

ZEITBEREICHS-MODELLREDUKTION UND SYLVESTER-GLEICHUNGEN

An der Fakultät für Mathematik
der Otto-von-Guericke-Universität Magdeburg
zur Erlangung des akademischen Grades
Master of Science
angefertigte

MASTERARBEIT

vorgelegt von
MANUELA HUND
geboren am 26.03.1991 in Nordhausen,
Studiengang Mathematik,
Studienrichtung Mathematik.

23. April 2015

Erstgutachter: DR. JENS SAAK
Zweitgutachter: PROF. DR. PETER BENNER

Inhaltsverzeichnis

Abbildungsverzeichnis	VI
Tabellenverzeichnis	VII
Symbolverzeichnis	IX
1 Einleitung	1
2 Grundlagen	4
2.1 Orthogonale Polynome	4
2.2 Matrixgleichungen	16
2.3 Modellreduktion durch Momentenabgleich	23
3 Modellreduktion von LTI-Systemen auf Basis von orthogonalen Polynomen	28
3.1 Grundidee	28
3.2 Algorithmus zur Modellreduktion	29
4 Lösen des linearen Gleichungssystems	32
4.1 Lösen mit Hilfe eines Splitting-Verfahrens	32
4.2 Lösen mit Hilfe eines vorkonditionierten Krylov-Unterraum-Verfahrens	33
4.2.1 GMRES-Verfahren	33
4.2.2 Vorkonditionierung	34
4.3 Umformulierung des Gleichungssystems als Sylvester-Gleichung	36
4.4 Varianten der Modellreduktion	39
4.4.1 Variante 1: Ausnutzen der gegebenen Bedingungen	39
4.4.2 Variante 2: Bedingungen an den Anfangswert	41
5 Experimentelle Auswertung	45
5.1 Rahmenbedingungen der Auswertung	45
5.2 Auswertung	48
5.2.1 Beispiel 1	48
5.2.2 Beispiel 2	54

5.2.3	Beispiel 3	57
5.2.4	Beispiel 4	60
5.2.5	Beispiel 5	65
5.3	Erklärung	67
6	Zusammenfassung und Ausblick	77
6.1	Zusammenfassung	77
6.2	Ausblick	79
	Literaturverzeichnis	81
	Anhang	i
A	Berechnung der Differentialrekursionsformel für einzelne orthogonale Polynome	i
A.1	Legendre-Polynome	i
A.2	Jacobi-Polynome	ii
A.3	Chebychev-Polynome erster Art	v
A.4	Chebychev-Polynome zweiter Art	vi
B	Weitere Abbildungen für die Auswertung	vii
B.1	Beispiel 1	viii
B.2	Beispiel 2	x
B.3	Beispiel 3	xii
B.4	Beispiel 4	xiv
B.5	Beispiel 5	xvii
C	Explizite Darstellung der Inversen von \tilde{E}	xix
C.1	Hermite-Polynome	xix
C.2	Laguerre-Polynome	xx
C.3	Struktur der Inversen für Legendre- und Chebychev-Polynome	xxi
C.4	Legendre-Polynome	xxiv
C.5	Chebychev Polynome erster Art	xxx
C.6	Chebychev Polynome zweiter Art	xxxiv
D	Weitere Abbildungen zu den verallgemeinerten Eigenwerten	xl
D.1	Verallgemeinerte Eigenwerte für die äquivalente Sylvester-Gleichung	xl
D.2	Verallgemeinerte Eigenwerte für die zusatzfreie Variante 2	xliv
E	CD	xlix

Abbildungsverzeichnis

5.1	Relativer Fehler durch Reduktion mit Hilfe vorgestellter Verfahren (Beispiel 1)	49
5.2	Kondition der H - und \hat{H} -Matrix für Beispiel 1	50
5.3	Trajektorien für die Anfangswert-unabhängige Variante 1 im Vergleich zur Originaltrajektorie für Beispiel 1 ($m = 40$)	51
5.4	Vergleich der Gesamtzeiten durch Reduktion mit Hilfe vorgestellter Verfahren (Beispiel 1)	53
5.5	Originaltrajektorie für Beispiel 2	55
5.6	Relativer Fehler durch die verbleibenden Lösungsstrategien (Beispiel 2)	56
5.7	Vergleich der Gesamtzeit verbleibender Lösungsstrategien (Beispiel 2) .	57
5.8	Relativer Fehler durch die verbleibenden Lösungsstrategien (Beispiel 3)	58
5.9	Trajektorie für die zusatzfreie Variante 2 (Hermite-Polynome) im Vergleich zur Originaltrajektorie für Beispiel 3 ($m = 40$)	59
5.10	Vergleich der Gesamtzeit verbleibender Lösungsstrategien (Beispiel 3) .	60
5.11	Relativer Fehler durch die verbleibenden Lösungsstrategien (Beispiel 4)	61
5.12	Trajektorien für Backslash und zusatzfreie Variante 2 (Hermite-, Laguerre-Polynome) im Vergleich zur Originaltrajektorie für Beispiel 4 ($m = 30$)	62
5.13	Vergleich der Gesamtzeit verbleibender Lösungsstrategien (Beispiel 4) .	63
5.14	Relativer Fehler für $m = 20, \dots, 150$ durch die verbleibenden Lösungsstrategien (Beispiel 4)	64
5.15	Originaltrajektorie für Beispiel 5	65
5.16	Relativer Fehler durch die verbleibenden Lösungsstrategien (Beispiel 5)	66
5.17	Vergleich der Gesamtzeit verbleibender Lösungsstrategien (Beispiel 5) .	67
5.18	Rang der Beobachtbarkeitsmatrix für Sylvester-Gleichung (4.4)	71
5.19	Rang der Beobachtbarkeitsmatrix für zusatzfreie Sylvester-Gleichung (4.12)	72
5.20	Endliche Eigenwerte des Matrixpaares $(-\tilde{A}, \tilde{E})$ für $m = 15, 30, 45, 60$. .	73
5.21	Endliche Eigenwerte des Matrixpaares $(-I_m, \mathcal{E})$ für $m = 15, 30, 45, 60$.	74
1	Absoluter Fehler durch Reduktion mit Hilfe vorgestellter Verfahren (Beispiel 1)	viii

2	Trajektorien für das GMRES-Verfahren mit beiden Vorkonditionierern für Beispiel 1	ix
3	Absoluter Fehler durch Reduktion mit Hilfe der verbleibenden Lösungsstrategien (Beispiel 2)	x
4	Relativer Fehler - Vergleich (Beispiel 2)	xi
5	Ausführungszeit - Vergleich (Beispiel 2)	xi
6	Absoluter Fehler durch Reduktion mit Hilfe der verbleibenden Lösungsstrategien (Beispiel 3)	xii
7	Relativer Fehler - Vergleich (Beispiel 3)	xiii
8	Ausführungszeit - Vergleich (Beispiel 3)	xiii
9	Absoluter Fehler durch Reduktion mit Hilfe der verbleibenden Lösungsstrategien (Beispiel 4)	xiv
10	Relativer Fehler - Vergleich (Beispiel 4)	xv
11	Ausführungszeit - Vergleich (Beispiel 4)	xv
12	Absoluter Fehler für $m = 20, \dots, 150$ durch die verbleibenden Lösungsstrategien (Beispiel 4)	xvi
13	Ausführungszeit für $m = 20, \dots, 150$ durch die verbleibenden Lösungsstrategien (Beispiel 4)	xvi
14	Absoluter Fehler durch Reduktion mit Hilfe der verbleibenden Lösungsstrategien (Beispiel 5)	xvii
15	Relativer Fehler - Vergleich (Beispiel 5)	xviii
16	Ausführungszeit - Vergleich (Beispiel 5)	xviii
17	Endliche Eigenwerte des Matrixpaares $(-\tilde{A}, \tilde{E})$ für Hermite-Polynome für $m = 15, 30, 45, 60$	xl
18	Endliche Eigenwerte des Matrixpaares $(-\tilde{A}, \tilde{E})$ für Laguerre-Polynome für $m = 15, 30, 45, 60$	xli
19	Endliche Eigenwerte des Matrixpaares $(-\tilde{A}, \tilde{E})$ für Legendre-Polynome für $m = 15, 30, 45, 60$	xlii
20	Endliche Eigenwerte des Matrixpaares $(-\tilde{A}, \tilde{E})$ für Chebychev-Polynome erster Art für $m = 15, 30, 45, 60$	xliii
21	Endliche Eigenwerte des Matrixpaares $(-\tilde{A}, \tilde{E})$ für Chebychev-Polynome zweiter Art für $m = 15, 30, 45, 60$	xliv
22	Endliche Eigenwerte des Matrixpaares $(-I_m, \mathcal{E})$ für Laguerre-Polynome für $m = 15, 30, 45, 60$	xlv
23	Endliche Eigenwerte des Matrixpaares $(-I_m, \mathcal{E})$ für Legendre-Polynome für $m = 15, 30, 45, 60$	xlvi
24	Endliche Eigenwerte des Matrixpaares für $(-I_m, \mathcal{E})$ Chebychev-Polynome erster Art für $m = 15, 30, 45, 60$	xlvii
25	Endliche Eigenwerte des Matrixpaares $(-I_m, \mathcal{E})$ für Chebychev-Polynome zweiter Art für $m = 15, 30, 45, 60$	xlviii

Tabellenverzeichnis

2.1	Differentialrekursionskoeffizienten	16
-----	---	----

Symbolverzeichnis

\mathbb{N}	Menge der natürlichen Zahlen $(1, 2, 3, \dots)$
\mathbb{N}_0	Menge der natürlichen Zahlen einschließlich 0 $(0, 1, 2, 3, \dots)$
\mathbb{R}	Menge der reellen Zahlen
\mathbb{R}^n	Vektorraum der reellwertigen n -dimensionalen Tupel
$\mathbb{R}^{n \times m}$	Menge der reellwertigen $(n \times m)$ -Matrizen
\mathbb{C}	Menge der komplexen Zahlen
\mathbb{C}^n	Vektorraum der komplexwertigen n -dimensionalen Tupel
$\mathbb{C}^{n \times m}$	Menge der komplexwertigen $(n \times m)$ -Matrizen
$L^2([a, b])$	Hilbertraum der auf dem Intervall $[a, b]$ definierten Lebesgue-meßbaren und quadratintegrierbaren Funktionen
$\operatorname{Re}(x)$	Realteil von $x \in \mathbb{C}$
$\operatorname{Im}(x)$	Imaginärteil von $x \in \mathbb{C}$
\bar{x}	komplex konjugierte Zahl zu $x \in \mathbb{C}$
$\ x\ _2$	$:= \sqrt{\sum_{i=1}^n x_i ^2}$ Euklidische Norm des Vektors $x \in \mathbb{R}^n$ oder $x \in \mathbb{C}^n$
I_n	$:= \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$ $(n \times n)$ -Einheitsmatrix
a_{ij}	(i, j) -ter Eintrag der Matrix $A \in \mathbb{R}^{n \times m}$ oder $A \in \mathbb{C}^{n \times m}$
A^T	Transponierte der Matrix $A \in \mathbb{R}^{n \times m}$
A^H	$:= \overline{a_{ji}}$ komplex konjugierte Transponierte der Matrix $A \in \mathbb{C}^{n \times m}$

A^{-1}	Inverse einer Matrix $A \in \mathbb{R}^{n \times n}$ oder $A \in \mathbb{C}^{n \times n}$
$\det(A)$	Determinante einer Matrix $A \in \mathbb{R}^{n \times n}$ oder $A \in \mathbb{C}^{n \times n}$
$\text{rank}(A)$	Rang einer Matrix $A \in \mathbb{R}^{n \times m}$ oder $A \in \mathbb{C}^{n \times m}$
$\text{span}(A)$	Spann einer Matrix $A \in \mathbb{R}^{n \times m}$ oder $A \in \mathbb{C}^{n \times m}$
$\text{colspan}(A)$	Spaltenspann einer Matrix $A \in \mathbb{R}^{n \times m}$ oder $A \in \mathbb{C}^{n \times m}$
$A \otimes B$	Kronecker-Produkt zweier Matrizen $A \in \mathbb{R}^{n \times m}$ und $B \in \mathbb{R}^{p \times q}$
$\sigma_{\max}(A)$	maximaler Singulärwert einer Matrix $A \in \mathbb{R}^{n \times n}$ oder $A \in \mathbb{C}^{n \times n}$
$\Lambda(A)$	Spektrum einer Matrix $A \in \mathbb{R}^{n \times n}$ oder $A \in \mathbb{C}^{n \times n}$
$\Lambda(A, E)$	Spektrum eines Matrixpaares $A, E \in \mathbb{R}^{n \times n}$ oder $A, E \in \mathbb{C}^{n \times n}$
$\ A\ _1$	$:= \max_{1 \leq j \leq m} \sum_{i=1}^n a_{ij} $ Spaltensummennorm der Matrix $A \in \mathbb{R}^{n \times m}$ oder $A \in \mathbb{C}^{n \times m}$
$\ A\ _F$	$:= \sqrt{\sum_{i=1}^n \sum_{j=1}^m a_{ij} ^2}$ Frobeniusnorm der Matrix $A \in \mathbb{R}^{n \times m}$ oder $A \in \mathbb{C}^{n \times m}$
$\text{cond}_1(A)$	$= \ A\ _1 \ A^{-1}\ _1$ Kondition der Matrix $A \in \mathbb{R}^{n \times m}$ oder $A \in \mathbb{C}^{n \times m}$
$\ G(s)\ _{\mathcal{H}_2}$	$:= \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} \ G(iw)\ _F^2 dw \right)^{\frac{1}{2}}$ \mathcal{H}_2 -Norm einer Übertragungsfunktion $G(s)$
$\ G(s)\ _{\mathcal{H}_\infty}$	$:= \max_{w \in \mathbb{R}} \sigma_{\max}(G(iw))$ \mathcal{H}_∞ -Norm einer Übertragungsfunktion $G(s)$
$\frac{\partial^n}{\partial x_i^n} f$	n -te partielle Ableitung der skalaren Funktion $f = f(x_1, \dots, x_n)$ nach x_i
$\dot{x}(t)$	$:= \frac{\partial}{\partial t} x(t)$ erste Ableitung der skalaren Funktion $x(t)$ nach der Zeit t
$\ddot{x}(t)$	$:= \frac{\partial^2}{\partial t^2} x(t)$ zweite Ableitung der skalaren Funktion $x(t)$ nach der Zeit t
Δf	$:= \sum_{i=1}^n \frac{\partial^2}{\partial x_i^2} f$ Laplace-Operator für eine skalare Funktion $f = f(x_1, \dots, x_n)$

Kapitel 1

Einleitung

Viele mathematische und physikalische Anwendungsbeispiele werden durch lineare zeitinvariante Eingangs-Ausgangs-Systeme modelliert, deren Ordnung oftmals sehr groß ist und weit über 10^3 hinausgeht. Daher ist eine numerische Simulation solcher Systeme aufgrund des immensen Rechenaufwandes und Speicherbedarfes sehr aufwändig oder sogar unmöglich. Diese Modelle besitzen die folgende Zustandsraumdarstellung

$$\begin{aligned} E\dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t), \end{aligned} \tag{1.1}$$

in der die Vektoren $x(t) \in \mathbb{R}^n$ als Zustands-, $u(t) \in \mathbb{R}^p$ als Eingangs- und $y(t) \in \mathbb{R}^q$ als Ausgangsvektor, die Matrizen $A, E \in \mathbb{R}^{n \times n}$ als System-, $B \in \mathbb{R}^{n \times p}$ als Eingangs- und $C \in \mathbb{R}^{q \times n}$ als Ausgangsmatrix bezeichnet werden und $t \in \mathbb{R}$ die Zeit repräsentiert.

Diese Systeme werden in der folgenden Arbeit als LTI-Systeme (aus dem Englischen: linear time-invariant) bezeichnet.

Eine Möglichkeit, (1.1) dennoch schnell und präzise zu lösen, bietet die Modellreduktion. Dabei wird zu einem gegebenen LTI-System ein Modell kleinerer Ordnung gesucht, welches das dynamische Verhalten sehr gut approximiert und die Lösung in kürzerer Rechenzeit ermittelt.

