu S.212 von [17]), für welches wiederum über die Ordnung der zugrundeliegenden Quadraturformel die Ordnung des Verfahrens gefolgert werden kann. Die Beweisidee lässt sich auf die verallgemeinerten Kollokationsverfahren wie z.B. Heun-Kollokation übertragen (siehe Abschnitt 5.5.3).

Bei m Kollokationsknoten hat das (stückweise definierte) Kollokationspolynom $p^{[C]}(t)$ Grad $\leq m$, die erste Ableitung $p^{[C]'}(t)$ hat Grad $\leq m - 1$. O.B.d.A. wird zur Vereinfachung der Notation der Index j bei den folgenden Ausführungen gleich 1 gesetzt und weggelassen, d.h. es wird das erste Kollokationspolynom betrachtet.

Entsprechend der Kollokationsbedingungen (5.5) kann $p^{[C]'}(t)$ mittels der Lagrange-Polynome

$$\Lambda_l(t) = \prod_{\mu=1, \mu \neq l}^m \frac{t - t_\mu}{t_l - t_\mu}, \quad l = 1, \dots, m \quad (5.16)$$

wie folgt angeschrieben werden

$$p^{[C]'}(t) = \sum_{\mu=1}^m \Lambda_\mu(t) p^{[C]'}(t_\mu) = \sum_{\mu=1}^m \Lambda_\mu(t) f(t_\mu, p^{[C]}(t_\mu)). \quad (5.17)$$

Aufintegration mit Anfangswert $p^{[C]}(t_0) = u_0$ ergibt allgemein

$$p^{[C]}(\tau) = u_0 + \sum_{\mu=1}^m \int_0^\tau \Lambda_l(\sigma) d\sigma f(t_\mu, p^{[C]}(t_\mu)), \quad (5.18)$$

und für die Kollokationsknoten $\tau = t_l$ bzw. das Intervallende $\tau = \mathbf{h}$ mit

$$a_{l,\mu} = \frac{1}{\mathbf{h}} \int_0^{t_l} \Lambda_l(\sigma) d\sigma \quad (5.19a)$$

$$b_\mu = \frac{1}{\mathbf{h}} \int_0^{\mathbf{h}} \Lambda_l(\sigma) d\sigma \quad (5.19b)$$

die (implizite) Runge-Kutta-Darstellung des Kollokationsverfahrens

$$p^{[C]}(t_l) = u_0 + \mathbf{h} \sum_{\mu=1}^m a_{l,\mu} f(t_\mu, p^{[C]}(t_\mu)), \quad (5.20a)$$

$$p^{[C]}(\mathbf{h}) = u_0 + \mathbf{h} \sum_{\mu=1}^m b_\mu f(t_\mu, p^{[C]}(t_\mu)). \quad (5.20b)$$

Mit $p^{[C]'}(t_l) = k_l$ gelangt man zur klassischen Schreibweise¹

$$k_l = f\left(t_0 + c_l \mathbf{h}, u_0 + \mathbf{h} \sum_{\mu=1}^m a_{l,\mu} k_\mu\right). \quad (5.21)$$

Setzt man in die letzte Beziehung die exakte Lösung $u^*(t_l)$ ein, erhält man mit $u^{*'}(t_l) = k_l^*$

$$k_l^* = f\left(t_0 + c_l \mathbf{h}, u_0 + \mathbf{h} \sum_{\mu=1}^m a_{l,\mu} k_\mu^* + \underbrace{\mathcal{O}(\mathbf{h}^m)}_{=\text{Quadraturfehler}}\right). \quad (5.22)$$

Subtrahiert man die linearen Beziehungen (5.22) von (5.21) (gemäß (1.44)), erhält man für die erste Ableitung des Fehlers

$$\begin{aligned} e_l^{*,[C]'} &= k_l - k_l^* \\ &= C_l \left(u_0 + \mathbf{h} \sum_{\mu=1}^m a_{l,\mu} k_\mu \right) + g_l \\ &\quad - C_l \left(u_0 + \mathbf{h} \sum_{\mu=1}^m a_{l,\mu} k_\mu^* + \mathcal{O}(\mathbf{h}^m) \right) + g_l \\ &= C_l \mathbf{h} \sum_{\mu=1}^m a_{l,\mu} \underbrace{(k_\mu - k_\mu^*)}_{=e_\mu^{*,[C]'}} + \underbrace{C_l \mathcal{O}(\mathbf{h}^m)}_{=\mathcal{O}(\mathbf{h}^m)}. \end{aligned} \quad (5.23)$$

In Matrixschreibweise folgt somit

$$\underbrace{\begin{pmatrix} e_1^{*,[C]'} \\ \vdots \\ e_m^{*,[C]'} \end{pmatrix}}_{=\varepsilon'} = \mathbf{h} \underbrace{\begin{pmatrix} C_1 a_{1,1} & \dots & C_1 a_{1,m} \\ \vdots & \ddots & \vdots \\ C_m a_{m,1} & \dots & C_m a_{m,m} \end{pmatrix}}_{=M} \underbrace{\begin{pmatrix} e_1^{*,[C]'} \\ \vdots \\ e_m^{*,[C]'} \end{pmatrix}}_{=\varepsilon'} + \mathcal{O}(\mathbf{h}^m). \quad (5.24)$$

Eine elementare Umformung ergibt (I =Einheitsmatrix)

$$(I - \mathbf{h}M)\varepsilon' = \mathcal{O}(\mathbf{h}^m), \quad (5.25a)$$

$$\varepsilon' = [I - \mathbf{h}M]^{-1} \mathcal{O}(\mathbf{h}^m), \quad (5.25b)$$

¹Im Fall der Kollokation ist die k -Schreibweise der Stufenschreibweise für Runge-Kutta-Verfahren vorzuziehen.

bzw.

$$\|\varepsilon'\| = \mathcal{O}(\mathbf{h}^m), \quad (5.26)$$

da man für hinreichend kleines \mathbf{h} mittels der Neumann-Reihe $[I - \mathbf{h}M]^{-1} = \mathcal{O}(1)$ folgern kann. Am Intervallende gilt

$$\begin{aligned} e_m^{*,[C]} = e^{*,[C]}(\mathbf{h}) &= \mathbf{h} \sum_{\mu=1}^m b_\mu e_\mu^{*,[C]'} + \mathcal{O}(\mathbf{h}^m) \\ &= \mathbf{h} \sum_{\mu=1}^m \mathcal{O}(1) \mathcal{O}(\mathbf{h}^m) + \mathcal{O}(\mathbf{h}^m) = \mathcal{O}(\mathbf{h}^m). \end{aligned} \quad (5.27)$$

Zusammen mit einem klassischen Stabilitätsargument und dem Lemma von Gronwall (siehe z.B. [4]) erhält man die Ordnung auch global.

5.5.2 EEUL-Kollokation – Konvergenz

Es gilt ein zu Satz 5.6 analoger Satz:

Satz 5.8. *Für die Fehlerfunktion $e^{*,[C]}(t) = p^{[C]}(t) - u^*(t)$ zwischen der Lösung $p^{[C]}(t)$ des auf dem Gitter 1.1 definierten EEUL-Kollokationsverfahrens (5.4) und der exakten Lösung $u^*(t)$ des Anfangswertproblems (1.1) gilt*

$$e^{*,[C]}(\mathbf{t}_j) = \mathcal{O}(\mathbf{h}^m). \quad (5.28)$$

Die EEUL-Kollokation kann wie in (5.4) angeschrieben als umindizierte IEUL-Kollokation gedeutet werden, $l = 0, \dots, m-1$. D.h., die Kollokationsknoten wären jeweils die linksseitigen Knoten. Der im Folgenden angeschriebene Fixpunktbeweis nützt diesen Zusammenhang jedoch als Vorbereitung auf den Konvergenzbeweis der Heun-Kollokation nicht aus, als Knoten werden weiterhin jene aus Definition 1.1 verwendet, $l = 1, \dots, m$.

Beweis von Satz 5.8 (lineare Analysis):

Ausgehend von den Lagrange-Polynomen (5.16) über die Knoten von 1 bis m , kann in die Kollokationsbedingungen für (5.4) (mit den Abkürzungen aus Abschnitt 5.5.1) wie folgt eingesetzt werden (linear)

$$\underbrace{\sum_{\mu=1}^m \Lambda_\mu(t_{l-1}) k_\mu}_{=p_{l-1}^{[C]'}} = C_{l-1} \left(u_0 + \mathbf{h} \underbrace{\sum_{\mu=1}^m a_{l-1,\mu} k_\mu}_{=p_{l-1}^{[C]}} \right) + g_{l-1}, \quad l = 1, \dots, m, \quad (5.29)$$

wobei zusätzlich zu den in (5.19a) definierten Koeffizienten $a_{0,\mu} = 0$, $\mu = 1, \dots, m$ gilt. Setzt man in Beziehung (5.29) die exakte Lösung ein

$$\sum_{\mu=1}^m \Lambda_\mu(t_{l-1}) k_\mu^* = C_{l-1} \left(u_0 + \mathbf{h} \sum_{\mu=1}^m a_{l-1,\mu} k_\mu^* + \mathcal{O}(h^m) \right) + g_{l-1}, \quad l = 1, \dots, m, \quad (5.30)$$

ergibt eine anschließende Subtraktion

$$\sum_{\mu=1}^m \Lambda_{\mu}(t_{l-1}) \underbrace{(k_{\mu} - k_{\mu}^*)}_{=e_{\mu}^{*,[C]'}} = C_{l-1} \mathbf{h} \sum_{\mu=1}^m a_{l-1,\mu} \underbrace{(k_{\mu} - k_{\mu}^*)}_{=e_{\mu}^{*,[C]'}} + \underbrace{C_{l-1} \mathcal{O}(h^m)}_{=\mathcal{O}(h^m)}. \quad (5.31)$$

In Matrixschreibweise folgt

$$\begin{aligned} & \underbrace{\begin{pmatrix} \Lambda_1(t_0) & \dots & \Lambda_m(t_0) \\ \vdots & \ddots & \vdots \\ \Lambda_1(t_{m-1}) & \dots & \Lambda_m(t_{m-1}) \end{pmatrix}}_{=M_{\Lambda}} \underbrace{\begin{pmatrix} e_1^{*,[C]'} \\ \vdots \\ e_m^{*,[C]'} \end{pmatrix}}_{=\varepsilon'} \\ = \mathbf{h} & \underbrace{\begin{pmatrix} C_0 a_{0,1} & \dots & C_0 a_{0,m} \\ \vdots & \ddots & \vdots \\ C_{m-1} a_{m-1,1} & \dots & C_{m-1} a_{m-1,m} \end{pmatrix}}_{=\bar{M}} \underbrace{\begin{pmatrix} e_1^{*,[C]'} \\ \vdots \\ e_m^{*,[C]'} \end{pmatrix}}_{=\varepsilon'} + \mathcal{O}(\mathbf{h}^m). \end{aligned} \quad (5.32)$$

Die Matrix M_{Λ} aus (5.32) vereinfacht sich zu

$$M_{\Lambda} = \begin{pmatrix} \Lambda_1(t_0) & \Lambda_2(t_0) & \dots & \dots & \Lambda_m(t_0) \\ 1 & 0 & \dots & 0 & 0 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}. \quad (5.33)$$

Für die Ordnung des Vektors ε' gilt

$$M_{\Lambda} \varepsilon' = \mathbf{h} \bar{M} \varepsilon' \mathcal{O}(\mathbf{h}^m), \quad (5.34a)$$

$$(M_{\Lambda} - \mathbf{h} \bar{M}) \varepsilon' = \mathcal{O}(\mathbf{h}^m), \quad (5.34b)$$

$$\varepsilon' = [M_{\Lambda} - \mathbf{h} \bar{M}]^{-1} \mathcal{O}(\mathbf{h}^m), \quad (5.34c)$$

bzw.

$$\|\varepsilon'\| = \mathcal{O}(\mathbf{h}^m), \quad (5.35)$$

da man für hinreichend kleines \mathbf{h} mittels einer Neumann-Reihe

$$[M_{\Lambda} - \mathbf{h} \bar{M}]^{-1} = [M_{\Lambda} (I - \mathbf{h} M_{\Lambda}^{-1} \bar{M})]^{-1} = [I - \mathbf{h} M_{\Lambda}^{-1} \bar{M}]^{-1} M_{\Lambda}^{-1} = \mathcal{O}(1) \quad (5.36)$$

folgern kann. Am Intervallende gilt Beziehung (5.27) und zusammen mit einem Stabilitätsargument sowie dem Lemma von Gronwall (siehe z.B. [4]) gilt die Ordnung auch global.

Zu zeigen bleibt, dass die Begleitmatrix (5.33) regulär ist, d.h. keinen Eigenwert $\lambda = 0$ besitzt. Das charakteristische Polynom $\pi(\lambda)$ lautet

$$\pi(\lambda) = \lambda^m - \Lambda_1(t_0) \lambda^{m-1} - \Lambda_2(t_1) \lambda^{m-2} - \dots - \Lambda_m(t_0). \quad (5.37)$$

$\lambda = 0$ kann keine Nullstelle von $\pi(\lambda)$ sein, da

$$\pi(\lambda = 0) = -\Lambda_m(t_0) \neq 0. \quad (5.38)$$

Letzterer Zusammenhang gilt, da das Lagrange-Polynom $\Lambda_m(t)$ den Grad $m - 1$ und die Nullstellen t_1, t_2, \dots, t_{m-1} hat. $\lambda = 0$ ist somit kein Eigenwert von M_{Λ} .