Ein solches reduziertes Modell besitzt die Zustandsraumdarstellung

$$\begin{aligned} E_r \dot{x}_r(t) &= A_r x_r(t) + B_r u(t) \\ y_r(t) &= C_r x_r(t), \end{aligned} \tag{1.2}$$

in welcher der Zustandsvektor $x_r(t) \in \mathbb{R}^m$ von Dimension $m \ll n$ ist und der Ausgangsvektor $y_r(t) \in \mathbb{R}^q$ durch die kleineren Matrizen $A_r, E_r \in \mathbb{R}^{m \times m}$, $B_r \in \mathbb{R}^{m \times p}$ und $C_r \in \mathbb{R}^{q \times m}$ bestimmt wird. Aufgrund der deutlich kleineren Ordnung des reduzierten Systems ist anzunehmen, dass weniger Speicherplatz benötigt wird.

Das eigentliche Ziel der Modellreduktion ist daher, ein System höherer Ordnung durch

ein System niedrigerer Ordnung zu approximieren, sodass die strukturellen Eigenschaften erhalten bleiben und der Approximationsfehler $y(t) - y_r(t)$ in einem gegebenen Zeitintervall t oder der Fehler der Übertragungsfunktionen $G(s) - G_r(s)$ für einen Frequenzbereich s klein sind.

Die *Übertragungsfunktionen* der Systeme (1.1) beziehungsweise (1.2) beschreiben die Beziehung zwischen dem Ein- und Ausgang des jeweiligen Systems und sind für eine Frequenz s gegeben durch

$$G(s) = C(sE - A)^{-1}B, \quad (1.3)$$

$$G_r(s) = C_r(sE_r - A_r)^{-1}B_r. \quad (1.4)$$

Desweiteren ist es wünschenswert, dass die Reduktionsverfahren numerisch stabil sind.

Um Systeme wie in (1.1) effizient zu reduzieren, können beispielsweise die \mathcal{H}_2 -optimalen Modellreduktionsverfahren wie der *zweiseitige Iterationsalgorithmus* (TSIA, aus dem Englischen: two-sided iteration algorithm) (siehe [3]) und der *iterative rationale Krylov Algorithmus* (IRKA) (siehe [20], Abschnitt 4.2.3) verwendet werden, bei denen der Fehler durch die \mathcal{H}_2 -Norm (siehe [2], Kapitel 5)

$$\|G(s) - G_r(s)\|_{\mathcal{H}_2}$$

approximiert wird. Eine weitere Methode ist das *balancierte Abschneiden* (siehe [2] Abschnitt 7.2), dessen Fehler in der \mathcal{H}_∞ -Norm (siehe [2], Kapitel 5)

$$\|G(s) - G_r(s)\|_{\mathcal{H}_\infty}$$

gemessen wird. Diese Verfahren schätzen den Fehler also im Frequenzbereich ab. Eine ausführliche Darstellung dieser Verfahren und Normen ist in den referenzierten Werken dargestellt und würde über den Rahmen dieser Arbeit hinausgehen.

Diese Arbeit beschäftigt sich mit dem in [13] präsentierten Verfahren zur Zeitbereichs-Modellreduktion von LTI-Systemen (1.1), das auf orthogonale Polynome basiert. Dabei soll es ausschließlich um *Single-Input Single-Output* (SISO)-Systeme gehen. Daher sind im Folgenden $p = q = 1$. Andernfalls wird ein solches Modell als *Multiple-Input Multiple-Output* (MIMO)-System bezeichnet.

Der Fehler, der durch diese Modellreduktion entsteht, wird in dieser Arbeit nicht abgeschätzt. Dieser wurde jedoch für den Zustands- und Ausgangsvektor des originalen und reduzierten Modells in [13] beziehungsweise [21] betrachtet.

Da eine Äquivalenz zwischen einer Modellreduktion mit speziellen orthogonalen Polynomen, den Laguerre-Polynomen, die in Abschnitt 2.1 näher betrachtet werden, und einem weiteren Reduktionsverfahren, dem Momentenabgleich, in [9] hergestellt wurde, wird auch die Methode des *Momentenabgleichs* in Abschnitt 2.3 präsentiert.

Die wesentlichen Kritikpunkte des in [13] präsentierten Modellreduktionsverfahrens sind die Abhängigkeiten von der Steuerung $u(t)$ und dem Anfangswert $x(t_0) = x_0$, bei dem t_0 den Startzeitpunkt und x_0 den Startvektor definieren.

Die Steuerung wird in der Praxis üblicherweise stückweise konstant gewählt, sodass dies im Folgenden ignoriert werden kann.

Eine vom Anfangswert abhängige Modellreduktion ist aber nicht erstrebenswert, da für jeden neu gewählten Anfangswert das reduzierte System neu berechnet werden muss und dies mit zusätzlichem Aufwand verbunden ist. Im Vergleich dazu sind die Modellreduktionsmethoden des balancierten Abschneidens oder des Momentenabgleichs entweder unabhängig vom Anfangswert oder setzen $x(t_0) = \begin{bmatrix} 0 & \cdots & 0 \end{bmatrix}^T$ voraus.

Desweiteren wird in [13] ein iteratives, Gauss-Seidel ähnliches Verfahren zum Erstellen des reduzierten Modells vorgestellt, welches oft nur langsam konvergiert.

Daher sind die Ziele dieser Arbeit:

- ein schnelles und direktes Verfahren zu entwickeln, das äquivalent zu der in [13] präsentierten Methode der Zeitbereichs-Modellreduktion ist,
- eine weitere effiziente, auf orthogonalen Polynomen basierende Variante der Modellreduktion herzuleiten, die wie das balancierte Abschneiden auf den Anfangswert verzichtet oder diesen wie beim Momentenabgleich festlegt,
- die in [13] präsentierte Methode der Modellreduktion für einige orthogonale Polynome ins Verhältnis zu dem Reduktionsverfahren des Momentenabgleichs zu setzen, da dies schon für die Laguerre-Polynome gelungen ist.

Nach der einleitenden Problemanalyse und Zielsetzung werden in Kapitel 2 zunächst orthogonale Polynome allgemein mit ihren wesentlichen Eigenschaften definiert. Anschließend werden die Polynome, die in der Arbeit eine Rolle spielen, genauer vorgestellt und diverse weitere Eigenschaften hergeleitet. Im weiteren Verlauf wird die Theorie der Eigenwerte und Matrixgleichungen kurz umrissen. Am Ende dieses Kapitels wird die projektionsbasierte Modellreduktion anhand des Momentenabgleichs erläutert. In Kapitel 3 wird das in [13] beschriebene Reduktionsverfahren, das auf orthogonalen Polynomen basiert, ausführlich hergeleitet.

Im nachfolgenden Kapitel 4 werden verschiedene Möglichkeiten präsentiert, um ein reduziertes Modell (1.2) zu berechnen.

Die Auswertung dieser Varianten wird in Kapitel 5 durchgeführt. Dabei wird zunächst auf die Rahmenbedingungen eingegangen. Danach werden die einzelnen Beispiele vorgestellt, ausführlich ausgewertet und die Resultate mit Hilfe von Theoremen erklärt.

In Kapitel 6 werden die Ergebnisse dieser Arbeit zusammengefasst und weitere Untersuchungen, die in dieser Arbeit nicht mehr aufgegriffen oder getestet werden konnten, vorgeschlagen.

Kapitel 2

Grundlagen

Im Folgenden werden Grundbegriffe und Zusammenhänge erläutert, die in den nachfolgenden Kapiteln Anwendung finden.

Die Grundlagen beziehen sich zum einen auf die Theorie der orthogonalen Polynome, wobei in diesem Abschnitt auch einige orthogonale Polynome mit ihren speziellen Eigenschaften vorgestellt werden sollen.

Zum anderen sollen Grundlagen, die beim Lösen und Umformulieren eines Gleichungssystems eine entscheidende Rolle spielen, aus der Theorie der Eigenwerte und Matrixgleichungen vorgestellt werden. Im Anschluss daran wird die projektionsbasierte Modellreduktion anhand einer speziellen Methode, dem Momentenabgleich, erläutert.

2.1 Orthogonale Polynome

Um die orthogonalen Polynome charakterisieren zu können, müssen zunächst der Begriff des Skalarproduktes und des Polynoms definiert werden.

Definition 2.1. (Skalarprodukt, z.B. [4] Abschnitt 5.3.8.5)

Es sei V ein reeller Vektorraum.

Ein *Skalarprodukt* ist eine positiv definite, symmetrische Bilinearform $\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbb{R}$, das heißt es gilt für alle $x, y, z \in V$ und $\lambda \in \mathbb{R}$:

$$\langle x, x \rangle > 0 \Leftrightarrow x \neq 0 \quad (2.1)$$

$$\langle x, y \rangle = \langle y, x \rangle \quad (2.2)$$

$$\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle \quad (2.3)$$

$$\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle \quad (2.4)$$

$$\langle \lambda x, y \rangle = \lambda \langle x, y \rangle = \langle x, \lambda y \rangle. \quad (2.5)$$

Definition 2.2. (Polynom, [15] Abschnitt 5.1)

Es sei \mathbb{K} ein Körper.

Eine Funktion p_n der Form

$$p_n(x) = \sum_{j=0}^n a_j x^j = a_0 + a_1 x + \dots + a_n x^n$$

heißt ein *Polynom* über \mathbb{K} in der Variablen x , wobei die Koeffizienten $a_j \in \mathbb{K}$ für $j = 0, \dots, n$ und $n \in \mathbb{N}_0$ sind.

Der *Grad* eines Polynoms $p_n(x)$ $\deg(p_n(x))$ ist der höchste Exponent n , für den der Koeffizient a_n des Monoms $a_n x^n$ nicht null wird.

Definition 2.3. (Orthogonale Polynome, [21], [1] Abschnitt 22.1)

Eine Menge von Polynomen $g_i(x)$ mit $\deg(g_i(x)) = i$ für $i = 0, 1, \dots$ aus dem orthogonalen Polynomraum

$$L_{w(x)}^2([a, b]) = \text{span}\{g_0(x), g_1(x), \dots, g_i(x), \dots\}$$

heißt bezüglich der Gewichtsfunktion $w(x)$ *orthogonal* auf dem Intervall $a \leq x \leq b$, wenn

$$\langle g_n, g_m \rangle := \int_a^b w(x) g_n(x) g_m(x) dx = \begin{cases} \xi_n, & n = m \\ 0, & n \neq m \end{cases}, \quad (2.6)$$

wobei die Gewichtsfunktion $w(x) \geq 0$ Lebesgue-messbar ist, das heißt es gilt

$$\int_a^b w(x) dx > 0.$$

Dabei bezeichnet $\langle g_n, g_m \rangle$ das gewichtete Skalarprodukt im Raum der quadratintegrierbaren Funktionen $L_{w(x)}^2$ und ξ_i ist für $i = 0, 1, \dots$ eine von null verschiedene Konstante.

Eine wichtige Eigenschaft der orthogonalen Polynome ist deren rekursive Darstellungsweise:

Satz 2.4. (Drei-Term-Rekursion, [17] Abschnitt 3.2)

Es sei eine Menge von orthogonalen Polynomen

$$g_i(x) = k_i x^i + k_i^{(1)} x^{i-1} + k_i^{(2)} x^{i-2} + \dots + k_i^{(i)} x^0, \quad i = 0, 1, \dots \quad (2.7)$$

gegeben. Die folgende Beziehung gilt für jeweils drei aufeinanderfolgende orthogonale Polynome:

$$g_n(x) = (A_n x + B_n)g_{n-1}(x) - C_n g_{n-2}(x), \quad n = 2, 3, \dots, \quad (2.8)$$

wobei A_n, B_n, C_n für $n = 2, 3, \dots$ konstant sind und sich darstellen lassen als

$$\begin{aligned} A_n &= \frac{k_n}{k_{n-1}}, \\ B_n &= A_n \left(\frac{k_n^{(1)}}{k_n} - \frac{k_{n-1}^{(1)}}{k_{n-1}} \right), \\ C_n &= \frac{A_n}{A_{n-1}} \frac{\xi_{n-1}}{\xi_{n-2}}. \end{aligned}$$

Beweis.

Es sei $n \geq 2$.

Wird $A_n = \frac{k_n}{k_{n-1}}$ gewählt, so gilt durch Einsetzen von (2.7)

$$\begin{aligned} g_n(x) - A_n x g_{n-1}(x) &\stackrel{(2.7)}{=} g_n(x) - \frac{k_n}{k_{n-1}} \left(k_{n-1} x^n + k_{n-1}^{(1)} x^{n-1} + \dots + k_{n-1}^{(n-1)} x \right) \\ &\stackrel{(2.7)}{=} \left(k_n x^n + k_n^{(1)} x^{n-1} + \dots + k_n^{(n)} x^0 \right) - k_n x^n - \frac{k_n}{k_{n-1}} \left(k_{n-1}^{(1)} x^{n-1} + \dots + k_{n-1}^{(n-1)} x \right) \\ &= \left(k_n^{(1)} - \frac{k_n}{k_{n-1}} k_{n-1}^{(1)} \right) x^{n-1} + \dots + \left(k_n^{(n-1)} - \frac{k_n}{k_{n-1}} k_{n-1}^{(n-1)} \right) x + k_n^{(n)} x^0 \end{aligned}$$

und damit ist $\deg(g_n(x) - A_n x g_{n-1}(x)) \leq n - 1$. Daher lässt sich dieses Polynom mit Koeffizienten $c_0, \dots, c_{n-1} \in \mathbb{R}$ darstellen als

$$g_n(x) - A_n x g_{n-1}(x) = c_{n-1} g_{n-1}(x) + \dots + c_0 g_0(x). \quad (2.9)$$

Es sei $k = 0, \dots, n - 3$.

Dann gilt aufgrund der Orthogonalität der Polynome (2.6) und aufgrund der Eigenschaften des Skalarproduktes:

$$\begin{aligned} c_k \xi_k &\stackrel{(2.6)}{=} c_k \langle g_k, g_k \rangle \\ &\stackrel{(2.5)}{=} \langle c_k g_k, g_k \rangle \\ &\stackrel{(2.6),(2.3)}{=} \langle c_{n-1} g_{n-1} + \dots + c_0 g_0, g_k \rangle \\ &\stackrel{(2.9)}{=} \langle g_n - A_n x g_{n-1}, g_k \rangle \\ &\stackrel{(2.3),(2.5)}{=} \langle g_n, g_k \rangle - A_n \langle x g_{n-1}, g_k \rangle \end{aligned}$$

$$\begin{aligned} &\stackrel{(2.6),(2.5)}{=} -A_n \langle g_{n-1}, xg_k \rangle \\ &\stackrel{(2.6)}{=} 0. \end{aligned}$$

Da ξ_k nach der Definition von null verschieden ist, folgt also $c_k = 0$ für $k = 0, \dots, n-3$ und zusammen mit (2.9) ist

$$g_n(x) - A_n x g_{n-1}(x) = c_{n-1} g_{n-1}(x) + c_{n-2} g_{n-2}(x).$$

Wird $c_{n-1} = B_n, c_{n-2} = -C_n$ gesetzt, so ergibt sich (2.8).

Um die Konstante B_n genauer zu bestimmen, werden die Koeffizienten von x^{n-1} aus der Rekursion (2.8) betrachtet. Mit der Darstellung der Polynome $g_{n-1}(x)$ und $g_n(x)$ aus (2.7) folgt daher

$$\begin{aligned} k_n^{(1)} &= A_n k_{n-1}^{(1)} + B_n k_{n-1} \\ \Leftrightarrow B_n &= \frac{k_n^{(1)} - A_n k_{n-1}^{(1)}}{k_{n-1}} \\ \Leftrightarrow B_n &= A_n \left(\frac{k_n^{(1)}}{k_n} - \frac{k_{n-1}^{(1)}}{k_{n-1}} \right). \end{aligned}$$

Mit Hilfe der Orthogonalitätsbedingung (2.6) ergibt sich weiter

$$\begin{aligned} 0 &\stackrel{(2.6)}{=} \langle g_n, g_{n-2} \rangle \\ &\stackrel{(2.8)}{=} \langle (A_n x + B_n) g_{n-1} - C_n g_{n-2}, g_{n-2} \rangle \\ &\stackrel{(2.3),(2.5)}{=} A_n \langle x g_{n-1}, g_{n-2} \rangle + B_n \langle g_{n-1}, g_{n-2} \rangle - C_n \langle g_{n-2}, g_{n-2} \rangle \\ &\stackrel{(2.6),(2.5)}{=} A_n \langle g_{n-1}, x g_{n-2} \rangle - C_n \langle g_{n-2}, g_{n-2} \rangle \\ &\stackrel{(2.6),(2.5)}{=} A_n \langle g_{n-1}, k_{n-2} x^{n-1} + \dots + k_{n-2}^{(n-2)} x \rangle - C_n \xi_{n-2} \\ &\stackrel{(2.5)}{=} A_n \frac{k_{n-2}}{k_{n-1}} \langle g_{n-1}, g_{n-1} \rangle - C_n \xi_{n-2} \\ &\stackrel{(2.6)}{=} A_n \frac{k_{n-2}}{k_{n-1}} \xi_{n-1} - C_n \xi_{n-2}. \end{aligned}$$

Somit liefert die letzte Gleichheit die Konstante C_n :

$$\begin{aligned} C_n &= A_n \frac{k_{n-2} \xi_{n-1}}{k_{n-1} \xi_{n-2}} \\ &= \frac{A_n}{A_{n-1}} \frac{\xi_{n-1}}{\xi_{n-2}}. \end{aligned}$$

Die Rekursionsformel (2.8) gilt auch für $n=1$, wenn $g_{-1}(x) = 0$ und C_1 beliebig ist. \square

Im Folgenden sollen nun einige spezielle orthogonale Polynome mit ihren Eigenschaften vorgestellt werden.

Dafür werden der folgende Satz, der die Verallgemeinerung der Produktregel beschreibt, und die Definition der Rodrigues-Formel benötigt:

Satz 2.5. (Leibnitz-Formel, [4] Abschnitt 6.1.3.3)

Es seien $u, v: \mathbb{R} \rightarrow \mathbb{R}$ n -mal differenzierbare Funktionen für eine natürliche Zahl $n \in \mathbb{N}$. Dann ist das Produkt $u \cdot v$ n -fach differenzierbar und es gilt:

$$\frac{d^n}{dx^n} [u \cdot v] = \sum_{m=0}^n \binom{n}{m} \frac{d^m}{dx^m} u \cdot \frac{d^{n-m}}{dx^{n-m}} v. \quad (2.10)$$

Beweis.

Der Beweis erfolgt per Induktion.

Der Induktionsanfang (IA) erfolgt für $n=1$. Dabei ergibt sich die in der Analysis 1 bewiesene Produktregel (z.B. [15] Kapitel IV Abschnitt 2.3)

$$\frac{d}{dx} [u \cdot v] = u \frac{d}{dx} v + \frac{d}{dx} [u] v.$$

Als Induktionsvoraussetzung (IV) gelte (2.10) für $n \in \mathbb{N}$.

Für den Induktionsschluss wird $n \rightarrow n + 1$ gesetzt.

Für diesen Beweis werden weitere Eigenschaften des Binomialkoeffizienten, der definiert wird als

$$\binom{n}{k} := \frac{n!}{k!(n-k)!},$$

benötigt:

$$\begin{aligned} \binom{n}{0} &= \binom{n}{n} = 1 & (2.11) \\ \binom{n}{m} + \binom{n}{m-1} &= \frac{n!}{m!(n-m)!} + \frac{n!}{(m-1)!(n-m+1)!} \\ &= \frac{n!(n-m+1) + n!m}{m!(n-m+1)!} \\ &= \frac{n!(n+1)}{m!(n+1-m)!} \\ &= \binom{n+1}{m}. & (2.12) \end{aligned}$$

Damit ergibt sich für die $(n + 1)$ -te Ableitung des Produktes:

$$\begin{aligned}
\frac{d^{n+1}}{dx^{n+1}} [u \cdot v] &= \frac{d}{dx} \left(\frac{d^n}{dx^n} [u \cdot v] \right) \\
&\stackrel{\text{IV}}{=} \frac{d}{dx} \left(\sum_{m=0}^n \binom{n}{m} \frac{d^m}{dx^m} u \cdot \frac{d^{n-m}}{dx^{n-m}} v \right) \\
&\stackrel{\text{IA}}{=} \sum_{m=0}^n \binom{n}{m} \left(\frac{d^m}{dx^m} u \cdot \frac{d^{n-m+1}}{dx^{n-m+1}} v + \frac{d^{m+1}}{dx^{m+1}} u \cdot \frac{d^{n-m}}{dx^{n-m}} v \right) \\
&\stackrel{(2.11)}{=} \binom{n+1}{0} u \frac{d^{n+1}}{dx^{n+1}} v + \binom{n+1}{n+1} \frac{d^{n+1}}{dx^{n+1}} [u] v + \\
&\quad + \sum_{m=1}^n \left(\binom{n}{m} + \binom{n}{m-1} \right) \frac{d^m}{dx^m} u \cdot \frac{d^{n+1-m}}{dx^{n+1-m}} v \\
&\stackrel{(2.12)}{=} \sum_{m=0}^{n+1} \binom{n+1}{m} \frac{d^m}{dx^m} u \cdot \frac{d^{n+1-m}}{dx^{n+1-m}} v
\end{aligned}$$

□

Definition 2.6. (Rodrigues-Formel, [1] Abschnitt 22.1)

Die *Rodrigues-Formel* zur Darstellung von orthogonalen Polynomen ist in allgemeiner Form gegeben durch

$$g_n(x) = \frac{1}{e_n w(x)} \frac{d^n}{dx^n} [w(x) f(x)^n],$$

wobei e_n ein Normierungsfaktor, $f(x)$ ein von $n \in \mathbb{N}_0$ unabhängiges Polynom und $w(x)$ die zu dem jeweiligen orthogonalen Polynom $g_n(x)$ gehörige Gewichtsfunktion ist.

Zunächst werden die Hermite-Polynome vorgestellt und deren Rekursionsformeln hergeleitet. Eine ausführliche Präsentation der darauffolgenden Polynome ist in [1] in Kapitel 22, in [17] in den Kapiteln 4 und 5 und in [6] in Kapitel 2 zu finden, in denen die nachfolgenden unbewiesenen Gleichungen und Rekursionen hergeleitet beziehungsweise bewiesen werden.

Hermite-Polynome ([6] Abschnitt 2.5)

Die Hermite-Polynome $H_n(x)$ sind mit $w(x) = \exp(-x^2)$ gewichtete orthogonale Polynome, die auf dem Intervall $(-\infty, \infty)$ für $n \in \mathbb{N}_0$ definiert sind und die hermitesche Differentialgleichung

$$\frac{d^2}{dx^2} y(x) - 2x \frac{d}{dx} y(x) + 2ny(x) = 0 \quad (2.13)$$

lösen.

Sie lassen sich mit Hilfe der Rodrigues-Formel

$$H_n(x) = (-1)^n \exp(x^2) \frac{d^n}{dx^n} \exp(-x^2) \quad (2.14)$$

darstellen.

Für diese Polynome soll im Folgenden die im Satz 2.4. vorgestellte Drei-Term-Rekursion hergeleitet werden.

Dafür wird $H_{n+1}(x)$ mit Hilfe der Rodrigues-Formel (2.14) unter der Voraussetzung dargestellt, dass $n \geq 1$ und $H_0(x) = 1$ und $H_1(x) = 2x$ ist, was sich durch Einsetzen von $n = 0$ und $n = 1$ in die Rodrigues-Formel (2.14) verifizieren lässt:

$$\begin{aligned} H_{n+1}(x) &\stackrel{(2.14)}{=} (-1)^{n+1} \exp(x^2) \frac{d^{n+1}}{dx^{n+1}} \exp(-x^2) \\ &= (-1)(-1)^n \exp(x^2) \frac{d^n}{dx^n} \left[\frac{d}{dx} \exp(-x^2) \right] \\ &= (-1)(-1)^n \exp(x^2) \frac{d^n}{dx^n} [-2x \exp(-x^2)] \\ &\stackrel{(2.10)}{=} (-1)^n \exp(x^2) \left[2x \frac{d^n}{dx^n} \exp(-x^2) + 2n \frac{d^{n-1}}{dx^{n-1}} \exp(-x^2) \right] \\ &= 2x(-1)^n \exp(x^2) \frac{d^n}{dx^n} \exp(-x^2) - 2n(-1)^{n-1} \exp(x^2) \frac{d^{n-1}}{dx^{n-1}} \exp(-x^2) \\ &\stackrel{(2.14)}{=} 2xH_n(x) - 2nH_{n-1}(x). \end{aligned}$$

Somit erfüllen die Hermite-Polynome für $n \in \mathbb{N}$ die Drei-Term-Rekursion

$$\begin{aligned} H_{n+1}(x) &= 2xH_n(x) - 2nH_{n-1}(x), \\ &\text{mit } H_0(x) = 1, H_1(x) = 2x. \end{aligned} \quad (2.15)$$

Wird die Lösung $H_n(x)$ in die Differentialgleichung (2.13) eingesetzt und diese nach der Variablen x differenziert, so ergibt sich

$$\begin{aligned} &\frac{d^3}{dx^3} H_n(x) - 2x \frac{d^2}{dx^2} H_n(x) - 2 \frac{d}{dx} H_n(x) + 2n \frac{d}{dx} H_n(x) = 0 \\ \Leftrightarrow &\frac{d^3}{dx^3} H_n(x) - 2x \frac{d^2}{dx^2} H_n(x) + 2(n-1) \frac{d}{dx} H_n(x) = 0, \end{aligned}$$

woraus sich schlussfolgern lässt, dass die Differentialgleichung

$$\frac{d^2}{dx^2} y(x) - 2x \frac{d}{dx} y(x) + 2(n-1)y(x) = 0$$

deren Lösung $H_{n-1}(x)$ ist, sich ebenfalls durch $\frac{d}{dx}H_n(x)$ lösen lässt.

Da die Differentialgleichung homogen ist, stimmen also die Lösungen bis auf einen Faktor c überein

$$H_{n-1}(x) = c \frac{d}{dx}H_n(x),$$

der sich mithilfe eines Koeffizientenvergleiches in der höchsten Potenz der Differentialgleichung, also x^{n-1} , bestimmen lässt.

Der Koeffizient der höchsten Potenz x^n von $H_n(x)$ lässt sich leicht mit Hilfe der Drei-Term-Rekursion

$$H_n(x) = 2xH_{n-1}(x) - 2(n-1)H_{n-2}(x)$$

bestimmen, da die höchste Potenz durch das Produkt $2xH_{n-1}(x)$ gebildet wird. Wird dieses Produkt mithilfe von $(n-2)$ -maliger Anwendung von (2.15) auf $H_1(x)$ zurückgeführt, so ergibt sich als Leitterm von $H_n(x)$

$$\begin{aligned} a_n x^n &\stackrel{(2.15)}{=} \underbrace{2x \cdot 2x \cdot \dots \cdot 2x}_{(n-1)\text{-mal}} H_1(x) \\ &= 2^{n-1} x^{n-1} H_1(x) \\ &\stackrel{(2.15)}{=} 2^n x^n. \end{aligned}$$

Somit ist der Koeffizient der höchsten Potenz von $H_n(x)$ $a_n = 2^n$ und ein Koeffizientenvergleich liefert:

$$\begin{aligned} 2^{n-1} x^{n-1} &= c \frac{d}{dx} (2^n x^n) \\ 2^{n-1} x^{n-1} &= cn 2^n x^{n-1} \\ \Leftrightarrow \quad c &= \frac{1}{2n}. \end{aligned}$$

Daraus resultiert eine weitere Rekursionsformel für $n \in \mathbb{N}_0$:

$$H_n(x) = \frac{1}{2(n+1)} \frac{d}{dx} H_{n+1}(x). \quad (2.16)$$

Laguerre-Polynome ([17] Abschnitt 5.1)

Die Laguerre-Polynome $L_n(x)$ sind auf dem Intervall $[0, \infty)$ mit der Gewichtsfunktion $w(x) = \exp(-x)$ definierte orthogonale Polynome, die die laguerreschen Differentialgleichung

$$x \frac{d^2}{dx^2} y(x) + (1-x) \frac{d}{dx} y(x) + ny(x) = 0$$

für $n \in \mathbb{N}_0$ lösen.

Sie sind ebenfalls mit Hilfe der Rodrigues-Formel

$$L_n(x) = \frac{\exp(x)}{n!} \frac{d^n}{dx^n} [x^n \exp(-x)]$$

darstellbar.

Die Rekursionsformeln

$$\begin{aligned} (n+1)L_{n+1}(x) &= (2n+1-x)L_n(x) - nL_{n-1}(x), \\ &\text{mit } L_0(x) = 1, L_1(x) = -x + 1, \\ x \frac{d}{dx} L_n(x) &= nL_n(x) - nL_{n-1}(x) \end{aligned} \quad (2.17)$$

sind für $n \in \mathbb{N}$ gültig und werden unter anderem in [17] im Abschnitt 5.1 ausführlich hergeleitet. Diese helfen bei der Konstruktion einer weiteren Rekursionsvorschrift.

Dafür wird (2.17) nach x abgeleitet

$$\frac{d}{dx} L_n(x) + x \frac{d^2}{dx^2} L_n(x) = n \frac{d}{dx} L_n(x) - n \frac{d}{dx} L_{n-1}(x). \quad (2.18)$$

Es soll nun die Differenz zwischen der laguerreschen Differentialgleichung, in die das Laguerre-Polynom n -ten Grades eingesetzt wurde

$$x \frac{d^2}{dx^2} L_n(x) + (1-x) \frac{d}{dx} L_n(x) + nL_n(x) = 0$$

und (2.18) gebildet werden:

$$\begin{aligned} -x \frac{d}{dx} L_n(x) + nL_n(x) &= -n \frac{d}{dx} L_n(x) + n \frac{d}{dx} L_{n-1}(x) \\ x \frac{d}{dx} L_n(x) &= nL_n(x) + n \frac{d}{dx} L_n(x) - n \frac{d}{dx} L_{n-1}(x) \\ \stackrel{(2.17)}{\Leftrightarrow} nL_n(x) - nL_{n-1}(x) &= nL_n(x) + n \frac{d}{dx} L_n(x) - n \frac{d}{dx} L_{n-1}(x) \\ \Leftrightarrow nL_{n-1}(x) &= -n \frac{d}{dx} L_n(x) + n \frac{d}{dx} L_{n-1}(x) \\ \Leftrightarrow L_{n-1}(x) &= -\frac{d}{dx} L_n(x) + \frac{d}{dx} L_{n-1}(x). \end{aligned} \quad (2.19)$$

Diese Rekursion gilt für Werte $n = 1, 2, \dots$

Legendre-Polynome ([6] Abschnitt 2.3, Abschnitt 2.13)

Die Legendre-Polynome $P_n(x)$ sind für $n \in \mathbb{N}_0$ spezielle orthogonale Polynome, die auf

dem Intervall $[-1, 1]$ mit der Gewichtsfunktion $w(x) = 1$ definiert und Lösungen der Legendreschen Differentialgleichung

$$(1 - x^2) \frac{d^2}{dx^2} y(x) - 2x \frac{d}{dx} y(x) + n(n+1)y(x) = 0$$

sind.

Sie lassen sich durch die Rodrigues-Formel

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} [(x^2 - 1)^n]$$

darstellen.

Weitere rekursive Darstellungsmöglichkeiten der Legendre-Polynome sind für $n \in \mathbb{N}$

$$P_{n+1}(x) = \frac{2n+1}{n+1} x P_n(x) - \frac{n}{n+1} P_{n-1}(x),$$

mit $P_0(x) = 1, P_1(x) = x,$

$$(x^2 - 1) \frac{d}{dx} P_n(x) = n x P_n(x) - n P_{n-1}(x),$$

deren Herleitung in [6] in den Abschnitten 2.3 beziehungsweise 2.13 nachgelesen werden kann.

Eine weitere Rekursionsvorschrift kann für $n = 1, 2, \dots$ ähnlich wie im Falle der Laguerre-Polynome bestimmt werden:

$$P_n(x) = \frac{1}{2n+1} \frac{d}{dx} P_{n+1}(x) - \frac{1}{2n+1} \frac{d}{dx} P_{n-1}(x). \quad (2.20)$$

Die ausführliche Herleitung kann dafür im Anhang A.1 dieser Arbeit nachgelesen werden.