Bemerkung 5.9. Mit Satz 5.8 ist auch die Konvergenz von IDEC-EEUL gegen die exakte Lösung aus Satz 1.16 gezeigt (Fixpunkt-Konvergenz von IDEC-EEUL + Konvergenz des Fixpunktes gegen die exakte Lösung \Rightarrow Konvergenz von IDEC-EEUL gegen die exakte Lösung).

5.5.3 Heun-Kollokation – Konvergenz

Satz 5.10. Für die Fehlerfunktion $e^{*,[C]}(t) = p^{[C]}(t) - u^*(t)$ zwischen der Lösung $p^{[C]}(t)$ des auf dem Gitter 1.1 definierten Heun-Kollokationsverfahrens (5.11) und der exakten Lösung $u^*(t)$ des Anfangswertproblems (1.1) gilt

$$e^{*,[C]}(\mathbf{t}_j) = \begin{cases} \mathcal{O}(\mathbf{h}^{m+1}), & m = \text{ungerade} \\ \mathcal{O}(\mathbf{h}^m), & m = \text{gerade}. \end{cases} \quad (5.39)$$

Beweis von Satz 5.10 (lineare Analysis):

Die lineare Fixpunkt-Kollokationsbedingung (5.11) (gemäß (1.44)) mit den Abkürzungen und Bezeichnungen aus den Abschnitten (5.5.1) und (5.5.2) (insbesondere wird an (5.21) erinnert) sowie

$$\hat{m} = \begin{cases} m + 1, & m = \text{ungerade} \\ m, & m = \text{gerade} \end{cases} \quad (5.40)$$

lautet

$$\begin{aligned} & \underbrace{\sum_{\mu=1}^m \Lambda_{\mu}(t_{l-1}) k_{\mu}}_{=p_{i-1}^{[C]'}} - C_{l-1} \left(\underbrace{u_0 + \mathbf{h} \sum_{\mu=1}^m a_{l-1,\mu} k_{\mu}}_{=p_{i-1}^{[C]}} \right) - g_{l-1} \\ + & \underbrace{k_l}_{=p_i^{[C]'}} - C_l \left(\underbrace{u_0 + \mathbf{h} \sum_{\mu=1}^m a_{l,\mu} k_{\mu}}_{=p_i^{[C]}} \right) - g_l \\ + & C_l \left(\underbrace{u_0 + \mathbf{h} \sum_{\mu=1}^m a_{l-1,\mu} k_{\mu}}_{=p_{i-1}^{[C]}} + (c_l - c_{l-1}) \mathbf{h} \underbrace{\sum_{\mu=1}^m \Lambda_{\mu}(t_{l-1}) k_{\mu}}_{=p_{i-1}^{[C]'}} \right) + g_l \\ - & C_l \left(\underbrace{u_0 + \mathbf{h} \sum_{\mu=1}^m a_{l-1,\mu} k_{\mu}}_{=p_{i-1}^{[C]}} + (c_l - c_{l-1}) \mathbf{h} C_{l-1} \left(\underbrace{u_0 + \mathbf{h} \sum_{\mu=1}^m a_{l-1,\mu} k_{\mu}}_{=p_{i-1}^{[C]}} \right) + g_{l-1} \right) - g_l = 0, \\ & l = 1, \dots, m, \end{aligned} \quad (5.41)$$

bzw. vereinfacht

$$\begin{aligned}
& \sum_{\mu=1}^m \Lambda_{\mu}(t_{l-1}) k_{\mu} - C_{l-1} \left(u_0 + \mathbf{h} \sum_{\mu=1}^m a_{l-1,\mu} k_{\mu} \right) - g_{l-1} \\
& + k_l - C_l \left(u_0 + \mathbf{h} \sum_{\mu=1}^m a_{l,\mu} k_{\mu} \right) - g_l \\
& + C_l (c_l - c_{l-1}) \mathbf{h} \left(\sum_{\mu=1}^m \Lambda_{\mu}(t_{l-1}) k_{\mu} - C_{l-1} \left(u_0 + \mathbf{h} \sum_{\mu=1}^m a_{l-1,\mu} k_{\mu} \right) + g_{l-1} \right) = 0, \\
& l = 1, \dots, m.
\end{aligned} \tag{5.42}$$

Analog folgt beim Einsetzen der exakten Lösung unter Zuhilfenahme von (5.40)

$$\begin{aligned}
& \sum_{\mu=1}^m \Lambda_{\mu}(t_{l-1}) k_{\mu}^* - C_{l-1} \left(u_0 + \mathbf{h} \sum_{\mu=1}^m a_{l-1,\mu} k_{\mu}^* + \mathcal{O}(\mathbf{h}^{\hat{m}}) \right) - g_{l-1} \\
& + k_l^* - C_l \left(u_0 + \mathbf{h} \sum_{\mu=1}^m a_{l,\mu} k_{\mu}^* + \mathcal{O}(\mathbf{h}^{\hat{m}}) \right) - g_l \\
& + C_l (c_l - c_{l-1}) \mathbf{h} \left(\sum_{\mu=1}^m \Lambda_{\mu}(t_{l-1}) k_{\mu}^* - C_{l-1} \left(u_0 + \mathbf{h} \sum_{\mu=1}^m a_{l-1,\mu} k_{\mu}^* + \mathcal{O}(\mathbf{h}^{\hat{m}}) \right) + g_{l-1} \right) = 0, \\
& l = 1, \dots, m.
\end{aligned} \tag{5.43}$$

Zieht man (5.43) von (5.42) ab, erhält man

$$\begin{aligned}
& \sum_{\mu=1}^m \Lambda_{\mu}(t_{l-1}) \underbrace{(k_{\mu} - k_{\mu}^*)}_{=e_{\mu}^{*,[C]'}} - C_{l-1} \left(\mathbf{h} \sum_{\mu=1}^m a_{l-1,\mu} \underbrace{(k_{\mu} - k_{\mu}^*)}_{=e_{\mu}^{*,[C]'}} + \mathcal{O}(\mathbf{h}^{\hat{m}}) \right) \\
& + \underbrace{(k_l - k_l^*)}_{=e_l^{*,[C]'}} - C_l \left(\mathbf{h} \sum_{\mu=1}^m a_{l,\mu} \underbrace{(k_{\mu} - k_{\mu}^*)}_{=e_{\mu}^{*,[C]'}} + \mathcal{O}(\mathbf{h}^{\hat{m}}) \right) \\
& + C_l (c_l - c_{l-1}) \mathbf{h} \left(\sum_{\mu=1}^m \Lambda_{\mu}(t_{l-1}) \underbrace{(k_{\mu} - k_{\mu}^*)}_{=e_{\mu}^{*,[C]'}} - C_{l-1} \left(\mathbf{h} \sum_{\mu=1}^m a_{l-1,\mu} \underbrace{(k_{\mu} - k_{\mu}^*)}_{=e_{\mu}^{*,[C]'}} + \mathcal{O}(\mathbf{h}^{\hat{m}}) \right) \right) = 0, \\
& l = 1, \dots, m
\end{aligned} \tag{5.44}$$

bzw. in Matrixschreibweise

$$\begin{aligned}
& \underbrace{\begin{pmatrix} \Lambda_1(t_0) & \dots & \Lambda_m(t_0) \\ \vdots & \ddots & \vdots \\ \Lambda_1(t_{m-1}) & \dots & \Lambda_m(t_{m-1}) \end{pmatrix}}_{=M_\Lambda} \underbrace{\begin{pmatrix} e_1^{*,[C]'} \\ \vdots \\ e_m^{*,[C]'} \end{pmatrix}}_{=\varepsilon'} \\
& - \mathbf{h} \underbrace{\begin{pmatrix} C_0 a_{0,1} & \dots & C_0 a_{0,m} \\ \vdots & \ddots & \vdots \\ C_{m-1} a_{m-1,1} & \dots & C_{m-1} a_{m-1,m} \end{pmatrix}}_{=\bar{M}} \underbrace{\begin{pmatrix} e_1^{*,[C]'} \\ \vdots \\ e_m^{*,[C]'} \end{pmatrix}}_{=\varepsilon'} + \mathcal{O}(\mathbf{h}^{\hat{m}}) \\
& + \underbrace{\begin{pmatrix} e_1^{*,[C]'} \\ \vdots \\ e_m^{*,[C]'} \end{pmatrix}}_{=\varepsilon'} - \mathbf{h} \underbrace{\begin{pmatrix} C_1 a_{1,1} & \dots & C_1 a_{1,m} \\ \vdots & \ddots & \vdots \\ C_m a_{m,1} & \dots & C_m a_{m,m} \end{pmatrix}}_{=M} \underbrace{\begin{pmatrix} e_1^{*,[C]'} \\ \vdots \\ e_m^{*,[C]'} \end{pmatrix}}_{=\varepsilon'} + \mathcal{O}(\mathbf{h}^{\hat{m}}) \\
& + \mathbf{h} \underbrace{\begin{pmatrix} C_1 (c_1 - c_0) \Lambda_1(t_0) & \dots & C_1 (c_1 - c_0) \Lambda_m(t_0) \\ \vdots & \ddots & \vdots \\ C_m (c_m - c_{m-1}) \Lambda_1(t_{m-1}) & \dots & C_m (c_m - c_{m-1}) \Lambda_m(t_{m-1}) \end{pmatrix}}_{=S_1} \underbrace{\begin{pmatrix} e_1^{*,[C]'} \\ \vdots \\ e_m^{*,[C]'} \end{pmatrix}}_{=\varepsilon'} \\
& - \mathbf{h}^2 \underbrace{\begin{pmatrix} C_1 (c_1 - c_0) C_0 a_{0,1} & \dots & C_1 (c_1 - c_0) C_0 a_{0,m} \\ \vdots & \ddots & \vdots \\ C_m (c_m - c_{m-1}) C_{m-1} a_{m-1,1} & \dots & C_m (c_m - c_{m-1}) C_{m-1} a_{m-1,m} \end{pmatrix}}_{=S_2} \underbrace{\begin{pmatrix} e_1^{*,[C]'} \\ \vdots \\ e_m^{*,[C]'} \end{pmatrix}}_{=\varepsilon'} \\
& + \mathcal{O}(\mathbf{h}^{\hat{m}}) = 0 \\
& \quad l = 1, \dots, m, \tag{5.45}
\end{aligned}$$

wobei für M_Λ Beziehung (5.33) gilt.

Analog zur IEUL- bzw. EEUL-Kollokation gilt bei hinreichend kleinem \mathbf{h} für die Ordnung des Vektors ε'

$$\underbrace{(I + M_\Lambda - \mathbf{h}(\bar{M} + M - S_1) - \mathbf{h}^2 S_2)}_{=:M_H} \varepsilon' = \mathcal{O}(\mathbf{h}^{\hat{m}}), \tag{5.46a}$$

$$\varepsilon' = M_H^{-1} \mathcal{O}(\mathbf{h}^{\hat{m}}), \tag{5.46b}$$

$$\|\varepsilon'\| = \mathcal{O}(1) \mathcal{O}(\mathbf{h}^{\hat{m}}) = \mathcal{O}(\mathbf{h}^{\hat{m}}), \tag{5.46c}$$

da unter Zuhilfenahme der Neumann-Reihe

$$\begin{aligned}
M_H^{-1} &= [I + M_\Lambda - \mathbf{h}(\bar{M} + M - S_1) - \mathbf{h}^2 S_2]^{-1} \\
&= [(I + M_\Lambda)(I - [I + M_\Lambda]^{-1}(\mathbf{h}(\bar{M} + M - S_1) - \mathbf{h}^2 S_2))]^{-1} \\
&= [I - \mathbf{h}[I + M_\Lambda]^{-1}(\bar{M} + M - S_1 - \mathbf{h} S_2)]^{-1} [I + M_\Lambda]^{-1} \\
&= \mathcal{O}(1) \tag{5.47}
\end{aligned}$$

gilt. Am Intervallende folgt analog zu Beziehung (5.27)

$$\begin{aligned} e_m^{*,[C]} = e^{*,[C]}(\mathbf{h}) &= \mathbf{h} \sum_{\mu=1}^m b_\mu e_\mu^{*,[C]'} + \mathcal{O}(\mathbf{h}^{\hat{m}}) \\ &= \mathbf{h} \sum_{\mu=1}^m \mathcal{O}(1) \mathcal{O}(\mathbf{h}^{\hat{m}}) + \mathcal{O}(\mathbf{h}^{\hat{m}}) = \mathcal{O}(\mathbf{h}^{\hat{m}}), \end{aligned} \quad (5.48)$$

und zusammen mit einem Stabilitätsargument sowie dem Lemma von Gronwall (siehe z.B. [4]) erhält man die Ordnung auch global.