Jacobi-Polynome ([17] Kapitel 4, [1])

Die Jacobi-Polynome $P_n^{(\alpha, \beta)}(x)$ sind Lösungen der Differentialgleichung

$$(1 - x^2) \frac{d^2}{dx^2} y(x) + (\beta - \alpha - (\alpha + \beta + 2)x) \frac{d}{dx} y(x) + n(n + \alpha + \beta + 1)y(x) = 0$$

und auf dem Intervall $[-1, 1]$ mit der Gewichtsfunktion $w(x) = (1-x)^\alpha (1+x)^\beta$ definiert, wobei $n \in \mathbb{N}_0$ und $\alpha, \beta > -1$.

Sie lassen sich wie folgt mit Hilfe der Rodrigues-Formel für Werte $\alpha, \beta > -1$

$$P_n^{(\alpha, \beta)}(x) = \frac{(-1)^n}{2^n n!} (1-x)^{-\alpha} (1+x)^{-\beta} \frac{d^n}{dx^n} [(1-x)^{\alpha+n} (1+x)^{\beta+n}]$$

darstellen.

Sie genügen desweiteren für $n \in \mathbb{N}$ folgenden Rekursionsbeziehungen

$$\begin{aligned} & 2(n+1)(n+\alpha+\beta+1)(2n+\alpha+\beta)P_{n+1}^{(\alpha,\beta)}(x) = \\ & = (2n+\alpha+\beta+1)(\alpha^2-\beta^2+(2n+\alpha+\beta)(2n+\alpha+\beta+2)x)P_n^{(\alpha,\beta)}(x) - \\ & - 2(n+\alpha)(n+\beta)(2n+\alpha+\beta+2)P_{n-1}^{(\alpha,\beta)}(x) \end{aligned}$$

$$\text{mit } P_0^{(\alpha,\beta)}(x) = 1, P_1^{(\alpha,\beta)}(x) = \frac{1}{2}(\alpha - \beta + (\alpha + \beta + 2)x)$$

für $\alpha, \beta > -1$,

$$\begin{aligned} (2n+\alpha+\beta)(1-x^2)\frac{d}{dx}P_n^{(\alpha,\beta)}(x) & = n(\alpha - \beta - (2n+\alpha+\beta)x)P_n^{(\alpha,\beta)}(x) + \\ & + 2(n+\alpha)(n+\beta)P_{n-1}^{(\alpha,\beta)}(x), \end{aligned}$$

die in [17] in Kapitel 4 aufgezeigt wurden.

Die ausführliche Herleitung einer weiteren Rekursionsformel ist im Anhang A.2 enthalten, deren Ergebnis für $n \in \mathbb{N}$ und $\alpha, \beta > -1$ wie folgt lautet:

$$\begin{aligned} P_n^{(\alpha,\beta)}(x) & = \frac{2(n+\alpha+\beta+1)}{(2n+\alpha+\beta+1)(2n+\alpha+\beta+2)}\frac{d}{dx}P_{n+1}^{(\alpha,\beta)}(x) + \\ & + \frac{2(\alpha-\beta)}{(2n+\alpha+\beta)(2n+\alpha+\beta+2)}\frac{d}{dx}P_n^{(\alpha,\beta)}(x) - \\ & - \frac{2(n+\alpha)(n+\beta)}{(n+\alpha+\beta)(2n+\alpha+\beta)(2n+\alpha+\beta+1)}\frac{d}{dx}P_{n-1}^{(\alpha,\beta)}(x). \end{aligned} \quad (2.21)$$

Chebyshev-Polynome erster Art ([6] Abschnitte 2.6 und 2.13)

Die Chebyshev-Polynome erster Art $T_n(x)$ sind für $n \in \mathbb{N}_0$ auf dem Intervall $[-1, 1]$ mit der Funktion $w(x) = (1-x^2)^{-\frac{1}{2}}$ gewichtete spezielle orthogonale Polynome, die die Chebyshev-Differentialgleichung

$$(1-x^2)\frac{d^2}{dx^2}y(x) - x\frac{d}{dx}y(x) + n^2y(x) = 0$$

lösen.

Die zugehörige Rodrigues-Formel ist in [6] Abschnitt 2.10 beschrieben.

Die rekursiven Zusammenhänge

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x),$$

$$\text{mit } T_0(x) = 1, T_1(x) = x,$$

$$(1-x^2)\frac{d}{dx}T_n(x) = -nxT_n(x) + nT_{n-1}(x)$$

gelten für $n \in \mathbb{N}$ und werden ausführlich in [6] in den Abschnitten 2.6 und 2.13 hergeleitet und helfen, um eine weitere Rekursionsvorschrift

$$T_n(x) = \frac{1}{2(n+1)} \frac{d}{dx} T_{n+1} - \frac{1}{2(n-1)} \frac{d}{dx} T_{n-1}(x) \quad (2.22)$$

zu bestimmen, deren ausführliche Herleitung im Anhang A.3 zu finden ist.

Chebyshev-Polynome zweiter Art ([6] Abschnitte 2.7 und 2.13)

Die Chebyshev-Polynome zweiter Art $U_n(x)$ sind ähnlich wie die Chebyshev-Polynome erster Art für $n \in \mathbb{N}_0$ auf dem Intervall $[-1, 1]$ definiert, allerdings mit der Gewichtsfunktion $w(x) = (1 - x^2)^{\frac{1}{2}}$ und lösen die Differentialgleichung

$$(1 - x^2) \frac{d^2}{dx^2} y(x) - 3x \frac{d}{dx} y(x) + n(n+2)y(x) = 0.$$

Die zugehörige Rodrigues-Formel ist in [6] Abschnitt 2.10 dargestellt. Es existieren ebenfalls Rekursionsvorschriften

$$\begin{aligned} U_{n+1}(x) &= 2xU_n(x) - U_{n-1}(x), \\ &\text{mit } U_0(x) = 1, U_1(x) = 2x, \\ (1 - x^2) \frac{d}{dx} U_n(x) &= -nxU_n(x) + (n+1)U_{n-1}(x), \end{aligned}$$

die für $n \in \mathbb{N}$ gelten und in [6] in den Abschnitten 2.7 beziehungsweise 2.13 aufgezeigt werden.

Die Herleitung einer weiteren Rekursionsvorschrift, die den Zusammenhang zwischen einem Chebyshev-Polynom zweiter Art und der Ableitung seines Vorgängers und Nachfolgers beschreibt, befindet sich im Anhang A.4.

Das Resultat lautet wie folgt:

$$U_n(x) = \frac{1}{2(n+1)} \frac{d}{dx} U_{n+1}(x) - \frac{1}{2(n-1)} \frac{d}{dx} U_{n-1}(x). \quad (2.23)$$

Mithilfe der Resultate der vorgestellten orthogonalen Polynome lässt sich ein weiterer Satz formulieren:

Satz 2.7. (Differentialrekursionsformel)

Es sei eine Menge der bereits vorgestellten orthogonalen Polynome $g_i(x)$ für $i \in \mathbb{N}_0$ gegeben.

Für drei aufeinanderfolgende orthogonale Polynome gilt die Beziehung:

$$g_n(x) = \alpha_n \frac{d}{dx} g_{n+1}(x) + \beta_n \frac{d}{dx} g_n(x) + \gamma_n \frac{d}{dx} g_{n-1}(x), \quad n = 1, 2, \dots, \quad (2.24)$$

wobei $\alpha_n, \beta_n, \gamma_n$ die Differentialrekursionskoeffizienten sind.

Diese sind in der Tabelle 2.1 für die vorgestellten orthogonalen Polynome aufgelistet.

Beweis.

Für die in diesem Abschnitt eingeführten orthogonalen Polynome wurden die einzelnen Rekursionen (2.16), (2.19), (2.20), (2.21), (2.22) und (2.23) in der Vorstellung der einzelnen Polynome beziehungsweise im Anhang hergeleitet. \square

Polynom	α_n	β_n	γ_n
$H_n(x)$	$\frac{1}{2(n+1)}$	0	0
$L_n(x)$	-1	1	0
$P_n(x)$	$\frac{1}{2n+1}$	0	$-\frac{1}{2n+1}$
$P_n^{(\alpha, \beta)}(x)$ ($\alpha, \beta > -1$)	$\frac{2(n+\alpha+\beta+1)}{(2n+\alpha+\beta+1)(2n+\alpha+\beta+2)}$	$\frac{2(\alpha-\beta)}{(2n+\alpha+\beta)(2n+\alpha+\beta+2)}$	$\frac{-2(n+\alpha)(n+\beta)}{(n+\alpha+\beta)(2n+\alpha+\beta)(2n+\alpha+\beta+1)}$
$T_n(x)$	$\frac{1}{2(n+1)}$	0	$-\frac{1}{2(n-1)}$
$U_n(x)$	$\frac{1}{2(n+1)}$	0	$-\frac{1}{2(n+1)}$

Tabelle 2.1: Differentialrekursionskoeffizienten

Die einzelnen Differentialrekursionskoeffizienten wurden in dieser Arbeit hergeleitet, da sie in [13] und [21] fehlerhaft waren.

2.2 Matrixgleichungen

In diesem Abschnitt soll eine Äquivalenzrelation zwischen einem linearen Gleichungssystem

$$Ax = b,$$

wobei $A \in \mathbb{R}^{n \times n}$ und $b \in \mathbb{R}^n$ gegeben und $x \in \mathbb{R}^n$ gesucht ist, und einer allgemeineren Gleichung

$$F(X) = C, \tag{2.25}$$

hergestellt werden, wobei $F: \mathbb{R}^{q \times p} \rightarrow \mathbb{R}^{p \times s}$ eine lineare Abbildung, $C \in \mathbb{R}^{p \times s}$ gegeben und $X \in \mathbb{R}^{q \times r}$ gesucht ist.

Dafür werden die Definitionen des Kronecker-Produktes und der Vektorisierung einer Matrix, sowie weitere Eigenschaften benötigt.

Definition 2.8. (Kronecker-Produkt, [7] Abschnitt 6.3.3)

Es seien $A \in \mathbb{R}^{p \times q}$ und $B \in \mathbb{R}^{r \times s}$.

Dann ist das *Kronecker-Produkt* definiert durch

$$A \otimes B := \begin{bmatrix} a_{1,1}B & \cdots & a_{1,q}B \\ \vdots & \ddots & \vdots \\ a_{p,1}B & \cdots & a_{p,q}B \end{bmatrix} \in \mathbb{R}^{pr \times qs}. \quad (2.26)$$

Satz 2.9. (Eigenschaften des Kronecker-Produktes, [7] Abschnitt 6.3.3)

Es seien $A, C \in \mathbb{R}^{n \times n}$ und $B, D \in \mathbb{R}^{m \times m}$.

Dann gilt:

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD). \quad (2.27)$$

Sind A und B invertierbar, so gilt desweiteren:

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}. \quad (2.28)$$

Beweis.

Für den ersten Teil des Beweises wird die Definition des Kronecker-Produktes genutzt, um die Äquivalenz beider Seiten zu zeigen:

$$\begin{aligned} (A \otimes B) \cdot (C \otimes D) &= \begin{bmatrix} a_{1,1}B & \cdots & a_{1,n}B \\ \vdots & \ddots & \vdots \\ a_{n,1}B & \cdots & a_{n,n}B \end{bmatrix} \cdot \begin{bmatrix} c_{1,1}D & \cdots & c_{1,n}D \\ \vdots & \ddots & \vdots \\ c_{n,1}D & \cdots & c_{n,n}D \end{bmatrix} \\ &= \begin{bmatrix} \sum_{i=1}^n a_{1,i}Bc_{i,1}D & \cdots & \sum_{i=1}^n a_{1,i}Bc_{i,n}D \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^n a_{n,i}Bc_{i,1}D & \cdots & \sum_{i=1}^n a_{n,i}Bc_{i,n}D \end{bmatrix} \\ &= \begin{bmatrix} \sum_{i=1}^n a_{1,i}c_{i,1} & \cdots & \sum_{i=1}^n a_{1,i}c_{i,n} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^n a_{n,i}c_{i,1} & \cdots & \sum_{i=1}^n a_{n,i}c_{i,n} \end{bmatrix} \otimes (BD) \\ &= (AC) \otimes (BD). \end{aligned}$$

Für den zweiten Teil des Beweises wird $C = A^{-1}$ und $D = B^{-1}$ unter der Voraussetzung, dass A und B invertierbar sind, gesetzt und (2.27) verwendet. Dann ist:

$$\begin{aligned} (A \otimes B) \cdot (A^{-1} \otimes B^{-1}) &\stackrel{(2.27)}{=} (AA^{-1}) \otimes (BB^{-1}) \\ &= I_n \otimes I_m \\ &= I_{nm}. \end{aligned}$$

Somit ist $(A^{-1} \otimes B^{-1})$ die Rechtsinverse von $(A \otimes B)$. Da der Beweis der Linksinversen äquivalent dazu erfolgt, wird dieser ausgelassen. Folglich ist die Inverse von $(A \otimes B)$

$$(A^{-1} \otimes B^{-1}) = (A \otimes B)^{-1}.$$

□

Definition 2.10. (Vektorisierung einer Matrix, [5] Abschnitt 2.7)

Es sei $X = \begin{bmatrix} x_1 & \dots & x_q \end{bmatrix} \in \mathbb{R}^{p \times q}$ mit den Spalten $x_1, \dots, x_q \in \mathbb{R}^p$. Dann ist die *Vektorisierung der Matrix X* definiert durch

$$\text{vec}(X) := \begin{bmatrix} x_1 \\ \vdots \\ x_q \end{bmatrix} \in \mathbb{R}^{pq}.$$

Satz 2.11. (Äquivalenzrelation, [7] Abschnitt 6.3.3)

Es seien $A \in \mathbb{R}^{p \times q}$, $B \in \mathbb{R}^{r \times s}$, $C \in \mathbb{R}^{p \times s}$ und $F(X) = AXB$ in (2.25).

Dann gilt die Äquivalenzrelation

$$AXB = C \quad \Leftrightarrow \quad (B^T \otimes A) \text{vec}(X) = \text{vec}(C). \quad (2.29)$$

Beweis.

Für den Beweis werden die Matrizen $B = \begin{bmatrix} b_1 & \dots & b_s \end{bmatrix}$, $C = \begin{bmatrix} c_1 & \dots & c_s \end{bmatrix}$, sowie $X = \begin{bmatrix} x_1 & \dots & x_r \end{bmatrix}$ der Gleichung $AXB = C$ nacheinander spaltenweise betrachtet:

$$\begin{aligned} &AXB = C \\ \Leftrightarrow &AX \begin{bmatrix} b_1 & \dots & b_s \end{bmatrix} = \begin{bmatrix} c_1 & \dots & c_s \end{bmatrix} \\ \Leftrightarrow &\begin{bmatrix} AXb_1 \\ \vdots \\ AXb_s \end{bmatrix} = \begin{bmatrix} c_1 \\ \vdots \\ c_s \end{bmatrix} \\ \Leftrightarrow &\begin{bmatrix} A \begin{bmatrix} x_1 & \dots & x_r \end{bmatrix} b_1 \\ \vdots \\ A \begin{bmatrix} x_1 & \dots & x_r \end{bmatrix} b_s \end{bmatrix} = \text{vec}(C) \end{aligned}$$

$$\begin{aligned}
&\Leftrightarrow \begin{bmatrix} A(b_{1,1}x_1 + \dots + b_{r,1}x_r) \\ \vdots \\ A(b_{1,s}x_1 + \dots + b_{r,s}x_r) \end{bmatrix} = \text{vec}(C) \\
&\Leftrightarrow \begin{bmatrix} b_{1,1}A & \cdots & b_{r,1}A \\ \vdots & \ddots & \vdots \\ b_{1,s}A & \cdots & b_{r,s}A \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_r \end{bmatrix} = \text{vec}(C) \\
&\Leftrightarrow (B^T \otimes A) \text{vec}(X) = \text{vec}(C).
\end{aligned}$$

□

Im Verlauf dieser Arbeit wird eine spezielle Gleichung von Bedeutung sein, die verallgemeinerte Sylvester-Gleichung:

Definition 2.12. (Verallgemeinerte Sylvester-Gleichung, z.B. [3])

Es seien $A, E \in \mathbb{R}^{n \times n}$, $F, H \in \mathbb{R}^{m \times m}$ und $R, X \in \mathbb{R}^{n \times m}$.

Eine Gleichung der Form

$$AXF + EXH = R \quad (2.30)$$

heißt *verallgemeinerte Sylvester-Gleichung*.

Im Folgenden wird die verallgemeinerte Sylvester-Gleichung der Einfachheit halber Sylvester-Gleichung genannt.

Mithilfe der hergeleiteten Äquivalenzbeziehung (2.29) kann die Sylvester-Gleichung (2.30) zu einem linearen Gleichungssystem umformuliert werden

$$\begin{aligned}
&AXF + EXH = R \\
&\stackrel{(2.29)}{\Leftrightarrow} \underbrace{(F^T \otimes A + H^T \otimes E)}_{\mathcal{A}} \underbrace{\text{vec}(X)}_x = \underbrace{\text{vec}(R)}_b.
\end{aligned} \quad (2.31)$$

Somit wird ersichtlich, dass die Sylvester-Gleichung (2.30) genau dann eine eindeutige Lösung besitzt, wenn das äquivalente lineare Gleichungssystem $\mathcal{A}x = b$ eindeutig lösbar ist, also wenn $\mathcal{A} = (F^T \otimes A + H^T \otimes E)$ invertierbar ist.

Die eindeutige Lösbarkeit des linearen Gleichungssystems $\mathcal{A}x = b$ lässt sich auch mit Hilfe von Eigenwerten charakterisieren. Dafür werden diese im Folgenden definiert und eine weitere wichtige Eigenschaft formuliert:

Definition 2.13. (Eigenwerte, Eigenvektoren, Spektrum, [10] Abschnitt 7.1.1)

Die *Eigenwerte* einer Matrix $A \in \mathbb{C}^{n \times n}$ sind die n Nullstellen des charakteristischen Polynoms

$$p_A(\lambda) := \det(A - \lambda I_n).$$

Die Menge der Eigenwerte von A wird *Spektrum* von A genannt und definiert als

$$\Lambda(A) := \{\lambda \in \mathbb{C} \mid \det(A - \lambda I) = 0\}.$$

Ist $\lambda \in \Lambda(A)$ ein Eigenwert von A , dann wird der von Null verschiedene Vektor $x \in \mathbb{C}^n$ (rechter) *Eigenvektor* genannt, wenn er

$$Ax = \lambda x$$

erfüllt beziehungsweise linker Eigenvektor im Falle von $x^H A = \lambda x^H$.

Definition 2.14. (Verallgemeinerte Eigenwerte [7] Abschnitt 4.5)

Es seien $A, E \in \mathbb{C}^{n \times n}$.

Ist $\det(A - \lambda E)$ nicht identisch null, so wird das Matrixbüschel $A - \lambda E$ für $\lambda \in \mathbb{C}$ *regulär* genannt. In diesem Fall heißt $p(\lambda) = \det(A - \lambda E)$ *charakteristisches Polynom*.

Die *verallgemeinerten Eigenwerte* sind für ein reguläres Matrixbüschel $A - \lambda E$ definiert durch

- $\det(A - \lambda E) = 0$ für $\lambda \in \mathbb{C}$,
- ∞ mit Vielfachheit $n - \deg p(\lambda)$, falls $\deg(p(\lambda)) < n$.

Das *Spektrum* des Matrixbüschels $A - \lambda E$ wird mit $\Lambda(A, E)$ bezeichnet.

Ist $\lambda \in \Lambda(A, E)$ ein endlicher Eigenwert des Matrixbüschels $A - \lambda E$ und ist

$$Ax = \lambda Ex, \quad x \in \mathbb{C}^n \setminus \{0\},$$

so wird x *Eigenvektor* zu $A - \lambda E$ genannt.

Ist andererseits $\lambda = \infty$ und $Ex = 0$, so ist x ebenfalls ein Eigenvektor.

Satz 2.15.

Es sei $A \in \mathbb{R}^{n \times n}$.

Dann gilt :

$$A \text{ invertierbar} \Leftrightarrow 0 \notin \Lambda(A). \quad (2.32)$$

Beweis.

Es sei $A \in \mathbb{R}^{n \times n}$.

Für den Beweis werden folgende nützliche Aussagen verwendet (siehe [10], Abschnitte 2.1.4 und 7.1.1):

$$\text{Eine Matrix } A \text{ ist invertierbar} \Leftrightarrow \det(A) \neq 0. \quad (2.33)$$

$$\text{Es gilt: } \det(A) = \prod_{i=1}^n \lambda_i \quad \text{mit } \lambda_i \in \Lambda(A) \text{ für } i = 1, \dots, n. \quad (2.34)$$

Damit ist:

$$\begin{aligned}
 A \text{ invertierbar} & \stackrel{(2.33)}{\Leftrightarrow} \det(A) \neq 0 \\
 & \stackrel{(2.34)}{\Leftrightarrow} \prod_{i=1}^n \lambda_i \neq 0 \quad \text{mit } \lambda_i \in \Lambda(A) \text{ f\"ur } i = 1, \dots, n \\
 & \Leftrightarrow \lambda_i \neq 0 \text{ f\"ur ein } i \in \{1, \dots, n\}.
 \end{aligned}$$

□

Aufgrund des vorangegangenen Satzes lässt sich die eindeutige Lösbarkeit der Sylvester-Gleichung (2.30) also mit Hilfe der Eigenwerte der Matrix \mathcal{A} , die durch eine Summe von Kronecker-Produkten dargestellt ist, voraussagen. Die Eigenwerte eines Kronecker-Produktes lassen sich auf folgende Weise bestimmen:

Satz 2.16. (Eigenwerte eines Kronecker-Produktes)

Es sei $A \in \mathbb{R}^{n \times n}$ eine Matrix mit Eigenwerten $\lambda_i \in \mathbb{C}$ für $i = 1, \dots, n$ und $B \in \mathbb{R}^{m \times m}$ eine Matrix mit Eigenwerten $\mu_j \in \mathbb{C}$ für $j = 1, \dots, m$.

Dann sind die Eigenwerte von $A \otimes B$ gegeben durch

$$\Lambda(A \otimes B) = \{\lambda_i \mu_j \mid i = 1, \dots, n, j = 1, \dots, m\}. \quad (2.35)$$

Beweis.

Es seien $x_i \in \mathbb{R}^n$ die zu $\lambda_i \in \mathbb{C}$ gehörigen Eigenvektoren für $i = 1, \dots, n$ und $y_j \in \mathbb{R}^m$ die zu $\mu_j \in \mathbb{C}$ gehörigen Eigenvektoren für $j = 1, \dots, m$.

Dann ist für $i = 1, \dots, n$ und $j = 1, \dots, m$:

$$\begin{aligned}
 (A \otimes B)(x_i \otimes y_j) & \stackrel{(2.27)}{=} Ax_i \otimes By_j \\
 & \stackrel{\substack{Ax_i = \lambda_i x_i \\ By_j = \mu_j y_j}}{=} \lambda_i x_i \otimes \mu_j y_j \\
 & \stackrel{(2.27)}{=} \lambda_i \mu_j (x_i \otimes y_j).
 \end{aligned}$$

Damit sind die Eigenwerte der Matrix $A \otimes B$ die nm Produkte $\lambda_i \mu_j$ für $i = 1, \dots, n$ und $j = 1, \dots, m$. □

Wird das vorangegangene Resultat auf die Matrix \mathcal{A} aus (2.31) übertragen, so ergibt sich folgender Satz:

Satz 2.17. (Eindeutige Lösbarkeit einer verallgemeinerten Sylvester-Gleichung)

Es seien $A, E \in \mathbb{R}^{n \times n}$, $F, H \in \mathbb{R}^{m \times m}$ und $R \in \mathbb{R}^{n \times m}$.

Die Sylvester-Gleichung

$$AXF + EXH = R$$

ist genau dann eindeutig lösbar, wenn alle Eigenwerte der Matrix

$$F^T \otimes A + H^T \otimes E$$

von Null verschieden sind, das heißt es gilt

$$\Lambda(F^T \otimes A) \cap \Lambda(-H^T \otimes E) = \emptyset, \quad (2.36)$$

vorausgesetzt, dass $F^T \otimes A$ und $H^T \otimes E$ gemeinsame Eigenvektoren haben.

Beweis.

Es seien $A, E \in \mathbb{R}^{n \times n}$, $F, H \in \mathbb{R}^{m \times m}$ und $R \in \mathbb{R}^{n \times m}$.

Die Sylvester-Gleichung

$$AXF + EXH = R$$

ist nach (2.31) genau dann eindeutig lösbar, wenn das lineare Gleichungssystem

$$\mathcal{A} \operatorname{vec}(X) = \operatorname{vec}(R)$$

mit $\mathcal{A} := F^T \otimes A + H^T \otimes E$ eindeutig lösbar ist.

Dafür wird die Invertierbarkeit der Matrix \mathcal{A} benötigt, die nach (2.32) äquivalent zu einem von Null verschiedenen Spektrum von \mathcal{A} ist.

Aufgrund der Eigenschaft (siehe [10], Abschnitt 2.1.4)

$$\det(B) = \det(B^T) \quad (2.37)$$

gilt für Eigenwerte $\lambda \in \Lambda(B^T)$ einer Matrix laut Definition

$$\begin{aligned} \lambda \in \Lambda(B^T) &\Leftrightarrow 0 = \det(B^T - \lambda I_n) \\ &\Leftrightarrow 0 = \det((B - \lambda I_n)^T) \\ &\stackrel{(2.37)}{\Leftrightarrow} 0 = \det(B - \lambda I_n) \\ &\Leftrightarrow \lambda \in \Lambda(B). \end{aligned} \quad (2.38)$$

Es seien x_i ein Eigenvektor zu den Eigenwerten $f_i, h_i \in \mathbb{C}$ der Matrizen F und H für $i = 1, \dots, m$ beziehungsweise y_j ein Eigenvektor zu den Eigenwerten $a_j, e_j \in \mathbb{C}$ der Matrizen A und E für $j = 1, \dots, n$. Dann lassen sich die Eigenwerte von \mathcal{A} wie folgt bestimmen:

$$\begin{aligned} (F^T \otimes A + H^T \otimes E)(x_i \otimes y_j) &= (F^T \otimes A)(x_i \otimes y_j) + (H^T \otimes E)(x_i \otimes y_j) \\ &\stackrel{(2.38)}{=} f_i a_j (x_i \otimes y_j) + h_i e_j (x_i \otimes y_j) \\ &\stackrel{(2.35)}{=} (f_i a_j + h_i e_j)(x_i \otimes y_j). \end{aligned}$$

Zusammen mit (2.32) folgt, dass

$$\Lambda(F^T \otimes A) \cap \Lambda(-H^T \otimes E) = \emptyset.$$

□

2.3 Modellreduktion durch Momentenabgleich

Das Modellreduktionsverfahren des Momentenabgleichs soll in diesem Abschnitt nur kurz umrissen werden, um einen Einblick in diese Methode zu erhalten. Eine ausführliche Darstellung ist in [20] in Kapitel 1 enthalten, an die diese Ausarbeitungen angelehnt sind.

Die Methode des Momentenabgleichs ist ein projektionsbasiertes Modellreduktionsverfahren, bei dem zunächst ein Unterraum gesucht wird, der den Raum, auf dem sich der Zustandsvektor $x(t)$ befindet, approximiert. Auf diesem gefundenen Unterraum wird eine orthogonale Basis $V \in \mathbb{R}^{n \times m}$ konstruiert, sodass $x(t) \approx Vx_r(t)$. Durch eine weitere Bedingung an den Fehler $Ae(t) - E\dot{e}(t)$, wobei $e(t) = x(t) - Vx_r(t)$ den Approximationsfehler der Zustandsvektoren beschreibt, kann eine zweite orthogonale Matrix $W \in \mathbb{R}^{n \times m}$ bestimmt werden, sodass ein reduziertes System durch Multiplikation mit W^T wie folgt entsteht:

$$\underbrace{W^T E V}_{E_r} \dot{x}_r(t) = \underbrace{W^T A V}_{A_r} x_r(t) + \underbrace{W^T B}_{B_r} u(t)$$

$$y_r(t) = \underbrace{C V}_{C_r} x_r(t).$$

Die exakte Herleitung der Projektionsmatrizen ist in [20] in Abschnitt 1.3 zu finden. Ein Unterraum, der beim Momentenabgleich eine wichtige Rolle spielt, ist der Krylov-Unterraum.

Definition 2.18. (Krylov-Unterraum, [20] Abschnitt 1.6)

Für eine gegebene Matrix $A \in \mathbb{R}^{n \times n}$ und einen Vektor $b \in \mathbb{R}^n$ wird ein *Krylov-Unterraum* der Ordnung r definiert als

$$\mathcal{K}_r(A, b) = \text{span}\{b, Ab, \dots, A^{r-1}b\}.$$

Beim Momentenabgleich wird die Übertragungsfunktion $G(s)$ durch eine Potenzreihe in einem Punkt $s_0 \in \mathbb{C} \cup \{\infty\}$ entwickelt. Dafür wird die Neumann-Reihe ([19], Kapitel II)

$$(I_n - T)^{-1} = \sum_{k=0}^{\infty} T^k \tag{2.39}$$

für eine Einheitsmatrix $I_n \in \mathbb{R}^{n \times n}$ und eine Matrix $T \in \mathbb{R}^{n \times n}$ verwendet.

Es sei $s_0 \in \mathbb{C}$ ein Entwicklungspunkt und $(s_0 E - A)$ invertierbar. Dann kann die Über-

tragungsfunktion $G(s)$ entwickelt werden durch

$$\begin{aligned}
 G(s) &\stackrel{(1.3)}{=} C(sE - A)^{-1}B \\
 &= C(s_0E - A - (s_0 - s)E)^{-1}B \\
 &= C\left(I - (s_0 - s)(s_0E - A)^{-1}E\right)^{-1}(s_0E - A)^{-1}B \\
 &\stackrel{(2.39)}{=} C \sum_{k=0}^{\infty} \left((s_0 - s)(s_0E - A)^{-1}E\right)^k (s_0E - A)^{-1}B \\
 &= \sum_{k=0}^{\infty} \underbrace{C \left(- (s_0E - A)^{-1}E\right)^k (s_0E - A)^{-1}B}_{=: M_k^{s_0}} (s - s_0)^k \\
 &= \sum_{k=0}^{\infty} M_k^{s_0} (s - s_0)^k,
 \end{aligned}$$

wobei $M_k^{s_0}$ die *Momente* der Übertragungsfunktion um s_0 genannt werden. Für einen Entwicklungspunkt im Nullpunkt vereinfachen sich die Momente zu

$$M_k^0 = -C(A^{-1}E)^k A^{-1}B$$

beziehungsweise falls $s_0 \rightarrow \infty$, so wird die Übertragungsfunktion durch

$$G(s) = \sum_{k=0}^{\infty} M_k^{\infty} (s)^{-(k+1)}$$

entwickelt mit den sogenannten *Markov-Parametern* von $G(s)$

$$M_k^{\infty} = C(E^{-1}A)^k E^{-1}B.$$

Das Ziel der Modellreduktion mit Hilfe des Momentenabgleichs ist, ein reduziertes System der Ordnung $m \ll n$ zu finden, dessen Momente einigen des originalen Systems gleichen. Bei diesem Verfahren wird also der Fehler im Frequenzbereich abgeschätzt. Damit die Momente der Übertragungsfunktionen um einen Punkt $s_0 \in \mathbb{C}$ übereinstimmen ohne diese explizit bestimmen zu müssen, sollten die Projektionsmatrizen $V \in \mathbb{R}^{n \times m}$ und $W \in \mathbb{R}^{n \times m}$ eine Basis der Krylov-Unterräume

$$\begin{aligned}
 \mathcal{K}_m \left((A - s_0E)^{-1}E, (A - s_0E)^{-1}B \right) &= \text{span}(V) \\
 \mathcal{K}_m \left((A - s_0E)^{-T}E^T, (A - s_0E)^{-T}C^T \right) &= \text{span}(W)
 \end{aligned}$$

bilden.