Darüber hinaus muss die Matrix $I + M_\Lambda$ invertierbar sein. Folgende Ausführungen für die erste Ableitung des Fehlerpolynoms $e^{*,[C]'}(t)$ mit Grad 3 zeigen, dass dies stets gewährleistet ist: Angenommen, es existiert ein Eigenwert $\lambda = 0$, d.h. es gilt

$$(I + M_\Lambda) \begin{pmatrix} e^{*,[C]'}(t_1) \\ e^{*,[C]'}(t_2) \\ e^{*,[C]'}(t_3) \\ e^{*,[C]'}(t_4) \end{pmatrix} = \begin{pmatrix} e^{*,[C]'}(t_0) + e^{*,[C]'}(t_1) \\ e^{*,[C]'}(t_1) + e^{*,[C]'}(t_2) \\ e^{*,[C]'}(t_2) + e^{*,[C]'}(t_3) \\ e^{*,[C]'}(t_3) + e^{*,[C]'}(t_4) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad (5.49)$$

woraus

$$e^{*,[C]'}(t_0) = -e^{*,[C]'}(t_1) = e^{*,[C]'}(t_2) = -e^{*,[C]'}(t_3) = e^{*,[C]'}(t_4) \quad (5.50)$$

folgt. Laut (5.50) müsste $e^{*,[C]'}(t)$ vier Nullstellen besitzen, was aber ein Widerspruch zu Grad 3 ist. D.h. $\lambda = 0$ ist kein Eigenwert von $I + M_\Lambda$.

5.5.4 Superkonvergenz

Angenommen, eine Quadraturformel liefert auf einem Gitter gemäß Definition 1.1 für eine beliebige hinreichend glatte Funktion $f(\sigma)$ die Ordnung s . D.h., es gilt unter Zuhilfenahme von (5.19b)

$$\frac{1}{\mathbf{h}} \int_{\mathbf{t}_{j-1}}^{\mathbf{t}_j} f(\sigma) d\sigma = \sum_{\mu=1}^m b_\mu f(t_{j,\mu}) + \mathcal{O}(\mathbf{h}^s), \quad (5.51)$$

wobei $s \leq 2m$ gilt. Ersetzt man die Ordnungen $\mathcal{O}(\mathbf{h}^m)$ bzw. $\mathcal{O}(\mathbf{h}^{\hat{m}})$ in den Abschnitten 5.5.1, 5.5.2 und 5.5.3 durch $\mathcal{O}(\mathbf{h}^s)$, so erhält man die hohe Ordnung s der Quadraturformel auch global für die bewiesenen (verallgemeinerten) Kollokationsverfahren. An dieser Stelle sei exemplarisch auf die numerischen Resultate der Beispiele 1.37, 1.55 und 4.7 verwiesen.

5.5.5 Gleichmäßige Beschränktheit des Kollokationspolynoms und dessen Ableitungen

Aufbauend auf Satz 5.6 aus Abschnitt 5.5.1 gilt unter Beibehaltung der dort verwendeten Notation folgender Satz für die gleichmäßige Konvergenz der Kollokationslösung $p^{[C]}(t)$ und ihrer Ableitungen. Daraus wird anschließend die gleichmäßige Beschränktheit für $\mathbf{h} \rightarrow 0$ gefolgert².

²Hierbei handelt es sich um ein bekanntes Resultat (siehe z.B. [1]), welches im Folgenden detailliert ausgeführt wird.

Satz 5.11. Für die Fehlerfunktion $e^{*,[C]}(t) = p^{[C]}(t) - u^*(t)$ zwischen der Lösung $p^{[C]}(t)$ des auf dem Gitter 1.1 definierten IEUL-Kollokationsverfahrens (5.5) und der exakten Lösung $u^*(t)$ bzw. für die zugehörigen Ableitungen $e^{*,[C](i)}(t) = p^{[C](i)}(t) - u^{*(i)}(t)$, $i = 1, \dots, m$ des Anfangswertproblems (1.1) gilt

$$e^{*,[C]}(t) = \mathcal{O}(\mathbf{h}^m), \quad (5.52a)$$

$$e^{*,[C](i)}(t) = \mathcal{O}(\mathbf{h}^{m+1-i}), \quad i = 1, \dots, m \quad (5.52b)$$

gleichmäßig in t .

Bemerkung 5.12. Die folgenden Ausführungen gelten ebenso für verallgemeinerte Kollokationsverfahren. Insbesondere wird hierbei auf die Abschnitte 5.5.2 und 5.5.3 hingewiesen.

Beweis von Satz 5.11 (lineare Analysis)³:

Mit den Quadraturkoeffizienten (5.19a) kann man durch numerische Integration der ersten Ableitung der Fehlerfunktion analog zu (5.27) die punktweise Konvergenz an den Kollokationsknoten t_l folgern. Es gilt

$$\begin{aligned} e_l^{*,[C]} &= \mathbf{h} \sum_{\mu=1}^m a_{l,\mu} e_{\mu}^{*,[C]'} + \mathcal{O}(\mathbf{h}^m) \\ &= \mathbf{h} \sum_{\mu=1}^m \mathcal{O}(1) \mathcal{O}(\mathbf{h}^m) + \mathcal{O}(\mathbf{h}^m) = \mathcal{O}(\mathbf{h}^m), \quad l = 1, \dots, m. \end{aligned} \quad (5.53)$$

Um zu einer gleichmäßigen Abschätzung von $e^{*,[C]}(t)$ zu gelangen, interpoliert man unter Zuhilfenahme von (5.16) (Grad $\leq m$) mittels Lagrange-Interpolation

$$e^{*,[C]}(t) = p^{[C]}(t) - u^*(t) \approx \sum_{\mu=0}^m e_{\mu}^{*,[C]} \Lambda_{\mu}(t) =: \tilde{e}^{*,[C]}(t). \quad (5.54)$$

Daraus folgt die *gleichmäßige* Abschätzung

$$e^{*,[C]}(t) = \tilde{e}^{*,[C]}(t) - \text{Interpolationsfehler} \quad (5.55a)$$

$$= \sum_{\mu=0}^m e_{\mu}^{*,[C]} \Lambda_{\mu}(t) + \prod_{\mu=0}^m (t - t_{\mu}) e^{*,[C]}[t_0, \dots, t_m, t] \quad (5.55b)$$

$$= \sum_{\mu=0}^m \underbrace{e_{\mu}^{*,[C]}}_{=\mathcal{O}(\mathbf{h}^m)} \Lambda_{\mu}(t) - \underbrace{(t - t_0) \cdot \dots \cdot (t - t_m)}_{=\mathcal{O}(\mathbf{h}^{m+1})} u^*[t_0, \dots, t_m, t] \quad (5.55c)$$

$$= \mathcal{O}(\mathbf{h}^m), \quad (5.55d)$$

und der erste Teil von Satz 5.11, also (5.52a), ist bewiesen.

Bemerkung 5.13. Beim Übergang von (5.55b) auf (5.55c) können die dividierten Differenzen von $e^{*,[C]}$ ($e^{*,[C]}[t_0, \dots, t_m, t]$) durch jene von u^* ($u^*[t_0, \dots, t_m, t]$) ersetzt werden, da $p^{[C]}(t)$ ein Polynom vom Grad m ist und die angeschriebenen dividierten Differenzen von $p^{[C]}(t)$ somit Null sind.

³Eine ähnliche Vorgangsweise, allerdings nur für die erste Ableitung und im Kontext von linearen Randwertproblemen, wird in [10] gezeigt.

Für den Beweis des zweiten Teils von Satz 5.11, also (5.52b), geht man wie folgt vor: Differentiation von (5.55c) ergibt

$$e^{*,[C]'}(t) = \sum_{\mu=0}^m e_{\mu}^{*,[C]} \Lambda'_{\mu}(t) - \left(\prod_{\mu=0}^m (t - t_{\mu}) u^{*}[t_0, \dots, t_m, t] \right)' \quad (5.56a)$$

$$\begin{aligned} &= \sum_{\mu=0}^m e_{\mu}^{*,[C]} \sum_{\eta=0, \eta \neq \mu}^m \frac{1}{t_{\mu} - t_{\eta}} \prod_{\zeta=0, \zeta \neq \mu, \eta}^m \frac{t - t_{\zeta}}{t_{\mu} - t_{\zeta}} \\ &\quad - \underbrace{\sum_{\eta=0}^m \prod_{\mu=0, \mu \neq \eta}^m (t - t_{\mu})}_{=\mathcal{O}(\mathbf{h}^m)} \underbrace{u^{*}[t_0, \dots, t_m, t]}_{=\mathcal{O}(1)} \\ &\quad - \underbrace{\prod_{\mu=0}^m (t - t_{\mu})}_{=\mathcal{O}(\mathbf{h}^{m+1})} \underbrace{u^{*'}[t_0, \dots, t_m, t]}_{=\mathcal{O}(1)} \\ &= \mathcal{O}(\mathbf{h}^m), \end{aligned} \quad (5.56b)$$

bzw.

$$e^{*,[C]''}(t) = \sum_{\mu=0}^m e_{\mu}^{*,[C]} \Lambda''_{\mu}(t) - \left(\prod_{\mu=0}^m (t - t_{\mu}) u^{*}[t_0, \dots, t_m, t] \right)'' \quad (5.57a)$$

$$\begin{aligned} &= \sum_{\mu=0}^m e_{\mu}^{*,[C]} \sum_{\beta=0, \beta \neq \mu}^m \frac{1}{t_{\mu} - t_{\beta}} \sum_{\eta=0, \eta \neq \mu, \beta}^m \frac{1}{t_{\mu} - t_{\eta}} \prod_{\zeta=0, \zeta \neq \mu, \eta, \beta}^m \frac{t - t_{\zeta}}{t_{\mu} - t_{\zeta}} \\ &\quad - \underbrace{\sum_{\beta=0}^m \sum_{\eta=0, \eta \neq \beta}^m \prod_{\mu=0, \mu \neq \eta, \beta}^m (t - t_{\mu})}_{=\mathcal{O}(\mathbf{h}^{m-1})} \underbrace{u^{*}[t_0, \dots, t_m, t]}_{=\mathcal{O}(1)} \\ &\quad - 2 \underbrace{\sum_{\eta=0}^m \prod_{\mu=0, \mu \neq \eta}^m (t - t_{\mu})}_{=\mathcal{O}(\mathbf{h}^m)} \underbrace{u^{*'}[t_0, \dots, t_m, t]}_{=\mathcal{O}(1)} \\ &\quad - \underbrace{\prod_{\mu=0}^m (t - t_{\mu})}_{=\mathcal{O}(\mathbf{h}^{m+1})} \underbrace{u^{*''}[t_0, \dots, t_m, t]}_{=\mathcal{O}(1)} \\ &= \mathcal{O}(\mathbf{h}^{m-1}). \end{aligned} \quad (5.57b)$$

Man sieht, dass bei jeder (zusätzlichen) Ableitung (durch die Anwendung der Produktregel) eine \mathbf{h} -Potenz „verloren“ geht. Eine Fortsetzung des gezeigten Weges ergibt

$$e^{*,[C]^{(i)}}(t) = \mathcal{O}(\mathbf{h}^{m+1-i}), \quad i = 1, \dots, m. \quad (5.58)$$

Um zu zeigen, dass die Ableitungen der dividierten Differenzen der exakten Lösung (= $u^{*}[t_0, \dots, t_m, t]$) in Bezug auf die Ordnungen als „unkritisch“ bezeichnet werden können, stellt

man diese mittels Integraldarstellung (entsprechende Glattheit von $u^*(t)$ vorausgesetzt) dar (siehe z.B. [20]):

$$\begin{aligned} u^*[t_0, \dots, t_m, t] &= \int_0^1 \int_0^{\tau_1} \dots \int_0^{\tau_{m-1}} \int_0^{\tau_m} u^{*,(m+1)}(t_0 + (t_1 - t_0) \tau_1 + (t_2 - t_1) \tau_2 + \dots \\ &+ (t_m - t_{m-1}) \tau_m + (t - t_m) \tau) d\tau d\tau_m \dots d\tau_2 d\tau_1. \end{aligned} \quad (5.59)$$

Bemerkung 5.14. Die dividierte Differenz $u^*[t_0, \dots, t_m, t]$ ist für $t = t_l, l = 0, \dots, m$, als konfluente dividierte Differenzen zu interpretieren. D.h., dass im Falle einer Division durch 0 der entsprechende Differenzenquotient der dividierten Differenz durch den zugehörigen Differentialquotienten zu ersetzen ist. Die Darstellung als Parameterintegral bezüglich t hat den Vorteil, dass die Differenzierbarkeit für beliebige Stellen t (unmittelbar) gegeben ist.

Für die ersten beiden Ableitungen von (5.59) gilt mit Hilfe der Kettenregel:

$$\begin{aligned} u^{*'}[t_0, \dots, t_m, t] &= \int_0^1 \int_0^{\tau_1} \dots \int_0^{\tau_{m-1}} \int_0^{\tau_m} [u^{*,(m+1)}(t_0 + (t_1 - t_0) \tau_1 + (t_2 - t_1) \tau_2 + \dots \\ &+ (t_m - t_{m-1}) \tau_m + (t - t_m) \tau)]' d\tau d\tau_m \dots d\tau_2 d\tau_1 \\ &= \int_0^1 \int_0^{\tau_1} \dots \int_0^{\tau_{m-1}} \int_0^{\tau_m} u^{*,(m+2)}(t_0 + (t_1 - t_0) \tau_1 + (t_2 - t_1) \tau_2 + \dots \\ &+ (t_m - t_{m-1}) \tau_m + (t - t_m) \tau) \tau d\tau d\tau_m \dots d\tau_2 d\tau_1, \end{aligned} \quad (5.60a)$$

$$\begin{aligned} u^{*''}[t_0, \dots, t_m, t] &= \int_0^1 \int_0^{\tau_1} \dots \int_0^{\tau_{m-1}} \int_0^{\tau_m} [u^{*,(m+2)}(t_0 + (t_1 - t_0) \tau_1 + (t_2 - t_1) \tau_2 + \dots \\ &+ (t_m - t_{m-1}) \tau_m + (t - t_m) \tau) \tau]' d\tau d\tau_m \dots d\tau_2 d\tau_1 \\ &= \int_0^1 \int_0^{\tau_1} \dots \int_0^{\tau_{m-1}} \int_0^{\tau_m} u^{*,(m+3)}(t_0 + (t_1 - t_0) \tau_1 + (t_2 - t_1) \tau_2 + \dots \\ &+ (t_m - t_{m-1}) \tau_m + (t - t_m) \tau) \tau^2 d\tau d\tau_m \dots d\tau_2 d\tau_1. \end{aligned} \quad (5.60b)$$

Partielle Integration von (5.60) ergibt mit Hilfe von (5.59)

$$u^{*'}[t_0, \dots, t_m, t] = u^*[t_0, \dots, t_m, t, t], \quad (5.61a)$$

$$u^{*''}[t_0, \dots, t_m, t] = 2! u^*[t_0, \dots, t_m, t, t, t] \quad (5.61b)$$

bzw. gilt für die i -te Ableitung

$$u^{*(i)}[t_0, \dots, t_m, t] = i! u^*[t_0, \dots, t_m, \underbrace{t, \dots, t}_{(i+1)\text{-mal}}], \quad i = 1, \dots, m. \quad (5.61c)$$

Folgerung: Aus Satz 5.11 folgt unter Zuhilfenahme der Dreiecksungleichung die gleichmäßige Beschränktheit (für $\mathbf{h} \rightarrow 0$) des Kollokationspolynoms $p^{[C]}(t)$ und seiner Ableitungen. Es

gilt

$$\begin{aligned}
\|p^{[C]}(t)\| &= \|u^*(t) + e^{*,[C]}(t)\| \\
&\leq \|u^*(t)\| + \|e^{*,[C]}(t)\| \\
&\leq C_0 + \mathcal{O}(\mathbf{h}^m) \leq \tilde{C}_0,
\end{aligned} \tag{5.62a}$$

$$\begin{aligned}
\|p^{[C]^{(i)}}(t)\| &= \|u^{*(i)}(t) + e^{*,[C]^{(i)}}(t)\| \\
&\leq \|u^{*(i)}(t)\| + \|e^{*,[C]^{(i)}}(t)\| \\
&\leq C_i + \mathcal{O}(\mathbf{h}^{m+1-i}) \leq \tilde{C}_i, \quad i = 1, \dots, m,
\end{aligned} \tag{5.62b}$$

wobei die $C_i, \tilde{C}_i \in \mathbb{R}^+$, $i = 0, \dots, m$, die jeweiligen Schranken bezeichnen.

Kapitel 6

Diskrete Defektkorrektur

Die Idee der Defektkorrektur ist weitaus weitreichender als die bisherigen Kapitel der vorliegenden Arbeit vermitteln. Um einen weiteren Aspekt der Defektkorrektur zu zeigen, werden das Verfahren von Heun aus Abschnitt 1.2.5 und das in Abschnitt 3.1.2 vorgestellte PC-Verfahren für Differentialgleichungen 2. Ordnung als *diskrete* Defektkorrekturverfahren interpretiert¹.

6.1 Verfahren von Heun als diskrete Defektkorrektur

Gegeben sei das Anfangswertproblem (1.1). Diskretisiert man dieses mit der impliziten Trapezregel auf einem simplen äquidistanten Gitter mit Schrittweite h und Index i , erhält man für einen Schritt bei bekanntem x_{i-1}

$$\frac{x_i - x_{i-1}}{h} = \frac{1}{2} (f(t_{i-1}, x_{i-1}) + f(t_i, x_i)) \quad (6.1)$$

mit (exakter diskreter) Lösung x_i .

Schreibt man den gleichen Schritt mit dem Verfahren von Heun² als 1-stufiges Runge-Kutta-Verfahren an, welches als Predictor-Corrector-Verfahren aufgefasst werden kann, führt dies für den Predictor-Schritt zu

$$x_i^P := X_1 = x_{i-1} + h f(t_{i-1}, x_{i-1}) \quad (6.2a)$$

$$\Rightarrow \frac{x_i^P - x_{i-1}}{h} = f(t_{i-1}, x_{i-1}). \quad (6.2b)$$

Für den zugehörigen Corrector-Schritt gilt

$$x_i^{PC,[1]} = x_{i-1} + \frac{h}{2} (f(t_{i-1}, x_{i-1}) + f(t_i, x_i^P)). \quad (6.3)$$

Bei Wiederholung des Corrector-Schritts in der Form

$$x_i^{PC,[\lambda]} = x_{i-1} + \frac{h}{2} (f(t_{i-1}, x_{i-1}) + f(t_i, x_i^{PC,[\lambda-1]})) \quad (6.4)$$

¹Es wird der von H.J. Stetter in [22] vorgestellte abstrakte Formalismus zweimal „mit Leben erfüllt“.

²Das Verfahren von Heun könnte man auch als explizite Trapezregel bezeichnen.

λ	$e_N^{*,PC,[\lambda]}$
1	4,0390835440E-05
2	2,5918923707E-05
3	2,5878147426E-05
4	2,5877999867E-05
5	2,5877999267E-05
6	2,5877999264E-05
ITR	$e_N^{*,ITR}$
-	2,5877999264E-05

Tabelle 6.1: Resultate zu Beispiel 6.1.

gelangt man zu einer Folge von Lösungen $x_i^{PC,[\lambda]}$, $\lambda = 1, \dots$, welche offenbar gegen die ITR-Lösung x_i (=exakte Lösung des diskreten Originalproblems) konvergiert, wie Beispiel 6.1 belegt.

Beispiel 6.1. Löst man das Anfangswertproblem (1.19) auf einem Gitter mit $h = 0,025$ im Intervall $[t_0, t_{end}] = [0, 1]$ (\Rightarrow Anzahl der Intervalle $N = 40$), und betrachtet man die globalen Fehler der ITR-Lösung x_i gemäß (6.1), sowie die Lösungen $x_i^{PC,[\lambda]}$, $\lambda = 1, \dots$ des Heun-PC-Verfahrens (6.2) und (6.3) bzw. (6.4) bezüglich der exakten Lösung $u^*(t)$, d.h. $e_N^{*,ITR} = x_N - u^*(t_{end})$ bzw. $e_N^{*,PC,[\lambda]} = x_N^{PC,[\lambda]} - u^*(t_{end})$, erhält man die in Tabelle 6.1 aufgelisteten Resultate. Man sieht, dass die Folge der Heun-PC-Lösungen offenbar gegen die ITR-Lösung konvergiert.

Wie die folgenden Ausführungen zeigen, kann das eben vorgestellte *iterierte* Heun-Verfahren (=PC-Verfahren) im Sinne der Defektkorrektur als *diskretes* Defektkorrekturverfahren aufgefasst werden. Dabei spielt die gemäß (6.1) diskretisierte Differentialgleichung die Rolle des Originalproblems, der P-Schritt des Heun-Verfahrens ist die Näherungslösung des Originalproblems. Setzt man die P-Lösung x_i^P in das Originalproblem (=ITR-Diskretisierung) ein, erhält man den diskreten Defekt

$$\delta_i^P = \frac{x_i^P - x_{i-1}}{h} - \frac{1}{2} (f(t_{i-1}, x_{i-1}) + f(t_i, x_i^P)). \quad (6.5)$$

Mit diesem kann man das Nachbarproblem wie folgt definieren

$$\frac{x_i^{NP} - x_{i-1}}{h} = \frac{1}{2} (f(t_{i-1}, x_{i-1}) + f(t_i, x_i^{NP})) + \delta_i^P. \quad (6.6)$$

Die exakte Lösung des Nachbarproblems ist $x_i^{NP} = x_i^P$, da mittels (6.5)

$$\begin{aligned} \frac{x_i^{NP} - x_{i-1}}{h} &= \frac{1}{2} (f(t_{i-1}, x_{i-1}) + f(t_i, x_i^{NP})) + \frac{x_i^P - x_{i-1}}{h} - \frac{1}{2} (f(t_{i-1}, x_{i-1}) + f(t_i, x_i^P)), \\ \frac{x_i^{NP} - x_{i-1}}{h} - \frac{1}{2} (f(t_{i-1}, x_{i-1}) + f(t_i, x_i^{NP})) &= \frac{x_i^P - x_{i-1}}{h} - \frac{1}{2} (f(t_{i-1}, x_{i-1}) + f(t_i, x_i^P)) \end{aligned} \quad (6.7)$$

gilt.

Löst man das Nachbarproblem durch Ausführen eines Predictor-Schritts mit $z_{i-1} = x_{i-1}$ und unter Zuhilfenahme von (6.2b)

$$\begin{aligned}
\frac{z_i - z_{i-1}}{h} &= f(t_{i-1}, z_{i-1}) + \delta_i^P \\
&= f(t_{i-1}, x_{i-1}) + \frac{x_i^P - x_{i-1}}{h} - \frac{1}{2} (f(t_{i-1}, x_{i-1}) + f(t_i, x_i^P)) \\
&= 2f(t_{i-1}, x_{i-1}) + \frac{1}{2} (f(t_{i-1}, x_{i-1}) + f(t_i, x_i^P)), \tag{6.8}
\end{aligned}$$

gelangt man zum diskreten Fehlerschätzer analog zu (1.11)

$$\underbrace{x_i^P}_{\text{Näherungslösung des Originalproblems}} - \underbrace{x_i}_{\text{exakte Lösung des Originalproblems}} \approx \underbrace{z_i}_{\text{Näherungslösung des Nachbarproblems}} - \underbrace{x_i^P}_{\text{exakte Lösung des Nachbarproblems}}. \tag{6.9}$$

Formt man (6.9) auf (die exakte Lösung des diskretisierten Problems) x_i um, erhält man

$$x_i \approx x_i^P - (z_i - x_i^P) = 2x_i^P - z_i =: x_i^{DC}, \tag{6.10}$$

wobei die rechte Seite in (6.10) als verbesserte Lösung x_i^{DC} eines Defektkorrekturschritts definiert wird. Wendet man auf (6.10) den Differenzenoperator L an und setzt man (6.6) und (6.5) ein, so führt dies mit Hilfe von (6.2b) zu

$$\begin{aligned}
\frac{x_i^{DC} - x_{i-1}}{h} &= 2 \frac{x_i^P - x_{i-1}}{h} - \frac{z_i - z_{i-1}}{h} \\
&= 2f(t_{i-1}, x_{i-1}) - f(t_{i-1}, \underbrace{z_{i-1}}_{=x_{i-1}}) - \frac{x_i^P - x_{i-1}}{h} + \frac{1}{2} (f(t_{i-1}, x_{i-1}) + f(t_i, x_i^P)) \\
&= \underbrace{f(t_{i-1}, x_{i-1}) - \frac{x_i^P - x_{i-1}}{h}}_{=0} + \frac{1}{2} (f(t_{i-1}, x_{i-1}) + f(t_i, x_i^P)) \\
&= \frac{1}{2} (f(t_{i-1}, x_{i-1}) + f(t_i, x_i^P)). \tag{6.11}
\end{aligned}$$

Vergleicht man (6.11) mit (6.3), so erkennt man, dass $x_i^{DC} = x_i^{PC,[1]}$ gilt. D.h., dass das Verfahren von Heun als diskretes Defektkorrekturverfahren aufgefasst werden kann.

Bemerkung 6.2. Ohne die Details dazu auszuführen, könnte man in gleicher Weise die explizite und die implizite Mittelpunktsregel zueinander in Beziehung setzen.

6.2 PC-Verfahren für Differentialgleichungen 2. Ordnung als diskrete Defektkorrektur

Analog zu Abschnitt 6.1 kann man bei dem in Abschnitt 3.1.2 vorgestellten PC-Verfahren vorgehen. Die Differentialgleichung (3.1) wird mittels des Basisverfahrens (3.2) (= *diskretes Originalproblem*) auf einem einfachen äquidistanten Gitter mit Schrittweite h diskretisiert (Index i). Für einen beliebigen Schritt bei gegebenem x_{i-1} und x_i gilt

$$\frac{x_{i+1} - 2x_i + x_{i-1}}{h^2} = f\left(t_i, x_i, \frac{x_{i+1} - x_{i-1}}{2h}\right) \tag{6.12}$$

mit der exakten Lösung x_{i+1} .

Man betrachte dazu einen Schritt des PC-Verfahrens (3.7) bzw. (3.8)

$$\frac{x_{i+1}^P - 2x_i + x_{i-1}}{h^2} = f_P \left(t_i, x_i, \frac{x_i - x_{i-1}}{h} \right) \quad (6.13)$$

mit der Lösung x_{i+1}^P , bzw.

$$\frac{x_{i+1}^{PC} - 2x_i + x_{i-1}}{h^2} = f \left(t_i, x_i, \frac{x_{i+1}^P - x_{i-1}}{2h} \right) \quad (6.14)$$

mit der Lösung x_{i+1}^{PC} .