Sind mehrere Entwicklungspunkte $s_1, \dots, s_k \in \mathbb{C}$ gegeben, so lassen sich die Projektionsmatrizen $V \in \mathbb{R}^{n \times m}$ und $W \in \mathbb{R}^{n \times m}$ als Basis der Vereinigung der zu den Entwicklungspunkten gehörigen Krylov-Unterräume $\mathcal{K}_{m_i}((A - s_i E)^{-1} E, (A - s_i E)^{-1} B)$ bestimmen, wobei $\sum_{i=1}^k m_i = m$

men, wobei $\sum_{i=1}^k m_i = m$

$$\bigcup_{i=1}^k \mathcal{K}_{m_i}((A - s_i E)^{-1} E, (A - s_i E)^{-1} B) = \text{span}(V)$$

$$\bigcup_{i=1}^k \mathcal{K}_{m_i}((A - s_i E)^{-T} E^T, (A - s_i E)^{-T} C^T) = \text{span}(W).$$

Unter dieser Voraussetzung stimmen für einen einzelnen Entwicklungspunkt s_0 die ersten $2m$ Momente überein. Im Falle von k Entwicklungspunkten passen die ersten $2m_i$ Momente um den jeweiligen Entwicklungspunkt s_i für $i = 1, \dots, k$ zusammen.

Die exakte Herleitung der Markov-Parameter und die Beweise zu der Anzahl der übereinstimmenden Momente im reduzierten und originalen System sind nicht Gegenstand dieser Arbeit, können aber in [8] Kapitel 3 nachgeschlagen werden.

Weitere Eigenschaften von Krylov-Unterräumen, insbesondere der Vereinigung von Krylov-Unterräumen, sind in [20] Abschnitt 1.6.1 zu finden.

Die nächsten beiden Theorem zeigen eine weitere Möglichkeit, eine Projektionsmatrix $V \in \mathbb{R}^{n \times m}$ mit Hilfe einer Sylvester-Gleichung zu bestimmen und stellen damit einen Zusammenhang zwischen der Modellreduktionsmethode des Momentenabgleichs und dem Lösen einer Sylvester-Gleichung her.

Da zur Reduktion des originalen Systems eine weitere Projektionsmatrix $W \in \mathbb{R}^{n \times m}$ benötigt wird, kann der Einfachheit halber $W = V$ gewählt werden.

Für das Verständnis beider Theoreme wird zunächst die Definition der Beobachtbarkeit benötigt.

Definition 2.19. (Beobachtbarkeit, [14] Abschnitt 3.2.2)

Für zwei Matrizen $S \in \mathbb{R}^{m \times m}$ und $L \in \mathbb{R}^{p \times m}$ heißt das Matrixpaar (S, L) *beobachtbar*, wenn die Beobachtbarkeitsmatrix

$$Ob(S, L) = \begin{bmatrix} L \\ LS \\ LS^2 \\ \vdots \\ LS^{m-1} \end{bmatrix}$$

vollen Rang m hat.

Theorem 2.20. (Dualität für einen Entwicklungspunkt, [20] Abschnitt 2.3)

Es sei $s_0 \in \mathbb{C}$ ein Entwicklungspunkt, der kein Eigenwert von $E^{-1}A$ ist.

Dann bilden die Spalten von $V \in \mathbb{R}^{n \times m}$ genau dann eine Basis des Krylov-Unterraumes

$$\text{span}(V) = \mathcal{K}_m \left((A - s_0 E)^{-1} E, (A - s_0 E)^{-1} B \right),$$

wenn es ein beobachtbares Paar (S, L) gibt, wobei $L \in \mathbb{C}^{1 \times m}$ und $S \in \mathbb{C}^{m \times m}$, für die die Jordan-Normalform J

$$T^{-1}ST = J = \begin{bmatrix} s_0 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & s_0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & s_0 & 1 & 0 \\ \vdots & & & \ddots & s_0 & 1 \\ 0 & \cdots & \cdots & \cdots & 0 & s_0 \end{bmatrix}$$

mit einer Transformationsmatrix $T \in \mathbb{C}^{m \times m}$ existiert, sodass die Sylvester-Gleichung

$$AV - EVS = BL$$

erfüllt ist.

Darüber hinaus stimmen die ersten m Momente des reduzierten Systems mit denen des originalen Systems überein, falls s_0 kein Pol von $G_r(s)$ ist.

Theorem 2.21. (Dualität für mehrere Entwicklungspunkte, [20] Abschnitt 2.3)

Es seien m verschiedene Entwicklungspunkte $s_1, \dots, s_m \in \mathbb{C}$ gegeben, die keine Eigenwerte von $E^{-1}A$ sind.

Dann bilden die Spalten von $V \in \mathbb{R}^{n \times m}$ genau dann eine Basis der Vereinigung der zu den Entwicklungspunkten gehörigen Krylov-Unterräume

$$\text{span}(V) = \text{span}\{(A - s_1 E)^{-1} B, \dots, (A - s_m E)^{-1} B\},$$

wenn es ein beobachtbares Paar (S, L) gibt, wobei $L \in \mathbb{C}^{1 \times m}$ und $S \in \mathbb{C}^{m \times m}$, für die die Jordan-Normalform J

$$T^{-1}ST = J = \begin{bmatrix} s_1 & 0 & \cdots & \cdots & 0 \\ 0 & s_2 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & s_{m-1} & 0 \\ 0 & \cdots & \cdots & 0 & s_m \end{bmatrix} \text{ und } LT = \begin{bmatrix} 1 & \cdots & 1 \end{bmatrix}$$

mit einer Transformationsmatrix $T \in \mathbb{C}^{m \times m}$ existiert, sodass die Sylvester-Gleichung

$$AV - EVS = BL$$

erfüllt ist.

Darüber hinaus stimmen die zu den Entwicklungspunkten gehörigen Momente des reduzierten Systems mit denen des originalen Systems überein, falls kein s_i für $i = 1, \dots, m$ ein Pol von $G_r(s)$ ist.

Der Beweis dieser beiden Theoreme ist in [20] Abschnitt 2.3 beziehungsweise [18] Abschnitt 3.4 zu finden.

Die Quintessenz beider Theoreme ist, dass jede Basis eines Krylov-Unterraumes eine gewisse Sylvester-Gleichung löst, die aus einem beobachtbaren Matrixpaar (S, L) besteht, dessen Matrix $S \in \mathbb{R}^{m \times m}$ die Entwicklungspunkte des Momentenabgleichs als Eigenwerte besitzt. Die Matrix $L \in \mathbb{C}^{p \times m}$ spielt im Falle eines MIMO-Systems eine Rolle, da in ihren Spalten die tangentialen Richtungen gespeichert sind.

Andererseits spannt jede Lösung einer Sylvester-Gleichung mit einem beobachtbaren Paar (S, L) einen Krylov-Unterraum auf, wobei die Entwicklungspunkte durch die Eigenwerte der Matrix S gegeben sind.

In beiden Theoremen wird die Beobachtbarkeit des Matrixpaares nur dafür gebraucht, um sicher zu gehen, dass s_0 beziehungsweise s_1, \dots, s_m Entwicklungspunkte in den Polen sind.

Nach Theorem 3.23 in [18] sind die Eigenwerte von S immer eine Teilmenge der Interpolationspunkte und Pole zwischen den Übertragungsfunktionen des originalen und reduzierten Systems.

Kapitel 3

Modellreduktion von LTI-Systemen auf Basis von orthogonalen Polynomen

Im Folgenden sei ein LTI-System

$$\begin{aligned} E\dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{aligned} \tag{3.1}$$

gegeben, wobei $A, E \in \mathbb{R}^{n \times n}$ die System-, $B \in \mathbb{R}^{n \times p}$ die Eingangs-, $C \in \mathbb{R}^{q \times n}$ die Ausgangs-Matrizen, $x(t) \in \mathbb{R}^n$ der Zustands-, $u(t) \in \mathbb{R}^p$ der Eingangs-, $y(t) \in \mathbb{R}^q$ der Ausgangsvektor und $t \in \mathbb{R}$ die Zeit sind.

Da diese Arbeit, wie schon in Kapitel 1 erwähnt, ausschließlich von SISO-Systeme handelt, ist im Folgenden $p = q = 1$.

3.1 Grundidee

Die Idee in [13] ist, eine orthogonale Projektionsmatrix $Q \in \mathbb{R}^{n \times m}$ durch Lösen eines linearen Gleichungssystems

$$HV = F$$

zu generieren, wobei $m \in \mathbb{N}$ die gewünschte reduzierte Ordnung ist und die Matrix $H \in \mathbb{R}^{mn \times mn}$ beziehungsweise die rechte Seite $F \in \mathbb{R}^{mn}$ durch orthogonale Polynome erzeugt werden sollen.

Die Lösung

$$V = \begin{bmatrix} v_0 \\ \vdots \\ v_{m-1} \end{bmatrix} \in \mathbb{R}^{mn}$$

wird dann zu einer rechteckigen $(n \times m)$ -Matrix

$$V = [v_0 \dots v_{m-1}] \in \mathbb{R}^{n \times m}$$

umgeformt, aus der die Projektionsmatrix $Q \in \mathbb{R}^{n \times m}$ durch ein Orthogonalisierungsverfahren gewonnen wird, sodass folgende Bedingungen erfüllt sind:

$$\begin{aligned} \text{colspan}\{V\} &= \text{colspan}\{Q\}, \\ Q^T Q &= I_m. \end{aligned}$$

Das dynamische System soll dann mit Hilfe dieser Matrix Q in ein System

$$\begin{aligned} \underbrace{Q^T E Q}_{E_r} \dot{x}_r(t) &= \underbrace{Q^T A Q}_{A_r} x_r(t) + \underbrace{Q^T B}_{B_r} u(t) \\ y_r(t) &= \underbrace{C Q}_{C_r} x_r(t) \end{aligned}$$

mit reduzierter Ordnung $m \ll n$ transformiert werden, wobei der Zustand durch $x(t) = Qx_r(t)$ approximiert wird.

3.2 Algorithmus zur Modellreduktion

Um die Matrix H und den Vektor F zu ermitteln, werden in [13] der Zustandsvektor $x(t)$ und der Eingangsvektor $u(t)$ mit Hilfe einer endlichen, durch orthogonale Polynome gewichteten Summe approximiert

$$\begin{aligned} x(t) \approx x_m(t) &= \sum_{i=0}^{m-1} v_i g_i(t), \\ u(t) \approx u_m(t) &= \sum_{i=1}^{m-1} w_i \dot{g}_i(t), \end{aligned} \tag{3.2}$$

wobei $v_i \in \mathbb{R}^n$ Lösungsvektoren, $w_i \in \mathbb{R}^p$ Koeffizientenvektoren und g_i allgemeine orthogonale Polynome sind (siehe Abschnitt 2.1) für $i = 0, \dots, m-1$.

Das Ziel soll es nun sein, die obere Gleichung von (3.1) mithilfe der ersten Ableitung $\dot{g}_i(t)$ der orthogonalen Polynome für $i = 1, \dots, m-1$ darzustellen, sodass die Matrizen A , E und B von derselben Basis abhängen. Aus dieser Gleichheit soll sich ein lineares Gleichungssystem ergeben, durch das die Lösungsvektoren v_i für $i = 0, \dots, m-1$ berechnet werden können.

Dafür werden zunächst die Approximationen des Zustandsvektor $x(t)$ und des Eingangsvektors $u(t)$ aus (3.2) in die erste Gleichung von (3.1) eingesetzt und die orthogonalen

Polynome anschließend mit Hilfe der Differentialrekursionsformel (2.24) dargestellt, wobei beachtet wird, dass $\dot{g}_0(t) = 0 \forall t$, da $g_0(t) = \text{const}$ für alle orthogonalen Polynome gilt:

$$\begin{aligned}
& Bu(t) = E\dot{x}(t) - Ax(t) \\
\stackrel{(3.2)}{\Leftrightarrow} & B \left(\sum_{i=1}^{m-1} w_i \dot{g}_i(t) \right) = E \left(\sum_{i=0}^{m-1} v_i \dot{g}_i(t) \right) - A \left(v_0 g_0(t) + \sum_{i=1}^{m-1} v_i g_i(t) \right) \\
\stackrel{(2.24)}{\Leftrightarrow} & B \left(\sum_{i=1}^{m-1} w_i \dot{g}_i(t) \right) = E \left(\sum_{i=0}^{m-1} v_i \dot{g}_i(t) \right) - Av_0 g_0(t) - \\
& \quad - A \sum_{i=1}^{m-1} v_i (\alpha_i \dot{g}_{i+1}(t) + \beta_i \dot{g}_i(t) + \gamma_i \dot{g}_{i-1}(t)) \\
\stackrel{\substack{\dot{g}_0(t)=0 \\ \Leftrightarrow \\ \text{Indexverschiebung}}}{\Leftrightarrow} & B \left(\sum_{i=1}^{m-1} w_i \dot{g}_i(t) \right) = \sum_{i=1}^{m-1} (E - \beta_i A) v_i \dot{g}_i(t) - Av_0 g_0(t) - \sum_{i=2}^m \alpha_{i-1} Av_{i-1} \dot{g}_i(t) - \\
& \quad - \sum_{i=1}^{m-2} \gamma_{i+1} Av_{i+1} \dot{g}_i(t). \tag{3.3}
\end{aligned}$$

Unter der Berücksichtigung, dass $\dot{g}_1(t) = \text{const} \forall t$ ist, liefert ein koeffizientenweiser Vergleich für:

$i = 1$:

$$\begin{aligned}
& Bw_1 \dot{g}_1(t) = (E - \beta_1 A) v_1 \dot{g}_1(t) - Av_0 g_0(t) - \gamma_2 Av_2 \dot{g}_1(t) \\
\stackrel{\dot{g}_1(t) \in \mathbb{R}}{\Leftrightarrow} & Bw_1 = (E - \beta_1 A) v_1 - \frac{g_0(t)}{\dot{g}_1(t)} Av_0 - \gamma_2 Av_2
\end{aligned}$$

$i = 2, \dots, m-2$:

$$\begin{aligned}
& Bw_i \dot{g}_i(t) = (E - \beta_i A) v_i \dot{g}_i(t) - \alpha_{i-1} Av_{i-1} \dot{g}_i(t) - \gamma_{i+1} Av_{i+1} \dot{g}_i(t) \\
\stackrel{\dot{g}_i(t) \text{ beidseitig}}{\Leftrightarrow} & Bw_i = (E - \beta_i A) v_i - \alpha_{i-1} Av_{i-1} - \gamma_{i+1} Av_{i+1}
\end{aligned}$$

$i = m-1$:

$$\begin{aligned}
& Bw_{m-1} \dot{g}_{m-1}(t) = (E - \beta_{m-1} A) v_{m-1} \dot{g}_{m-1}(t) - \alpha_{m-2} Av_{m-2} \dot{g}_{m-1}(t) \\
\stackrel{\dot{g}_{m-1}(t) \text{ beidseitig}}{\Leftrightarrow} & Bw_{m-1} = (E - \beta_{m-1} A) v_{m-1} - \alpha_{m-2} Av_{m-2}.
\end{aligned}$$

Zusammen mit einer zusätzlichen Bedingung an den Anfangswert

$$x(t_0) = x_0,$$

in die die Approximation (3.2) eingesetzt wird

$$x_0 = x(t_0) \approx x_m(t_0) = \sum_{i=0}^{m-1} v_i g_i(t_0), \quad (3.4)$$

ist ein lineares Gleichungssystem der Form

$$HV = F \quad (3.5)$$

gegeben, wobei

$$H = \begin{bmatrix} g_0(t_0)I_n & g_1(t_0)I_n & g_2(t_0)I_n & g_3(t_0)I_n & g_4(t_0)I_n & g_5(t_0)I_n & \cdots & g_{m-1}(t_0)I_n \\ -\frac{g_0(t)}{\dot{g}_1(t)}A & E - \beta_1 A & -\gamma_2 A & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & -\alpha_1 A & E - \beta_2 A & -\gamma_3 A & \ddots & & & \vdots \\ \vdots & \ddots & -\alpha_2 A & E - \beta_3 A & -\gamma_4 A & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & -\alpha_{m-4} A & E - \beta_{m-3} A & -\gamma_{m-2} A & 0 \\ \vdots & & & & \ddots & -\alpha_{m-3} A & E - \beta_{m-2} A & -\gamma_{m-1} A \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & -\alpha_{m-2} A & E - \beta_{m-1} A \end{bmatrix},$$

$$V = \begin{bmatrix} v_0 \\ \vdots \\ v_{m-1} \end{bmatrix},$$

$$F = \begin{bmatrix} x_0 \\ Bw_1 \\ \vdots \\ Bw_{m-1} \end{bmatrix},$$

die Dimensionen $H \in \mathbb{R}^{mn \times mn}$, $F \in \mathbb{R}^{mn}$, $V \in \mathbb{R}^{mn}$ besitzen.