Definiert man den diskreten Defekt δ_{i+1}^P als jenes Residuum, welches beim Einsetzen der Predictor-Lösung (=Näherungslösung des Originalproblems=Basisverfahren der diskreten Defektkorrektur) in das Originalproblem (=implizite Diskretisierung) entsteht, erhält man

$$\begin{aligned} \delta_{i+1}^P &= \frac{x_{i+1}^P - 2x_i + x_{i-1}}{h^2} - f \left(t_i, x_i, \frac{x_{i+1}^P - x_{i-1}}{2h} \right) \\ &= f_P \left(t_i, x_i, \frac{x_i - x_{i-1}}{h} \right) - f \left(t_i, x_i, \frac{x_{i+1}^P - x_{i-1}}{2h} \right). \end{aligned} \quad (6.15)$$

Mit diesem Defekt kann das zugehörige Nachbarproblem angeschrieben werden

$$\frac{x_{i+1}^{NP} - 2x_i + x_{i-1}}{h^2} = f \left(t_i, x_i, \frac{x_{i+1}^{NP} - x_{i-1}}{2h} \right) + \delta_{i+1}^P. \quad (6.16)$$

Dieses hat die exakte Lösung $x_{i+1}^{NP} = x_{i+1}^P$, da durch Einsetzen des Defekts (6.15) in (6.16) und elementarer Umformung

$$\begin{aligned} &\frac{x_{i+1}^{NP} - 2x_i + x_{i-1}}{h^2} - f \left(t_i, x_i, \frac{x_{i+1}^{NP} - x_{i-1}}{2h} \right) \\ &= \frac{x_{i+1}^P - 2x_i + x_{i-1}}{h^2} - f \left(t_i, x_i, \frac{x_{i+1}^P - x_{i-1}}{2h} \right) \end{aligned} \quad (6.17)$$

gilt.

Löst man das Nachbarproblem (6.16) mit dem Predictor (6.13), $z_{i-1} = x_{i-1}$ und $z_i = x_i$, sowie durch Einsetzen von (6.15)

$$\begin{aligned} \frac{z_{i+1} - 2z_i + z_{i-1}}{h^2} &= f_P \left(t_i, z_i, \frac{z_i - z_{i-1}}{h} \right) + \delta_{i+1}^P \\ &= 2f_P \left(t_i, x_i, \frac{x_i - x_{i-1}}{h} \right) - f \left(t_i, x_i, \frac{x_i^P - x_{i-1}}{h} \right), \end{aligned} \quad (6.18)$$

gelangt man zum Fehlerschätzer für die diskrete Defektkorrektur analog zu (1.11)

$$\underbrace{x_{i+1}^P}_{\text{Näherungslösung des Originalproblems}} - \underbrace{x_{i+1}}_{\text{exakte Lösung des Originalproblems}} \approx \underbrace{z_{i+1}}_{\text{Näherungslösung des Nachbarproblems}} - \underbrace{x_{i+1}^P}_{\text{exakte Lösung des Nachbarproblems}}. \quad (6.19)$$

Formt man (6.19) auf x_{i+1} um, und definiert man die rechte Seite als verbesserte Lösung x_{i+1}^{DC} eines Defektkorrekturschritts, erhält man

$$x_{i+1} \approx x_{i+1}^P - (z_{i+1} - x_{i+1}^P) = 2x_{i+1}^P - z_{i+1} =: x_{i+1}^{DC}. \quad (6.20)$$

Bildet man auf Basis der Definition von (6.20) den 2. Differenzenquotienten und setzt man zuerst (6.18) und danach (6.13) ein, erhält man

$$\begin{aligned} \frac{x_{i+1}^{DC} - 2x_i + x_{i-1}}{h^2} &= 2 \frac{x_{i+1}^P - 2x_i + x_{i-1}}{h^2} - \frac{z_{i+1} - 2z_i + z_{i-1}}{h^2} \\ &= 2 \frac{x_{i+1}^P - 2x_i + x_{i-1}}{h^2} - f_P \left(t_i, z_i, \frac{x_i - x_{i-1}}{h} \right) - \delta_{i+1}^P \\ &= 2 \frac{x_{i+1}^P - 2x_i + x_{i-1}}{h^2} - \frac{x_{i+1}^P - 2x_i + x_{i-1}}{h^2} - \delta_{i+1}^P \\ &= \underbrace{\frac{x_{i+1}^P - 2x_i + x_{i-1}}{h^2} - f_P \left(t_i, x_i, \frac{x_i - x_{i-1}}{h} \right)}_{=0} + f \left(t_i, x_i, \frac{x_{i+1}^P - x_{i-1}}{2h} \right) \\ &= f \left(t_i, x_i, \frac{x_{i+1}^P - x_{i-1}}{2h} \right), \end{aligned} \quad (6.21)$$

woraus durch Vergleich mit dem Corrector-Schritt (6.14) $x_{i+1}^{DC} = x_{i+1}^{PC}$ folgt. D.h. PC-Verfahren sind diskrete Defektkorrekturverfahren.

Bemerkung 6.3. Das aus explizitem und implizitem Adams-Verfahren gebildete klassische PC-Verfahren kann ebenfalls als diskrete Defektkorrektur aufgefasst werden [11].

Bemerkung 6.4. Man könnte aufgrund der obigen Ausführungen die diskrete Defektkorrektur auch als eine verfahrensimmanente *innere* Defektkorrektur bezeichnen, die „aufgesetzte“ *polynomiale* Defektkorrektur aus den Kapiteln 1 bis 4 als *äußere*.

Kapitel 7

Schlussbemerkungen und Ausblick

7.1 Grenzen der Defektkorrektur

Die folgenden 2 Abschnitte zeigen die Grenzen der Defektkorrektur für Integralmittel- und interpolierten Defekt auf. Obwohl IDeC-RK4 tadellos funktioniert (siehe Abschnitt 1.2.8), zeigen die Verfahren mit den genannten Defekten trotz Konvergenz gegen den korrekten Fixpunkt schon am äquidistanten Gitter nicht das erwartete (Ordnungs-)Verhalten.

7.1.1 IQDeC-RK4

Mit dem Basisverfahren (1.97) bzw. (1.98) und dem Integralmitteldefekt (1.120) kann das IQDeC-RK4-Nachbarproblem wie folgt angeschrieben werden

$$Z_1^{[0]} = z_{j,l-1}^{[0]}, \quad (7.1a)$$

$$Z_2^{[0]} = z_{j,l-1}^{[0]} + \frac{h_{j,l}}{2} \left[f \left(t_{j,l-1}, Z_1^{[0]} \right) + \bar{d}_{j,l}^{[\nu]} \right], \quad (7.1b)$$

$$Z_3^{[0]} = z_{j,l-1}^{[0]} + \frac{h_{j,l}}{2} \left[f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, Z_2^{[0]} \right) + \bar{d}_{j,l}^{[\nu]} \right], \quad (7.1c)$$

$$Z_4^{[0]} = z_{j,l-1}^{[0]} + h_{j,l} \left[f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, Z_3^{[0]} \right) + \bar{d}_{j,l}^{[\nu]} \right], \quad (7.1d)$$

und

$$\begin{aligned} z_{j,l}^{[0]} = z_{j,l-1}^{[0]} &+ h_{j,l} \left[\frac{1}{6} \left[f \left(t_{j,l-1}, Z_1^{[0]} \right) \right] \right. \\ &+ \frac{1}{3} \left[f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, Z_2^{[0]} \right) \right] \\ &+ \frac{1}{3} \left[f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, Z_3^{[0]} \right) \right] \\ &\left. + \frac{1}{6} \left[f \left(t_{j,l}, Z_4^{[0]} \right) \right] + \bar{d}_{j,l}^{[\nu]} \right]. \end{aligned} \quad (7.2)$$

Beispiel 7.1. Wie Beispiel 1.24, jedoch mit Integralmitteldefekt. Tabelle 7.1 zeigt die Resultate.

h	RK4	IQDeC1	IQDeC2	IQDeC3	COLL
0,1	3,70E-10	7,61E-16	2,22E-17	2,23E-17	2,23E-17
0,05	2,32E-11	1,16E-17	8,64E-20	8,69E-20	8,69E-20
0,025	1,45E-12	1,80E-19	3,37E-22	3,39E-22	3,39E-22
0,0125	9,05E-14	2,81E-21	1,32E-24	1,32E-24	1,32E-24
0,1	4,00	6,04	8,00	8,00	8,00
0,05	4,00	6,01	8,00	8,00	8,00
0,025	4,00	6,00	8,00	8,00	8,00
0,0125	4,00	6,00	8,00	8,00	8,00

Tabelle 7.1: Resultate zu Beispiel 7.1.

Bemerkung 7.2. IQDeC-RK4-Lösung konvergiert zwar gegen den Fixpunkt (=IEUL-Kollokation), jedoch mit der Ordnungsfolge 4,6,8,... anstatt 4,8,8,....

7.1.2 IPDeC-RK4

Mit dem Basisverfahren (1.97) bzw. (1.98) und den interpolierten Defekten (1.146) und (1.150) folgt für das Nachbarproblem

$$Z_1^{[0]} = z_{j,l-1}^{[0]}, \quad (7.3a)$$

$$Z_2^{[0]} = z_{j,l-1}^{[0]} + \frac{h_{j,l}}{2} \left[f \left(t_{j,l-1}, Z_1^{[0]} \right) + \hat{d}_{j,l-1}^{[\nu]} \right], \quad (7.3b)$$

$$Z_3^{[0]} = z_{j,l-1}^{[0]} + \frac{h_{j,l}}{2} \left[f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, Z_2^{[0]} \right) + \hat{d}_{\frac{1}{2}}^{[\nu]} \right], \quad (7.3c)$$

$$Z_4^{[0]} = z_{j,l-1}^{[0]} + h_{j,l} \left[f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, Z_3^{[0]} \right) + \hat{d}_{\frac{1}{2}}^{[\nu]} \right], \quad (7.3d)$$

und

$$\begin{aligned}
z_{j,l}^{[0]} = z_{j,l-1}^{[0]} &+ h_{j,l} \left[\frac{1}{6} \left[f \left(t_{j,l-1}, Z_1^{[0]} \right) + \hat{d}_{j,l-1}^{[\nu]} \right] \right. \\
&+ \frac{1}{3} \left[f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, Z_2^{[0]} \right) + \hat{d}_{\frac{1}{2}}^{[\nu]} \right] \\
&+ \frac{1}{3} \left[f \left(t_{j,l-1} + \frac{h_{j,l}}{2}, Z_3^{[0]} \right) + \hat{d}_{\frac{1}{2}}^{[\nu]} \right] \\
&\left. + \frac{1}{6} \left[f \left(t_{j,l}, Z_4^{[0]} \right) + \hat{d}_{j,l}^{[\nu]} \right] \right]. \quad (7.4)
\end{aligned}$$

Beispiel 7.3. Wie Beispiel 1.24, jedoch mit interpoliertem Defekt mit 5 äquidistanten und 5 Gauß-Knoten. Tabelle 7.2 zeigt die Resultate.

Bemerkung 7.4. Die IPDeC-RK4-Lösung konvergiert zwar gegen den Fixpunkt (=Gauß-Kollokation), jedoch (ähnlich wie die IQDeC-RK4-Lösung) mit der Ordnungsfolge 4,6,8,... anstatt 4,8,10,....

h	RK4	IPDeC1	IPDeC2	IPDeC3	IPDeC4	G-COLL
0,1	2,43E-09	4,42E-15	3,02E-19	1,18E-19	1,18E-19	1,18E-19
0,05	1,52E-10	6,90E-17	8,34E-22	1,15E-22	1,15E-22	1,15E-22
0,025	9,49E-12	1,08E-18	2,92E-24	1,12E-25	1,12E-25	1,12E-25
0,0125	5,93E-13	1,68E-20	1,11E-26	1,09E-28	1,09E-28	1,09E-28
0,1	4,00	6,00	8,50	10,00	10,00	10,00
0,05	4,00	6,00	8,16	10,00	10,00	10,00
0,025	4,00	6,00	8,04	10,00	10,00	10,00
0,0125	4,00	6,00	8,04	10,00	10,00	10,00

Tabelle 7.2: Resultate zu Beispiel 7.3.

Ohne eine Analyse des Problems durchgeführt zu haben, liegt die Ursache für die langsamere Konvergenz gegen den Fixpunkt in der „Verschmierung“ des Defekts durch die Bildung des Integralmittels (bei IQDeC) bzw. die Interpolation (bei IPDeC). Bei Basisverfahren niedrigerer Ordnung (bis einschließlich 2) sind diese Verschmierungen offensichtlich irrelevant¹.

7.2 Fehlende Algorithmen

Im Zuge der Erstellung der vorliegenden Arbeit wurden große Anstrengungen unternommen einen IIQDeC2-DQ2-Algorithmus zu finden. Das Problem dabei ist, dass man bei der Diskretisierung der 2. Ableitung durch Integration (siehe Abschnitt 3.2.1) einen Sprungdefekt zum (äquidistant angeschriebenen) Integral

$$\int_{-1}^1 (1 - |\vartheta|) F(t, u(t + \vartheta h), u'(t + \vartheta h), u''(t + \vartheta h)) d\vartheta \quad (7.5)$$

„basteln“ müsste, was aufgrund der „Verschachtelung“ der zweiten Ableitung $u''(t)$ in F leider nicht gelungen ist bzw. nicht gelingen kann.

Für IIPDeC2-DQ2 auf einem Gauß-Gitter konnte ebenfalls kein korrekt arbeitender Sprungdefekt gefunden werden. Abhilfe schaffte, wie in Abschnitt 4.2 beschrieben, der Umstieg auf ein Lobatto-Gitter.

7.3 Ausblick auf partielle Differentialgleichungen

Die nächsten beiden Abschnitte zeigen, dass eine Anwendung der Defektkorrektur auf partielle Differentialgleichungen möglich ist².

¹Bei Basisverfahren der Ordnung 1 (=Euler-Verfahren) tritt beim Integralmittelddefekt gar keine Verschmierung auf (exaktes Differenzenschema).

²Weiterführende Analysen wurden nicht durchgeführt.

7.3.1 Wärmeleitungsgleichung mit IQDeC-IEUL

Ziel dieses Teils war, zu untersuchen, ob die Defektkorrektur mit Integralmitteldefekt die (inhomogene lineare bzw.) nichtlineare Wärmeleitungsgleichung

$$\frac{\partial u(x, t)}{\partial t} = \frac{\partial^2 u(x, t)}{\partial x^2} + \frac{1}{1 + u(x, t)^2} + \Phi(x, t) = f \left(t, x, u(x, t), \frac{\partial^2 u(x, t)}{\partial x^2} \right) \quad (7.6)$$

für bestimmte Anfangsbedingungen und Randwerte (siehe Beispiele 7.5 und 7.6) „zufriedenstellend“ lösen kann. D.h., sind die zu erwartenden Ordnungsfolgen tatsächlich beobachtbar? Die Funktion $\Phi(x, t)$ wurde so bestimmt, dass die Differentialgleichung eine vorgegebene exakte Lösung $u^*(x, t)$ hat.

Für die äquidistante Ortsdiskretisierung wird der Index $i = 1, 2, \dots, \tilde{N} - 1$ eingeführt ($\Rightarrow y_{i,(j,l)}^{[\nu]}$), die Schrittweite der Ortsdiskretisierung wird mit Δx bezeichnet. Als Basisverfahren für die Defektkorrektur wurde (aufgrund der Steifheit des Problems) das implizite Euler-Verfahren gewählt

$$\frac{y_{i,(j,l)}^{[0]} - y_{i,(j,l-1)}^{[0]}}{h_{j,l}} = f \left(t_{j,l}, x_i, y_{i,(j,l)}^{[0]}, \frac{y_{i+1,(j,l)}^{[0]} - 2y_{i,(j,l)}^{[0]} + y_{i-1,(j,l)}^{[0]}}{\Delta x^2} \right). \quad (7.7)$$

D.h., dass für jeden Zeitpunkt $t_{j,l}$ des Gitters ein $(\tilde{N} - 2)$ -dimensionales nichtlineares Gleichungssystem gelöst werden muss. Die Lösung erfolgt durch das in Maple mit *fsolve* implementierte Newtonverfahren mit der Option *fulldigits*.

Um den (störenden) Einfluss der Ortsdiskretisierung auf das Ergebnis zu unterdrücken, wurde dieser „ausgeblendet“. Dazu subtrahiert man in jedem Schritt den durch das Verfahren verursachten Ortsdiskretisierungsfehler (bei bekannter Lösung durch Abtastung an den Gitterpunkten)

$$F_{i,(j,l)} = \frac{\partial u^*(x_i, t_{j,l})}{\partial t} - f \left(t_{j,l}, x_i, u^*(x_i, t_{j,l}), \frac{u^*(x_{i+1}, t_{j,l}) - 2u^*(x_i, t_{j,l}) + u^*(x_{i-1}, t_{j,l})}{\Delta x^2} \right) \quad (7.8)$$

von Gleichung (7.7). Für das resultierende *semidiskrete* Problem erhält man somit

$$\frac{y_{i,(j,l)}^{[\nu]} - y_{i,(j,l-1)}^{[\nu]}}{h_{j,l}} = f \left(t_{j,l}, x_i, y_{i,(j,l)}^{[\nu]}, \frac{y_{i+1,(j,l)}^{[\nu]} - 2y_{i,(j,l)}^{[\nu]} + y_{i-1,(j,l)}^{[\nu]}}{\Delta x^2} \right) - F_{i,(j,l)}. \quad (7.9)$$

Der zugehörige Integralmitteldefekt $\bar{d}_{i,(j,l)}^{[\nu]}$ wird analog zu Beziehung (1.120) unter Zuhilfenahme der entsprechenden Koeffizienten (1.117) bei Ausblendung der Ortsdiskretisierung wie folgt berechnet

$$\bar{d}_{i,(j,l)}^{[\nu]} = \frac{y_{i,(j,l)}^{[\nu]} - y_{i,(j,l-1)}^{[\nu]}}{h_{j,l}} - \sum_{\mu=1}^m \left(f \left(t_{j,\mu}, x_i, y_{i,(j,\mu)}^{[\nu]}, \frac{y_{i+1,(j,\mu)}^{[\nu]} - 2y_{i,(j,\mu)}^{[\nu]} + y_{i-1,(j,\mu)}^{[\nu]}}{\Delta x^2} \right) - F_{i,(j,\mu)} \right). \quad (7.10)$$

Für das Nachbarproblem gilt (ebenfalls unter Ausblendung des Ortsdiskretisierungsfehlers)

$$\frac{z_{i,(j,l)}^{[\nu]} - z_{i,(j,l-1)}^{[\nu]}}{h_{j,l}} = f \left(t_{j,l}, x_i, z_{i,(j,l)}^{[\nu]}, \frac{z_{i+1,(j,l)}^{[\nu]} - 2z_{i,(j,l)}^{[\nu]} + z_{i-1,(j,l)}^{[\nu]}}{\Delta x^2} \right) - F_{i,(j,l)} + \bar{d}_{i,(j,l)}^{[\nu]}. \quad (7.11)$$

Die verbesserte Näherung erhält man mittels

$$y_{i,(j,l)}^{[\nu]} = y_{i,(j,l)}^{[0]} - \left(z_{i,(j,l)}^{[\nu-1]} - y_{i,(j,l)}^{[\nu-1]} \right), \quad i = 1, \dots, \tilde{N} - 1. \quad (7.12)$$

Beispiel 7.5. Die exakte Lösung des Beispiels³ lautet $u^*(x, t) = x(1 - x)e^t$. Gerechnet wird sowohl zeitlich als auch räumlich im Einheitsintervall. Die Anfangsbedingungen lauten $u^*(x, 0) = x(1 - x)$, die Randbedingungen ergeben sich mit $u(0, t) = u(1, t) = 0$. Für die zugehörige Funktion $\Phi(x, t)$ gilt

$$\Phi(x, t) = x(1 - x)e^t + 2e^t - \frac{1}{1 + x^2(1 - x)^2 e^{2t}}. \quad (7.13)$$

Weiters wurden die Ortsdiskretisierungsschrittweite $\Delta x = 0,1$ und am größten Gitter die Intervalllänge $\mathbf{h} = 0,05$ gewählt. Gerechnet wird auf einem nicht äquidistanten Gitter mit 5 Gitterpunkten, davon 4 Kollokationsknoten folgender relativer Lagen: ($c_0 = 0$;) $c_1 = 0,02$; $c_2 = 0,5$; $c_3 = 0,7$; $c_4 = 1$.

Tabelle 7.3 zeigt im oberen Abschnitt neben den Schrittweiten die (mittels euklidischer Norm über alle Gitterpunkte an der Stelle $t_{end} = 1$) berechneten globalen Fehler für das Basisverfahren (IEUL) und die einzelnen Iterationen der Defektkorrektur (IQDeC). Im unteren Abschnitt sind die daraus berechneten Konvergenzordnungen aufgelistet.

Abbildung 7.1 zeigt die globalen Fehler aufgetragen über der Schrittweite mit doppelt logarithmischem Maßstab.

Beispiel 7.6. behandelt die lineare inhomogene Wärmeleitungsgleichung der Bauart

$$\frac{\partial u(x, t)}{\partial t} = \frac{\partial^2 u(x, t)}{\partial x^2} + \Phi(x, t). \quad (7.14)$$

Die exakte Lösung lautet $u^*(x, t) = \sin(\pi x) \cdot e^t + \sin(20\pi x) \cdot e^{-20t}$. Die zugehörigen Anfangsbedingungen lauten $u(x, 0) = \sin(\pi x) + \sin(20\pi x)$. Die Randbedingungen sind ident mit jenen aus Beispiel 7.5. Die Anzahl der Gitterpunkte beträgt 7 (davon 6 Kollokationsknoten), diese sind äquidistant innerhalb der Teilintervalle verteilt. Die analog zu Beispiel 7.5 aufgelisteten Resultate können Tabelle 7.4 und Abbildung 7.2 entnommen werden.

7.3.2 Wellengleichung mit IQDeC2-IEUL

Analog zur im Abschnitt 7.3.1 behandelten Wärmeleitungsgleichung, allerdings mit Rückführung auf ein System 1. Ordnung, wurde auch die Wellengleichung

$$\frac{\partial^2 u(x, t)}{\partial t^2} = c^2 \frac{\partial^2 u(x, t)}{\partial x^2} \quad (7.15)$$

³Bei diesem Fall tritt kein Ortsdiskretisierungsfehler auf. Die beschriebene Ausblendung wäre somit nicht erforderlich.

behandelt, wobei c die Ausbreitungsgeschwindigkeit der Welle bezeichnet.

Eine zum vorigen Abschnitt analoge Diskretisierung und Defektdefinition, allerdings ohne Ausblendung des Ortdiskretisierungsfehlers, liefert, wie folgendes Beispiel zeigt, akzeptable Resultate.

Beispiel 7.7. Die exakte Lösung für $c = 2$, die Anfangsbedingungen $u^*(x, 0) = \sin(\pi x)$ und $u^{*'}(x, 0) = 0$, sowie für die Randbedingungen $u(0, t) = u(1, t) = 0$ lautet $u^*(x, t) = \sin(\pi x) \cos(2\pi t)$. Gerechnet wird sowohl zeitlich als auch räumlich im Einheitsintervall. Weiters wurden die Ortdiskretisierungsschrittweite Δx mit 0,1 und die Intervalllänge \mathbf{h} am größten Gitter mit 0,01 festgelegt. Das zugrunde liegende zeitliche Gitter $t_{j,l}$ ist äquidistant mit 5 Gitterpunkten, d.h. 4 Kollokationsknoten. Tabelle 7.5 zeigt im oberen Abschnitt neben den Schrittweiten die (mittels euklidischer Norm über alle Punkte an der Stelle $t_{end} = 1$) berechneten globalen Fehler für das Basisverfahren (IEUL) und die einzelnen Iterationen der Defektkorrektur (IQDeC). Im unteren Abschnitt sind die daraus berechneten Konvergenzordnungen aufgelistet.

Abbildung 7.3 zeigt die globalen Fehler aufgetragen über der Schrittweite mit doppelt logarithmischem Maßstab.

h	IEUL	IQDeC1	IQDeC2	IQDeC3	IQDeC4
0,05	1,19E-03	3,52E-06	3,07E-07	9,17E-08	3,61E-08
0,025	6,06E-04	1,01E-06	4,39E-08	8,06E-09	3,01E-09
0,0125	3,06E-04	2,76E-07	5,77E-09	5,55E-10	1,96E-10
0,00625	1,54E-04	7,26E-08	7,50E-10	2,56E-11	7,57E-11
0,05	0,97	1,80	2,81	3,51	3,58
0,025	0,98	1,87	2,93	3,86	3,94
0,0125	0,99	1,93	2,94	4,44	4,69

Tabelle 7.3: Resultate zu Beispiel 7.5.

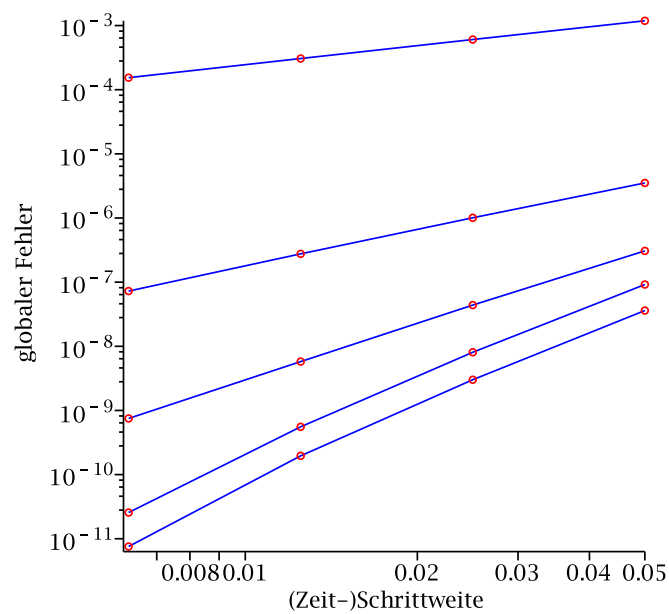


Abbildung 7.1: Der Graph zeigt (doppelt logarithmisch) die globalen Fehler aufgetragen über der Schrittweite für Beispiel 7.5.

h	IEUL	IQDeC1	IQDeC2	IQDeC3	IQDeC4	IQDeC5	IQDeC6
0,05	2,34E-03	8,86E-06	2,97E-08	3,60E-10	4,20E-12	1,83E-13	2,32E-15
0,025	1,17E-03	2,22E-06	3,77E-09	2,21E-11	1,44E-13	2,75E-15	8,24E-17
0,0125	5,87E-04	5,55E-07	4,74E-10	1,37E-12	4,74E-15	4,18E-17	1,74E-18
0,00625	2,93E-04	1,39E-07	5,95E-11	8,50E-14	1,51E-16	6,43E-19	3,10E-20
0,05	1,00	2,00	2,98	4,03	4,86	6,06	4,81
0,025	1,00	2,00	2,99	4,01	4,94	6,04	5,57
0,0125	1,00	2,00	3,00	4,00	4,97	6,02	5,81

Tabelle 7.4: Resultate zu Beispiel 7.6.