Kapitel 4

Lösen des linearen Gleichungssystems

In diesem Kapitel soll die in [13] präsentierte Methode zur Bestimmung der Projektionsmatrix erläutert werden. Ferner werden weitere Verfahren vorgestellt, um das Gleichungssystem (3.5) effizienter zu lösen.

Außerdem werden Alternativen aufgezeigt, um die Projektionsmatrix unabhängig von der Bedingung an den Anfangswert (3.4) aufzustellen.

Die Auswertung zu den in diesem Kapitel neu aufgezeigten Lösungsstrategien befindet sich im Kapitel 5.

4.1 Lösen mit Hilfe eines Splitting-Verfahrens

Um das Gleichungssystem (3.5) zu lösen wird in [13] ein iteratives Gauss-Seidel ähnliches Verfahren verwendet, bei dem die Matrix H in eine untere Block-Dreiecksmatrix $L_H \in \mathbb{R}^{mn \times mn}$ und eine strikte obere Block-Dreiecksmatrix $U_H \in \mathbb{R}^{mn \times mn}$ aufgeteilt wird:

$$H = \underbrace{\begin{bmatrix} g_0(t_0)I_n & 0 & \dots & \dots & \dots & 0 \\ -\frac{g_0(t)}{\dot{g}_1(t)}A & E - \beta_1 A & \ddots & & & \vdots \\ 0 & -\alpha_1 A & E - \beta_2 A & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & -\alpha_{m-3} A & E - \beta_{m-2} A & 0 \\ 0 & \dots & \dots & 0 & -\alpha_{m-2} A & E - \beta_{m-1} A \end{bmatrix}}_{=: L_H} +$$

$$+ \underbrace{\begin{bmatrix} 0 & g_1(t_0)I_n & g_2(t_0)I_n & g_3(t_0)I_n & \cdots & g_{m-1}(t_0)I_n \\ \vdots & \ddots & -\gamma_2 A & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & -\gamma_{m-2} A & 0 \\ \vdots & & & & \ddots & -\gamma_{m-1} A \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 \end{bmatrix}}_{=: U_H}.$$

Anschließend wird das resultierende Gleichungssystem mit einem Startvektor $x_0 \in \mathbb{R}^n$ und einer gegebenen Toleranzschranke tol für $k \geq 1$ iterativ gelöst

$$L_H x_{k+1} = F - U_H x_k$$

bis $\|x_{k+1} - x_k\|_2 < tol$ erfüllt ist.

Genauer zum Gauss-Seidel-Verfahren kann in [10] Abschnitt 10.1 nachgelesen werden.

Da dieses Verfahrens oftmals nur sehr langsam konvergiert, werden im Folgenden weitere Lösungsstrategien vorgestellt, die die Konvergenzgeschwindigkeit erhöhen sollen.

4.2 Lösen mit Hilfe eines vorkonditionierten Krylov-Unterraum-Verfahrens

Da die Matrix H sehr groß, aber dennoch dünnbesetzt ist, ist die erste Idee das Gleichungssystem (3.5) mit Hilfe des Krylov-Unterraum-Verfahrens Generalized Minimum Residual (GMRES) zu lösen, welches mit Hilfe eines geeigneten Vorkonditionierers noch effizienter ist.

Dieses Verfahren wird zunächst erläutert.

4.2.1 GMRES-Verfahren

Das GMRES-Verfahren ist ein iteratives Verfahren zum Lösen großer, dünnbesetzter Gleichungssysteme

$$Ax = b$$

mit $A \in \mathbb{R}^{n \times n}$ und $b \in \mathbb{R}^n$.

Die Idee dieses Verfahrens ist, eine unsymmetrische und invertierbare Matrix A mit Hilfe einer orthogonalen Matrix $Q_k = [q_1, \dots, q_k] \in \mathbb{R}^{n \times k}$ in eine Matrix $H_k \in \mathbb{R}^{k+1 \times k}$,

die in Hessenberg-Gestalt ist, zu transformieren

$$AQ_k = Q_{k+1}H_k = Q_{k+1} \begin{bmatrix} * & * & * & \cdots & * \\ * & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & * & * \\ 0 & \cdots & \cdots & 0 & * \end{bmatrix}.$$

Dieser Vorgang wird mit dem Arnoldi-Verfahren realisiert, bei dem zu der Matrix A und dem Vektor $q_1 := \frac{b - Ax_0}{\|b - Ax_0\|_2} \in \mathbb{R}^n$ mit Startvektor $x_0 \in \mathbb{R}^n$ eine orthogonale Basis des Krylov-Unterraumes $\mathcal{K}_k(A, q_1)$ berechnet wird. Weitere Informationen und Details zu diesem Verfahren können in [10] Abschnitt 9.4 nachgelesen werden.

Beim GMRES-Verfahren wird nun die Norm des Residuums $\|b - Ax_k\|_2$ über dem Krylov-Unterraum $x_0 + \mathcal{K}_k(A, q_1)$ minimiert. Dies wird insbesondere beim Betrachten des Residuums deutlich:

$$\begin{aligned} \|b - Ax_k\|_2 &= \|b - A(x_0 + Q_k y_k)\|_2 \\ &= \|r_0 - AQ_k y_k\|_2 \\ &= \|r_0 - Q_{k+1} H_k y_k\|_2 \\ &\stackrel{q_1 = \frac{r_0}{\|r_0\|}}{=} \left\| Q_{k+1} (\|r_0\|_2 e_1 - H_k y_k) \right\|_2 \\ &= \left\| \|r_0\|_2 e_1 - H_k y_k \right\|_2, \end{aligned}$$

wobei $e_1 \in \mathbb{R}^{k+1}$ der erste Einheitsvektor ist.

Das GMRES-Verfahren konvergiert nach spätestens n Schritten.

Eine ausführliche Präsentation dieses Verfahrens kann beispielsweise in [10] Abschnitt 10.4.4 gefunden werden.

4.2.2 Vorkonditionierung

Um das GMRES-Verfahren noch effizienter zu machen, werden in der Praxis oftmals Vorkonditionierer eingesetzt.

Definition 4.1. (Vorkonditionierer, [7] Abschnitt 6.6.5)

Für ein Gleichungssystem

$$Ax = b$$

mit einer Matrix $A \in \mathbb{R}^{n \times n}$ und Vektoren $b, x \in \mathbb{R}^n$ ist ein *Vorkonditionierer* eine invertierbare Matrix $P \in \mathbb{R}^{n \times n}$, die die Matrix A sehr gut approximiert, sodass sich das Gleichungssystem

$$P^{-1}Ax = P^{-1}b$$

mit wenig Aufwand lösen lässt und die Matrix $P^{-1}A$ besser konditioniert ist. Der ideale Vorkonditionierer wäre daher $P = A$.

Da die Matrix H des hergeleiteten Gleichungssystems (3.5) sehr groß, aber dennoch dünnbesetzt ist, ist das GMRES-Verfahren zur Lösung ideal geeignet. Da diese Matrix oftmals auch schlecht konditioniert ist, kann ein Vorkonditionierer sehr nützlich sein.

In dieser Arbeit wurden zwei Vorkonditionierer getestet.

Einer der verwendeten Vorkonditionierer wurde als untere Block-Dreiecksmatrix von H gewählt

$$P_1 = \begin{bmatrix} g_0(t_0)I_n & 0 & \cdots & \cdots & \cdots & 0 \\ -\frac{g_0(t)}{\dot{g}_1(t)}A & E - \beta_1A & \ddots & & & \vdots \\ 0 & -\alpha_1A & E - \beta_2A & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & -\alpha_{m-3}A & E - \beta_{m-2}A & 0 \\ 0 & \cdots & \cdots & 0 & -\alpha_{m-2}A & E - \beta_{m-1}A \end{bmatrix} \in \mathbb{R}^{mn \times mn} \quad (4.1)$$

und entspricht damit der Matrix L_H aus dem Gauss-Seidel ähnlichen Verfahren.

Ein weiterer getesteter Vorkonditionierer lässt sich aus dem kommenden Abschnitt 4.3 ableiten und hat die Gestalt

$$P_2 = \begin{bmatrix} g_0(t_0)A & g_1(t_0)A & g_2(t_0)A & g_3(t_0)A & \cdots & \cdots & g_{m-1}(t_0)A \\ -\frac{g_0(t)}{\dot{g}_1(t)}A & -\beta_1A & -\gamma_2A & 0 & \cdots & \cdots & 0 \\ 0 & -\alpha_1A & -\beta_2A & -\gamma_3A & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & -\alpha_{m-4}A & -\beta_{m-3}A & -\gamma_{m-2}A & 0 \\ \vdots & & & \ddots & -\alpha_{m-3}A & -\beta_{m-2}A & -\gamma_{m-1}A \\ 0 & \cdots & \cdots & \cdots & 0 & -\alpha_{m-2}A & -\beta_{m-1}A \end{bmatrix} \in \mathbb{R}^{mn \times mn}. \quad (4.2)$$

4.3 Umformulierung des Gleichungssystems als Sylvester-Gleichung

Da das direkte Lösen des linearen Gleichungssystems (3.5) für große Werte von n sehr viel Speicher beansprucht und für wachsendes m sehr aufwändig ist, wird in diesem Abschnitt das Gleichungssystem (3.5) umgeformt.

Dafür wird zunächst die Matrix H als Summe von Kronecker-Produkten dargestellt:

$$H = H_1 \otimes I_n + H_2 \otimes A + H_3 \otimes E \quad (4.3)$$

mit den Matrizen

$$H_1 := \begin{bmatrix} g_0(t_0) & \cdots & g_{m-1}(t_0) \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \in \mathbb{R}^{m \times m},$$

$$H_2 := - \begin{bmatrix} 0 & 0 & 0 & \cdots & \cdots & \cdots & 0 \\ \frac{g_0(t)}{g_1(t)} & \beta_1 & \gamma_2 & \ddots & & & \vdots \\ 0 & \alpha_1 & \beta_2 & \gamma_3 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \alpha_{m-4} & \beta_{m-3} & \gamma_{m-2} & 0 \\ \vdots & & & \ddots & \alpha_{m-3} & \beta_{m-2} & \gamma_{m-1} \\ 0 & \cdots & \cdots & \cdots & 0 & \alpha_{m-2} & \beta_{m-1} \end{bmatrix} \in \mathbb{R}^{m \times m},$$

$$H_3 := \begin{bmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & 0 \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix} \in \mathbb{R}^{m \times m}.$$

Für die weiteren Umformungen seien

$$\text{vec}(X) := V \in \mathbb{R}^{mn},$$

$$\text{vec}(R) := F \in \mathbb{R}^{mn},$$

wobei $\text{vec}(\cdot)$ die im Abschnitt 2.2 definierte Vektorisierung einer Matrix ist.

Anschließend wird die in (2.29) aufgezeigte Äquivalenzrelation ausgenutzt, um das lineare Gleichungssystem (3.5) in eine Sylvester-Gleichung zu transformieren:

$$HV = F$$

$$\begin{aligned}
&\Leftrightarrow H\text{vec}(X) = \text{vec}(R) \\
(4.3) \quad &\Leftrightarrow (H_1 \otimes I_n + H_2 \otimes A + H_3 \otimes E)\text{vec}(X) = \begin{bmatrix} x_0 \\ Bw_1 \\ \vdots \\ Bw_{m-1} \end{bmatrix} \\
&\stackrel{\star}{\Leftrightarrow} (H_1 \otimes A + H_2 \otimes A + H_3 \otimes E)\text{vec}(X) = \underbrace{\begin{bmatrix} Ax_0 \\ Bw_1 \\ \vdots \\ Bw_{m-1} \end{bmatrix}}_{\text{vec}(\tilde{R})} \\
&\Leftrightarrow ((H_1 + H_2) \otimes A + H_3 \otimes E)\text{vec}(X) = \text{vec}(\tilde{R}) \\
(2.29) \quad &\Leftrightarrow AX \underbrace{(H_1 + H_2)^T}_{\tilde{E}} + EX \underbrace{H_3^T}_{\tilde{A}} = \tilde{R} \\
&\Leftrightarrow AX\tilde{E} + EX\tilde{A} = \tilde{R}. \tag{4.4}
\end{aligned}$$

★Da nur H_1 für die erste Komponente von $\text{vec}(R)$ verantwortlich ist, kann diese Zeile mit A durchmultipliziert werden.

Dabei sind die Matrizen gegeben durch

$$\tilde{A} = \begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & 0 \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix} \in \mathbb{R}^{m \times m},$$

$$\tilde{E} = \begin{bmatrix} g_0(t_0) & -\frac{g_0(t)}{\dot{g}_1(t)} & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ g_1(t_0) & -\beta_1 & -\alpha_1 & \ddots & & & & \vdots \\ g_2(t_0) & -\gamma_2 & -\beta_2 & -\alpha_2 & \ddots & & & \vdots \\ g_3(t_0) & 0 & -\gamma_3 & -\beta_3 & -\alpha_3 & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ g_{m-3}(t_0) & \vdots & & \ddots & -\gamma_{m-3} & -\beta_{m-3} & -\alpha_{m-3} & 0 \\ g_{m-2}(t_0) & \vdots & & & \ddots & -\gamma_{m-2} & -\beta_{m-2} & -\alpha_{m-2} \\ g_{m-1}(t_0) & 0 & \cdots & \cdots & \cdots & 0 & -\gamma_{m-1} & -\beta_{m-1} \end{bmatrix} \in \mathbb{R}^{m \times m},$$

$$\tilde{R} = \begin{bmatrix} Ax_0 & Bw_1 & \cdots & Bw_{m-1} \end{bmatrix} \in \mathbb{R}^{n \times m},$$

$$X = \begin{bmatrix} v_0 & \dots & v_{m-1} \end{bmatrix} \in \mathbb{R}^{n \times m}.$$

Also kann das lineare Gleichungssystem (3.5) zu einer Sylvester-Gleichung (4.4) umformuliert werden, die sich aufgrund ihrer besonderen Struktur effizient mit einem speziellen Verfahren für dünnbesetzte-dichtbesetzte (sparse-dense) Sylvester-Gleichungen lösen lässt.

Definition 4.2. (sparse-dense Sylvester-Gleichung, [3])

Eine Sylvester-Gleichung wird *sparse-dense* genannt, wenn für eine Sylvester-Gleichung

$$AXF + EXH = M$$

$M, X \in \mathbb{R}^{n \times m}$, die Matrizen $A \in \mathbb{R}^{n \times n}$ und $E \in \mathbb{R}^{n \times n}$ groß und dünnbesetzt und die Matrizen $F \in \mathbb{R}^{m \times m}$ beziehungsweise $H \in \mathbb{R}^{m \times m}$ klein und dichtbesetzt sind.

Weitere Informationen zu einem solchen sparse-dense Löser können in [3] gefunden werden.

Der in Abschnitt 4.2.2 vorgestellte Vorkonditionierer P_2 ergibt sich demnach aus dem Kronecker-Produkt $\tilde{E}^T \otimes A$.

Mit Hilfe der Sylvester-Gleichung (4.4) und Satz 2.17. ist es möglich, eine Aussage über die Eindeutigkeit der Lösung X und damit auch der Lösung V zu treffen.

Da die Matrix \tilde{A} eine Diagonalmatrix ist, kann das Spektrum direkt abgelesen werden:

$$\Lambda(\tilde{A}) = \{0, 1\}.$$

Damit ist nach (2.35)

$$\Lambda(\tilde{A}^T \otimes E) = \Lambda(E) \cup \{0\}.$$

Für eine eindeutige Lösbarkeit der Sylvester-Gleichung muss für gemeinsame Eigenvektoren von $\tilde{A}^T \otimes E$ und $\tilde{E}^T \otimes A$ nach Satz 2.17.

$$\Lambda(\tilde{E}^T \otimes A) \cap \Lambda(-\tilde{A}^T \otimes E) = \emptyset$$

gelten.

Allerdings kann es aufgrund der zusätzlichen Bedingung an den Anfangswert (3.4) dazu kommen, dass \tilde{E} nicht mehr invertierbar und damit mindestens ein Eigenwert 0 ist und somit die Sylvester-Gleichung beziehungsweise das lineare Gleichungssystem (3.5) nicht mehr eindeutig lösbar ist.