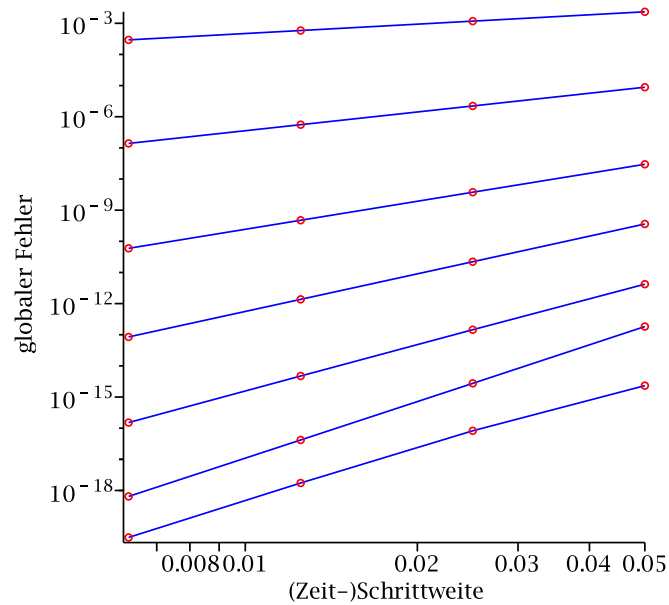


Abbildung 7.2: Der Graph zeigt (doppelt logarithmisch) die globalen Fehler aufgetragen über der Schrittweite für Beispiel 7.6.

h	IEUL	IQDeC21	IQDeC22	IQDeC23	IQDeC24
0,01	2,73E-00	6,38E-01	1,06E-01	9,96E-03	9,85E-04
0,005	1,54E-00	1,87E-01	1,47E-02	4,77E-04	7,16E-05
0,0025	8,20E-01	5,03E-02	1,88E-03	2,02E-05	2,88E-06
0,00125	4,23E-01	1,30E-02	2,36E-04	1,08E-06	1,24E-07
0,01	0,83	1,77	2,85	4,39	3,78
0,005	0,91	1,89	2,97	4,56	4,64
0,0025	0,96	1,95	3,00	4,22	4,54
0,00125					

Tabelle 7.5: Resultate zu Beispiel 7.7.

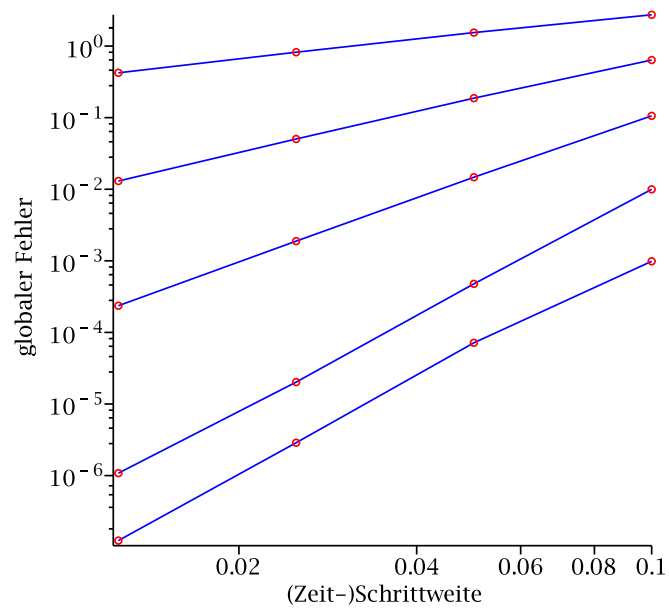


Abbildung 7.3: Der Graph zeigt (doppelt logarithmisch) die globalen Fehler aufgetragen über der Schrittweite für Beispiel 7.7.

Anhang A

Maple-Code von IIPDeC2-DQ2-Lobatto

Es folgt ein kommentierter Maple-Code des IIPDeC2-DQ2-Lobatto-Algorithmus. Siehe dazu Abschnitt 4.2.

```
>restart;
>with(plots):
>printlevel:=4;
>nber:=4; # Anzahl der Gitter, für welche die Konvergenzordnungen berechnet
  werden
>m:=6; # Anzahl der äquidistanten Kollokationsknoten
  (Anzahl der Gitterpunkte = m+1)
>mr:=m-1; # Anzahl der Lobatto-Knoten
>nymax:=4; # Anzahl der Iterationen
>tend:=3.; # Integrationsdauer (beginnend bei Null)
>hmax:=0.1; # Teilintervalllänge I am größten Gitter
>N:=round(tend/hmax); # Anzahl der Intervalle
>f(1):=(t,y1,y1dt,y1ddt)->exp(y1ddt)+y1dt+y1-exp(-sin(t))-cos(t)-1-sin(t); #
  zu lösende Differentialgleichung
loes(1):=t->1+sin(t); # exakte Lösung
eta  (1,1,1,1,1):=1; # Anfangswerte
etadot(1,1,1,1,1):=1;
>if (mr=4) then # relative Lage der Lobatto-Knoten
  cr(2):=0;
  cr(3):=evalf((5-sqrt(5))/10);
  cr(4):=evalf((5+sqrt(5))/10);
  cr(5):=1;
end if;
if (mr=5) then
  cr(2):=0.;
  cr(3):=evalf(1/2-sqrt(21)/14);
  cr(4):=evalf(1/2);
  cr(5):=evalf(1/2+sqrt(21)/14);
  cr(6):=1.;
end if;
>if m=5 then # relative Lage der äquidistanten Knoten
  c(1):=0;
```

```

c(2):=1/5;
c(3):=2/5;
c(4):=3/5;
c(5):=4/5;
c(6):=1;
end if;
if m=6 then
c(1):=0;
c(2):=1/6;
c(3):=1/3;
c(4):=1/2;
c(5):=2/3;
c(6):=5/6;
c(7):=1;
end if;
>for iber from 1 by 1 to nber do # Vervielfältigung der Anfangswerte für
das Basisverfahren und IIPDeC2-DQ2
for ny from 1 by 1 to (nymax+1) do
eta (1,1,1,ny,iber):=eta (1,1,1,1,1);
pi (1,1,1,ny,iber):=eta (1,1,1,1,1);
etadot(1,1,1,ny,iber):=etadot(1,1,1,1,1);
pidot (1,1,1,ny,iber):=etadot(1,1,1,1,1);
end do;
end do:
>for iber from 1 by 1 to nber do # Vervielfältigung der Anfangswerte für
das Kollokationsverfahren
z (1,1,iber):=eta (1,1,1,1,1);
zdt(1,1,iber):=etadot(1,1,1,1,1);
end do:
>for i from 2 by 1 to (m+1) do # Aufstellen der Lagrange-Polynome für
das Kollokationsverfahren
if (i=2) then
lc(i):=mul((t-c(j))/(c(i)-c(j)),j=i+1..m+1);
end if;
if ((i>2) and (i<m+1)) then
lc(i):=mul((t-c(j))/(c(i)-c(j)),j=2..i-1)
*mul((t-c(j))/(c(i)-c(j)),j=i+1..m+1);
end if;
if (i=m+1) then
lc(i):=mul((t-c(j))/(c(i)-c(j)),j=2..m);
end if;
end do:
>for i from 1 by 1 to (m+1) do # Aufstellen der Lagrange-Polynome für poly(t)
if (i=1) then
lp(i):=mul((t-c(j))/(c(i)-c(j)),j=i+1..m+1);
end if;
if ((i>1) and (i<m+1)) then

```

```

        lp(i):=mul((t-c(j))/(c(i)-c(j)),j=1..i-1)
        *mul((t-c(j))/(c(i)-c(j)),j=i+1..m+1);
    end if;
    if (i=m+1) then
        lp(i):=mul((t-c(j))/(c(i)-c(j)),j=1..m);
    end if;
end do:
>for i from 2 by 1 to (mr+1) do # Aufstellen der Lagrange polynome für
die Lobatto-Knoten
    if (i=2) then
        lr(i):=mul((t-cr(j))/(cr(i)-cr(j)),j=i+1..mr+1);
    end if;
    if ((i>2) and (i<mr+1)) then
        lr(i):=mul((t-cr(j))/(cr(i)-cr(j)),j=2..i-1)
        *mul((t-cr(j))/(cr(i)-cr(j)),j=i+1..mr+1);
    end if;
    if (i=mr+1) then
        lr(i):=mul((t-cr(j))/(cr(i)-cr(j)),j=2..mr);
    end if;
end do:
>for iber from 1 by 1 to nber do # IIPDeC2-DQ2

    h:=hmax/2^(iber-1); # Berechnung der Intervalllängen
    ny:=1;

    for j from 1 by 1 to N*2^(iber-1) do # Basisverfahren - Lösung
        for l from 2 by 1 to (m+1) do

            etahilf:='etahilf';

            if (l=2) then
                if (j=1) then # Startschritt

                    glg1:=f(1)(h*((j-1)+c(l-1)),eta(1,j,l-1,ny,iber),
                    etadot(1,j,l-1,ny,iber),((etahilf-eta(1,j,l-1,ny,iber))/
                    (h*(c(l)-c(l-1))))
                    -etadot(1,j,l-1,ny,iber))/((h*(c(l)-c(l-1)))/2));

                    eta(1,j,l,ny,iber):=fsolve(glg1,etahilf,fulldigits);

                end if;

                if (j>1) then # Schritt über eine Intervallgrenze hinweg

                    glg1:=f(1)(h*((j-1)+c(l-1)),eta(1,j,l-1,ny,iber),
                    (etahilf-eta(1,j-1,m,ny,iber))/(h*(c(l)-c(l-1)+c(m+1)-c(m))),
                    ((etahilf-eta(1,j,l-1,ny,iber))/(h*(c(l)-c(l-1))))

```



```

        -(eta(1,j,l-1,ny,iber)-eta(1,j-1,m,ny,iber))/
        (h*(c(m+1)-c(m))))/(h/2*(c(2)-c(1)+c(m+1)-c(m)))));

        eta(1,j,l,ny,iber):=fsolve(glg1,etahilf,fulldigits);

    end if;
end if;

if (l>2) then # Schritt für einen inneren Knoten

    glg1:=f(1)(h*((j-1)+c(l-1)),eta(1,j,l-1,ny,iber),
    (etahilf-eta(1,j,l-2,ny,iber))/(h*(c(1)-c(l-2))),
    ((etahilf-eta(1,j,l-1,ny,iber))/(h*(c(1)-c(l-1))))
    -(eta(1,j,l-1,ny,iber)-eta(1,j,l-2,ny,iber))/
    (h*(c(l-1)-c(l-2))))/(h/2*(c(1)-c(l-2)))));

    eta(1,j,l,ny,iber):=fsolve(glg1,etahilf,fulldigits);

end if;

if (l=m+1) then
    eta(1,j+1,1,ny,iber):=eta(1,j,l,ny,iber);
end if;

end do;
end do;

for ny from 1 by 1 to (nymax+1) do

    for j from 1 by 1 to N*2^(iber-1) do # Aufstellen der Interpolanten
    und deren Ableitungen

        poly    (1,j,ny,iber):=sum(eta(1,j,phi,ny,iber)
        *lp(phi),phi=1..(m+1));
        polyfkt(1,j,ny,iber):=unapply(poly(1,j,ny,iber),t);
        polydt (1,j,ny,iber):=unapply(diff(poly(1,j,ny,iber)/h,t),t);
        polyddt(1,j,ny,iber):=unapply(diff(diff(poly(1,j,ny,iber)/h^2,t),t),t);

    end do;

    for j from 1 by 1 to N*2^(iber-1) do # Berechnung des Defekts
        for l from 2 by 1 to (mr+1) do

            dquer(1,j,l,ny,iber):=evalf(f(1)(h*((j-1)+cr(l)),
            polyfkt(1,j,ny,iber)(cr(l)),polydt(1,j,ny,iber)(cr(l)),
            polyddt(1,j,ny,iber)(cr(l))));

```

```

    if ((j=1) and (l=2)) then # Berechnung des B-Teils
    des Startschritt-Defekts

        dquerB(1,j,l-1,ny,iber):=evalf(polydt(1,j,ny,iber)(c(l-1))
        -etadot(1,j,l-1,ny,iber));

    end if;

end do;
end do;

for j from 1 by 1 to N*2^(iber-1) do # Interpolation des Defekts

    polydefr (1,j,ny,iber):=sum(dquer(1,j,phi,ny,iber)
    *lr(phi),phi=2..(mr+1));
    polydeffktr(1,j,ny,iber):=unapply(polydefr(1,j,ny,iber),t);

end do;

for j from 1 by 1 to N*2^(iber-1) do
    for l from 1 by 1 to (m+1) do # Auswertung für einen inneren Knoten

        dquerr(1,j,l,ny,iber):=evalf(polydeffktr(1,j,ny,iber)(c(l)));

    end do;
end do;

for j from 2 by 1 to N*2^(iber-1) do # Auswertung an den Intervallenden
= Sprungdefekt

    dquerr(1,j,1,ny,iber):=f(1)(h*((j-1)+c(1)),eta(1,j,1,ny,iber),
    0.5*(polydt (1,j,ny,iber)(c(1))+polydt (1,j-1,ny,iber)(c(m+1))),
    (polydt (1,j,ny,iber)(c(1))-polydt (1,j-1,ny,iber)(c(m+1)))/
    (h*(c(2)-c(1)))
    +0.5*(polyddt(1,j,ny,iber)(c(1))+polyddt(1,j-1,ny,iber)(c(m+1))));

end do;

for j from 1 by 1 to N*2^(iber-1) do # Lösung des Nachbarproblems
    for l from 2 by 1 to (m+1) do

        pihilf:='pihilf';

        if (l=2) then
            if (j=1) then # Startschritt

                glg1:=f(1)(h*((j-1)+c(l-1)),pi(1,j,l-1,ny,iber),

```

```

        pidot(1,j,l-1,ny,iber)+dquerB(1,j,l-1,ny,iber),
        ((pihilf-pi(1,j,l-1,ny,iber))/(h*(c(l)-c(l-1))))
        -(pidot(1,j,l-1,ny,iber)+dquerB(1,j,l-1,ny,iber)))/
        ((h*(c(l)-c(l-1)))/2))
        -dquerr(1,j,l-1,ny,iber);

        pi(1,j,l,ny,iber):=fsolve(glg1,pihilf,fulldigits);

    end if;

    if (j>1) then # Schritt über eine Intervallgrenze hinweg

        glg1:=f(1)(h*((j-1)+c(l-1)),pi(1,j,l-1,ny,iber),
        (pihilf-pi(1,j-1,m,ny,iber))/(h*(c(l)-c(l-1)+c(m+1)-c(m))),
        ((pihilf-pi(1,j,l-1,ny,iber))/(h*(c(l)-c(l-1))))
        -(pi(1,j,l-1,ny,iber)-pi(1,j-1,m,ny,iber))/
        (h*(c(m+1)-c(m))))/(h/2*(c(2)-c(1)+c(m+1)-c(m))))
        -dquerr(1,j,l-1,ny,iber);

        pi(1,j,l,ny,iber):=fsolve(glg1,pihilf,fulldigits);

    end if;
end if;

if (l>2) then # Schritt für einen inneren Knoten

    glg1:=f(1)(h*((j-1)+c(l-1)),pi(1,j,l-1,ny,iber),
    (pihilf-pi(1,j,l-2,ny,iber))/(h*(c(l)-c(l-2))),
    ((pihilf-pi(1,j,l-1,ny,iber))/(h*(c(l)-c(l-1))))
    -(pi(1,j,l-1,ny,iber)-pi(1,j,l-2,ny,iber))/
    (h*(c(l-1)-c(l-2))))/(h/2*(c(l)-c(l-2))))
    -dquerr(1,j,l-1,ny,iber);

    pi(1,j,l,ny,iber):=fsolve(glg1,pihilf,fulldigits);

end if;

    if (l=m+1) then
        pi(1,j+1,1,ny,iber):=pi(1,j,l,ny,iber);
    end if;

end do;
end do;

for j from 1 by 1 to N*2^(iber-1) do # Ausführen eines
Iterationsschritts von \nu nach \nu+1
    for l from 2 by 1 to (m+1) do

```

```

        eta(1,j,l,ny+1,iber):=eta(1,j,l,1,iber)
        -(pi(1,j,l,ny,iber)-eta(1,j,l,ny,iber));

        if (l=m+1) then
            eta(1,j+1,1,ny+1,iber):=eta(1,j,l,ny+1,iber);
        end if;

    end do;
end do;

end do;
end do:
>for i from 2 by 1 to (mr+1) do # Berechnung der Butcher-Koeffizienten
für die Kollokationslösung
    for j from 2 by 1 to (mr+1) do
        a(i,j):=int(lr(j),t=0..cr(i));
    end do;
end do:
>for i from 2 by 1 to (mr+1) do
    b(i):=int(lr(i),t=0..1);
end do:
>for i from 2 by 1 to (mr+1) do
    for j from 2 by 1 to (mr+1) do
        aquer(i,j):=int((cr(i)-t)*lr(j),t=0..cr(i));
    end do;
end do:
>for i from 2 by 1 to (mr+1) do
    bqwer(i):=int((1-t)*lr(i),t=0..1);
end do:
>for iber from 1 by 1 to nber do # Berechnung der Kollokationslösung
    h:=hmax/2^(iber-1);
    for j from 2 by 1 to (N*2^(iber-1)+1) do

        for i from 2 by 1 to (mr+1) do
            glgc(i):=f(1)(h*((j-2)+cr(i)),
                z(1,j-1,iber)+h*cr(i)*zdt(1,j-1,iber)
                +h^2*sum(kc[my]*aquar(i,my),my=2..(mr+1)),
                zdt(1,j-1,iber)
                +h *sum(kc[my]*a(i,my),my=2..(mr+1)),kc[i]);
        end do;

        if (mr=4) then loesung:=fsolve({glgc(2),glgc(3),glgc(4),glgc(5)},
            {kc[2],kc[3],kc[4],kc[5]}); end if;
        if (mr=5) then loesung:=fsolve({glgc(2),glgc(3),glgc(4),glgc(5),glgc(6)},
            {kc[2],kc[3],kc[4],kc[5],kc[6]}); end if;
    end do;
end do;

```

```

for i from 2 by 1 to (mr+1) do
    kcc(i):=rhs(op(select(has,loesung,kc[i])));
end do;

z(1,j,iber) :=z (1,j-1,iber)+h*zdt(1,j-1,iber)
+h^2*sum(bquer(my)*kcc(my),my=2..(mr+1));
zdt(1,j,iber):=zdt(1,j-1,iber)
+h *sum(b(my)*kcc(my),my=2..(mr+1));

end do;
end do:
>h:=hmax/2^(nber-1): # Erzeugung der Daten für die graphische Ausgabe
von IIPDeC2-DQ2
>i:=1:
>etaout(1,i,nymax,nber):=eta(1,1,1,nymax,nber):
>tout(i):=0:
>i:=2:
>for j from 1 by 1 to N*2^(nber-1) do
    for l from 2 by 1 to (m+1) do
        tout(i):=h*((j-1)+c(l)):
        etaout(1,i,nymax,nber):=eta(1,j,l,nymax,nber);
        i:=i+1;
    end do;
end do:
>etaoutlist1:=seq([tout(ii),etaout(1,ii,nymax,nber)],ii=1..N*m*2^(nber-1)+1):
>h:hmax/2^(nber-1): # Erzeugung der Daten für die graphische Ausgabe
des Kollokationsverfahrens
>zlist1:=seq([(j-1)*h,z(1,j,nber)],j=1..N*2^(nber-1)+1):
>p1:=plot(loes(1)(t),t=0..tend,legend=["y_exakt"]):
>p2:=plot([etaoutlist1],[zlist1],legend=["IIPDeC2-DQ2","COLL"],
color=[blue,green]):
>display([p1,p2]);
>y1exakt(1):=loes(1)(tend); # Berechnung der exakten Lösung zum
Zeitpunkt t_end
>"BV";
>for k from 2 by 1 to nber do # Berechnung der globalen Fehler und
der Konvergenzordnungen für das Basisverfahren
    hi(k-1):=hmax/2^(k-2);
    hi(k) :=hmax/2^(k-1);
    e(1,k-1):=eta(1,N*2^(k-2),m+1,1,k-1)-y1exakt(1);
    eges(k-1):=evalf(sqrt(e(1,k-1)^2));
    e(1,k) :=eta(1,N*2^(k-1),m+1,1,k )-y1exakt(1);
    eges(k):=evalf(sqrt(e(1,k)^2));
    p(k):=evalf(ln(eges(k-1)/eges(k))/ln(hi(k-1)/hi(k)));
    print(eges(k-1),eges(k),p(k));
end do:

```

```

>"IIPDEC2-DQ2-Lobatto";
>for ny from 2 by 1 to (nymax+1) do
  print(ny-1);
  for k from 2 by 1 to nber do # Berechnung der globalen Fehler und
  der Konvergenzordnungen für IIPDeC2-DQ2
    hi(k-1):=hmax/2^(k-2);
    hi(k) :=hmax/2^(k-1);
    e(1,k-1):=eta(1,N*2^(k-2),m+1,ny,k-1)-y1exakt(1);
    eges(k-1):=evalf(sqrt(e(1,k-1)^2));
    e(1,k) :=eta(1,N*2^(k-1),m+1,ny,k )-y1exakt(1);
    eges(k):=evalf(sqrt(e(1,k)^2));
    p(k):=evalf(ln(eges(k-1)/eges(k))/ln(hi(k-1)/hi(k)));
    print(eges(k-1),eges(k),p(k));
  end do;
end do:
>"COLL";
>for k from 2 by 1 to nber do # Berechnung der globalen Fehler und
der Konvergenzordnungen für das Kollokationsverfahren
  hi(k-1):=hmax/2^(k-2);
  hi(k) :=hmax/2^(k-1);
  e(1,k-1):=z(1,N*2^(k-2)+1,k-1)-y1exakt(1);
  eges(k-1):=evalf(sqrt(e(1,k-1)^2));
  e(1,k) :=z(1,N*2^(k-1)+1,k )-y1exakt(1);
  eges(k):=evalf(sqrt(e(1,k)^2));
  p(k):=evalf(ln(eges(k-1)/eges(k))/ln(hi(k-1)/hi(k)));
  print(eges(k-1),eges(k),p(k));
end do:

```

Literaturverzeichnis

- [1] U.M. Ascher, R.M.M. Mattheij, R.D. Russell, *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations*, Prentice-Hall International, 1988.
- [2] W. Auzinger, H. Hofstätter, W. Kreuzer, E. Weinmüller, *Modified Defect Correction Algorithms for ODEs. Part I: General Theory*, Erschienen in Numerical Algorithms 36, Preprint online unter: <http://www.math.tuwien.ac.at/~winfried/papers/anumpp0203.pdf>, 21.12.2012.
- [3] W. Auzinger, O. Koch, E. Weinmüller, *Efficient Collocation Schemes for Singular Boundary Value Problems*, Numerical Algorithms 85, S.5-25, 2002.
- [4] W. Auzinger, *Numerical Solutions of Ordinary Differential Equations*, Skriptum zur gleichnamigen Vorlesung, Technische Universität Wien, 2003.
- [5] W. Auzinger, H. Hofstätter, W. Kreuzer, E. Weinmüller, *Modified Defect Correction Algorithms for ODEs. Part II: Stiff Initial Value Problems*, Numerical Algorithms 40, S.285-303, 2005.
- [6] W. Auzinger, D. Praetorius, *Numerische Mathematik*, Skriptum zur gleichnamigen Vorlesung, Technische Universität Wien, 2007.
- [7] W. Auzinger, *Error Estimation Techniques Based on Defect Computation and Global Reconstruction*, Proceedings of the International Conference on Numerical Analysis and Applied Mathematics, Rhodos, Greece, September 2010 (AIP Conference Proceedings 1281), Vol.1, S.210-213, 2010.
- [8] W. Auzinger, *Defect Correction Techniques for a Posteriori Error Estimation*, ICIAM-Vortragsfolien, Vancouver, 18.-22. Juli 2011.
- [9] W. Auzinger, *Defect Correction Methods*, In Encyclopedia of Applied and Computational Mathematics, Springer-Verlag, online unter: <http://www.springerreference.com/docs/html/chapterdbid/333743.html>, 21.12.2012.
- [10] W. Auzinger, M. Tutz, *A review of stability and error theory for collocation methods applied to linear boundary value problems*, ASC Report Nr. 35, Technische Universität Wien, 2012.
- [11] W. Auzinger, *Notizen zu Mehrschrittverfahren*, persönliche Kommunikation, 2012.
- [12] R. Frank, *The Method of Iterated Defect-Correction and its Application to Two-Point Boundary Value Problems, Part I*, Numerische Mathematik, S.409-419, 1975.

- [13] R. Frank, C. Überhuber, *Iterated Defect Correction for Runge-Kutta-Methods*, Bericht, Institut für Numerische Mathematik, Technische Universität Wien, 1975.
- [14] R. Frank, C. Überhuber, *Collocation and Iterated Defect Correction*, Bericht, Institut für Numerische Mathematik, Technische Universität Wien, 1976.
- [15] R. Frank, C. Überhuber, *Iterated Defect Correction for Differential Equations, Part I: Theoretical Results*, Computing 20, S.207-228, 1978.
- [16] C. Großmann, H.-G. Roos, *Numerik partieller Differentialgleichungen*, Teubner Verlag, 1992.
- [17] E. Hairer, S.P. Nørsett, G. Wanner, *Solving Ordinary Differential Equations I. Nonstiff Problems*, 2nd ed., Springer-Verlag, 1993.
- [18] B. Hulme, *One-Step Piecewise Polynomial Galerkin Methods for Initial Value Problems*, Mathematics of Computation 118, S.415-426, 1972.
- [19] G. Kitzler, *A posteriori Fehlerschätzer für Zweipunkt-Randwertprobleme mittels Defektkorrektur*, Diplomarbeit, Technische Universität Wien, 2010.
- [20] L.M. Milne-Thomson, *The calculus of finite differences*, 2nd ed., Chelsea Publishing Company, 1981.
- [21] A. Saboor Bagherzadeh, *Defect-Based Error Estimation for Higher Order Differential Equations*, Dissertation, Technische Universität Wien, 2011.
- [22] H.J. Stetter, *The Defect Correction Principle and Discretization Methods*, Numerische Mathematik 29, S.425-443, 1978.
- [23] P.E. Zadunaisky, *On the Estimation of Errors Propagated in the Numerical Integration of ODEs*, Numerische Mathematik 27, S.21-39, 1976.