Daher werden in dem folgenden Abschnitt zwei Varianten der auf orthogonalen Polynomen basierenden Modellreduktion vorgestellt, die wie das Verfahren des balancierten Abschneidens unabhängig vom gewählten Anfangswert sind oder wie beim Momentenabgleich eine Bedingung an den Anfangswert stellen.

4.4 Varianten der Modellreduktion

Da wie im vorangegangenen Abschnitt erörtert eine Bedingung an den Anfangswert nicht immer wünschenswert ist, werden im Folgenden zwei Varianten vorgestellt.

Dies sind zum einen eine Modellreduktion, bei der auf den Anfangswert ganz verzichtet wird und zum anderen eine Variante, die grundsätzlich eine Bedingung an den Anfangswert stellt und damit insgesamt unabhängig von der künstlich geschaffenen Zusatzbedingung ist.

4.4.1 Variante 1: Ausnutzen der gegebenen Bedingungen

Um eine quadratische Matrix für ein Gleichungssystem zu erhalten, braucht nicht wie in der in [13] vorgeschlagenen Modellreduktion eine künstliche Bedingung hinzugefügt werden.

Beim Betrachten der Approximationsergebnisse (3.3) wird deutlich, dass es eine weitere Bedingung für $i = m$ gibt, die bisher nicht berücksichtigt wurde:

$$0 = \alpha_{m-1}Av_{m-1}.$$

Anhand der im Abschnitt 3.2 hergeleiteten weiteren Bedingungen für $i = 1, \dots, m - 1$, kann ein Gleichungssystem

$$\hat{H}V = \hat{F} \tag{4.5}$$

mit

$$\hat{H} = \begin{bmatrix} -\frac{g_0(t)}{\dot{g}_1(t)}A & E - \beta_1A & -\gamma_2A & 0 & \cdots & \cdots & 0 \\ 0 & -\alpha_1A & E - \beta_2A & -\gamma_3A & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & -\alpha_{m-4}A & E - \beta_{m-3}A & -\gamma_{m-2}A & 0 \\ \vdots & & & \ddots & -\alpha_{m-3}A & E - \beta_{m-2}A & -\gamma_{m-1}A \\ \vdots & & & & \ddots & -\alpha_{m-2}A & E - \beta_{m-1}A \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & -\alpha_{m-1}A \end{bmatrix} \in \mathbb{R}^{mn \times mn},$$

$$V = \begin{bmatrix} v_0 \\ \vdots \\ v_{m-1} \end{bmatrix} \in \mathbb{R}^{mn},$$

$$\hat{F} = \begin{bmatrix} Bw_1 \\ \vdots \\ Bw_{m-1} \\ 0 \end{bmatrix} \in \mathbb{R}^{mn}$$

definiert werden.

Dieses Gleichungssystem kann wie schon zuvor mithilfe des Kroneckerproduktes und der Vektorisierung einer Matrix in eine Sylvester-Gleichung überführt werden.

Für die nachfolgenden Äquivalenzumformungen sei die Matrix \hat{H} durch eine Summe von Kronecker-Produkten dargestellt

$$\hat{H} = \hat{H}_2 \otimes A + \hat{H}_3 \otimes E \quad (4.6)$$

und die Vektoren

$$\text{vec}(X) := V,$$

$$\text{vec}(\hat{\mathcal{R}}) := \hat{F}$$

definiert, sodass das Gleichungssystem (4.5) folgendermaßen umformuliert werden kann:

$$\begin{aligned} & \hat{H}V = \hat{F} \\ \stackrel{(4.6)}{\Leftrightarrow} & (\hat{H}_2 \otimes A + \hat{H}_3 \otimes E) \text{vec}(X) = \text{vec}(\hat{\mathcal{R}}) \\ \stackrel{(2.29)}{\Leftrightarrow} & AX \underbrace{\hat{H}_2^T}_{=: \hat{E}} + EX \underbrace{\hat{H}_3^T}_{=: \hat{A}} = \hat{\mathcal{R}} \\ \Leftrightarrow & AX\hat{E} + EX\hat{A} = \hat{\mathcal{R}}. \end{aligned} \quad (4.7)$$

Dabei sind

$$\hat{E} = \hat{H}_2^T = - \begin{bmatrix} \frac{g_0(t)}{\dot{g}_1(t)} & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ \beta_1 & \alpha_1 & \ddots & & & & \vdots \\ \gamma_2 & \beta_2 & \alpha_2 & \ddots & & & \vdots \\ 0 & \gamma_3 & \beta_3 & \alpha_3 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \gamma_{m-2} & \beta_{m-2} & \alpha_{m-2} & 0 \\ 0 & \cdots & \cdots & 0 & \gamma_{m-1} & \beta_{m-1} & \alpha_{m-1} \end{bmatrix} \in \mathbb{R}^{m \times m},$$

$$\hat{A} = \hat{H}_3^T = \begin{bmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ 1 & \ddots & & & \vdots \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix} \in \mathbb{R}^{m \times m},$$

$$\hat{R} = \begin{bmatrix} Bw_1 & \cdots & Bw_{m-1} & 0 \end{bmatrix} \in \mathbb{R}^{n \times m}.$$

Diese Sylvester-Gleichung hat den Vorteil, dass die Lösung der Sylvester-Gleichung im Falle einer invertierbaren Matrix A immer eindeutig ist. Der Grund dafür ist, dass das Produkt aller möglichen Eigenwerte von E und \hat{A}

$$\lambda = 0$$

ist, denn \hat{A} ist eine untere Diagonalmatrix, auf deren Hauptdiagonale die Eigenwerte 0 zu finden sind.

Da \hat{E} ebenfalls eine untere Diagonalmatrix ist und für jede beliebige Wahl von orthogonalen Polynomen $\alpha_i \neq 0$ für $i = 1, \dots, m$ und der Quotient $\frac{g_0(t)}{\dot{g}_1(t)} \neq 0 \forall t$ ist, sind alle Eigenwerte von 0 verschieden.

Somit ist die Bedingung (2.36) an die eindeutige Lösbarkeit einer Sylvester-Gleichung nur im Falle einer singulären Matrix A nicht erfüllt.

4.4.2 Variante 2: Bedingungen an den Anfangswert

Für eine weitere Variante der Modellreduktion mit Hilfe von orthogonalen Polynomen, die auf eine künstliche Zusatzbedingung verzichtet, wird sich an einem konkurrierenden Modellreduktionsverfahren, dem Momentenabgleich, orientiert.

In diesem Verfahren wird davon ausgegangen, dass

$$x(t_0) = x_0 \quad \Leftrightarrow \quad x(0) = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{R}^n,$$

gilt, da dies für LTI-Systeme, wie sie in dieser Arbeit eine Rolle spielen, keine Einschränkung bedeutet und durch eine Transformation immer erreicht werden kann.

Aufgrund dieser Bedingung ist es möglich, die Approximation mit dem linearen Polynom $g_1(t)$ statt mit dem konstanten Polynom $g_0(t)$ zu beginnen, da ein konstanter Faktor $g_0(t) \neq 0 \forall t$ nicht mehr benötigt wird.

Da das Weglassen der künstlichen Bedingung (3.4) und der Verzicht auf das konstante Polynom $g_0(t)$ ein Gleichungssystem mit nur $m - 1$ Lösungsvektoren erzeugen würde, die Projektionsmatrix aber m Spalten haben soll, muss die reduzierte Ordnung um 1

erhöht werden, also $m \rightarrow m + 1$.

Somit werden der Zustandsvektor $x(t)$ und der Eingangsvektor $u(t)$ im Folgenden durch

$$\begin{aligned} x(t) &\approx x_{m+1}(t) = \sum_{i=1}^m v_i g_i(t), \\ u(t) &\approx u_{m+1}(t) = \sum_{i=1}^m w_i \dot{g}_i(t) \end{aligned} \quad (4.8)$$

approximiert.

Ein Einsetzen dieser Bedingungen und der Differentialrekursionsformel (2.24) in das LTI-System (3.1) liefert dann:

$$\begin{aligned} &Bu(t) = E\dot{x}(t) - Ax(t) \\ \stackrel{(4.8)}{\Leftrightarrow} &B \sum_{i=1}^m w_i \dot{g}_i(t) = E \sum_{i=1}^m v_i \dot{g}_i(t) - A \sum_{i=1}^m v_i g_i(t) \\ \stackrel{(2.24)}{\Leftrightarrow} &B \sum_{i=1}^m w_i \dot{g}_i(t) = E \sum_{i=1}^m v_i \dot{g}_i(t) - A \sum_{i=1}^m v_i (\alpha_i \dot{g}_{i+1}(t) + \beta_i \dot{g}_i(t) + \gamma_i \dot{g}_{i-1}(t)) \\ \stackrel{\substack{\dot{g}_0(t)=0 \\ \text{Indexverschiebung}}}{\Leftrightarrow} &B \sum_{i=1}^m w_i \dot{g}_i(t) = \sum_{i=1}^m (E - \beta_i A) v_i \dot{g}_i(t) - \sum_{i=2}^{m+1} \alpha_{i-1} A v_{i-1} \dot{g}_i(t) - \sum_{i=1}^{m-1} \gamma_{i+1} A v_{i+1} \dot{g}_i(t). \end{aligned}$$

Unter Berücksichtigung, dass für jedes $i = 1, \dots, m$ $\dot{g}_i(t)$ auf beiden Seiten ausgeklammert werden kann, ergibt sich beim koeffizientenweisen Vergleich für $i = 1$:

$$Bw_1 = (E - \beta_1 A) v_1 - \gamma_2 A v_2$$

$i = 2, \dots, m - 1$:

$$Bw_i = (E - \beta_i A) v_i - \alpha_{i-1} A v_{i-1} - \gamma_{i+1} A v_{i+1}$$

$i = m$:

$$Bw_m = (E - \beta_m A) v_m - \alpha_{m-1} A v_{m-1},$$

auf dessen Basis sich ein lineares Gleichungssystem

$$\mathcal{H}\mathcal{V} = \mathcal{F} \quad (4.9)$$

aufstellen lässt mit

$$\mathcal{H} = \begin{bmatrix} E - \beta_1 A & -\gamma_2 A & 0 & \cdots & \cdots & \cdots & 0 \\ -\alpha_1 A & E - \beta_2 A & -\gamma_3 A & \ddots & & & \vdots \\ 0 & -\alpha_2 A & E - \beta_3 A & -\gamma_4 A & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & -\alpha_{m-3} A & E - \beta_{m-2} A & -\gamma_{m-1} A & 0 \\ \vdots & & & \ddots & -\alpha_{m-2} A & E - \beta_{m-1} A & -\gamma_m A \\ 0 & \cdots & \cdots & \cdots & 0 & -\alpha_{m-1} A & E - \beta_m A \end{bmatrix} \in \mathbb{R}^{mn \times mn},$$

$$\mathcal{V} = \begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix} \in \mathbb{R}^{mn},$$

$$\mathcal{F} = \begin{bmatrix} Bw_1 \\ \vdots \\ Bw_m \end{bmatrix} \in \mathbb{R}^{mn}.$$

Aufgrund der speziellen Struktur kann die Matrix \mathcal{H} als Summe von Kronecker-Produkten

$$\mathcal{H} = \mathcal{H}_2 \otimes A + \mathcal{H}_3 \otimes E \quad (4.10)$$

dargestellt werden. Zusammen mit denen durch die Vektorisierung $\text{vec}(\cdot)$ definierten Vektoren

$$\begin{aligned} \text{vec}(\mathcal{X}) &:= \mathcal{V}, \\ \text{vec}(\mathcal{R}) &:= \mathcal{F} \end{aligned}$$

kann das lineare Gleichungssystem (4.9) in eine Sylvester-Gleichung umgeformt werden

$$\begin{aligned} &\mathcal{H}\mathcal{V} = \mathcal{F} \\ \stackrel{(4.10)}{\Leftrightarrow} &(\mathcal{H}_2 \otimes A + \mathcal{H}_3 \otimes E) \text{vec}(\mathcal{X}) = \text{vec}(\mathcal{R}) \\ \stackrel{(2.29)}{\Leftrightarrow} &A\mathcal{X} \underbrace{\mathcal{H}_2^T}_{=: \mathcal{E}} + E\mathcal{X} \underbrace{\mathcal{H}_3^T}_{=: \mathcal{A}} = \mathcal{R} \\ \Leftrightarrow &A\mathcal{X}\mathcal{E} + E\mathcal{X}\mathcal{A} = \mathcal{R}. \end{aligned} \quad (4.11)$$

Die Matrizen haben dabei folgende Gestalt:

$$\mathcal{A} = \mathcal{H}_3^T = \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & 0 \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix} \in \mathbb{R}^{m \times m},$$

$$\mathcal{E} = \mathcal{H}_2^T = - \begin{bmatrix} \beta_1 & \alpha_1 & 0 & \cdots & \cdots & \cdots & 0 \\ \gamma_2 & \beta_2 & \alpha_2 & \ddots & & & \vdots \\ 0 & \gamma_3 & \beta_3 & \alpha_3 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \gamma_{m-2} & \beta_{m-2} & \alpha_{m-2} & 0 \\ \vdots & & & \ddots & \gamma_{m-1} & \beta_{m-1} & \alpha_{m-1} \\ 0 & \cdots & \cdots & \cdots & 0 & \gamma_m & \beta_m \end{bmatrix} \in \mathbb{R}^{m \times m},$$

$$\mathcal{R} = \begin{bmatrix} Bw_1 & \cdots & Bw_m \end{bmatrix} \in \mathbb{R}^{n \times m}.$$

Aufgrund der Einheitsmatrix \mathcal{A} lässt sich die Sylvester-Gleichung vereinfacht schreiben als

$$AX\mathcal{E} + EX = \mathcal{R}. \quad (4.12)$$

Die Eindeutigkeit der Lösung dieser Sylvester-Gleichung (4.12) und damit auch die des linearen Gleichungssystems (4.9) lässt sich mit Hilfe des Satzes 2.17. über die eindeutige Lösbarkeit einer verallgemeinerten Sylvester-Gleichung bestimmen und ist für gemeinsame Eigenvektoren zu den Eigenwerten der Matrizen A und E gegeben, falls

$$\Lambda(\mathcal{E}^T \otimes A) \cap \Lambda(-E) = \emptyset.$$

Kapitel 5

Experimentelle Auswertung

In diesem Kapitel sollen einige der in Kapitel 4 vorgestellten Ideen anhand von verschiedenen Beispielen ausgewertet werden.

Die Rahmenbedingungen werden dafür im folgenden Abschnitt vorgestellt. Im Anschluss daran werden zunächst die verschiedenen Beispiele kurz vorgestellt und die getesteten Verfahren bewertet.

Abschließend werden die in der Auswertung aufgetretenen Auffälligkeiten erklärt und eine Äquivalenz zum Modellreduktionsverfahren des Momentenabgleichs in speziellen Fällen aufgezeigt.

5.1 Rahmenbedingungen der Auswertung

Das Testsystem für die Auswertung war das Folgende:

- Asus F750JB-TY010H,
- Intel Quad Core i7-4700HQ ($4 \times 2.40\text{GHz}$, 6 MB Cache),
- 8 GB DDR3-1600,
- Betriebssystem: Windows 7,
- Matlab 7.9.0 (R2009b).

Zu den ausgewerteten Verfahren gehören die Methoden, die die zur Modellreduktion benötigte Projektionsmatrix durch

- direktes Lösen des linearen Gleichungssystems (3.5),
- Lösen des linearen Gleichungssystems (3.5) mit Hilfe des vorkonditionierten GM-RES (Vorkonditionierer: Block-Dreiecksmatrix (4.1)),
- Lösen des linearen Gleichungssystems (3.5) mit Hilfe des vorkonditionierten GM-RES (Vorkonditionierer: Kronecker-Produkt (4.2)),

- Lösen der zu (3.5) äquivalenten Sylvester-Gleichung (4.4),
- Lösen der Anfangswert-unabhängigen Sylvester-Gleichung (4.7) (Variante 1),
- Lösen der von der Zusatzbedingung unabhängigen Sylvester-Gleichung (4.12) (Variante 2)

bestimmen.

Auf das in Abschnitt 4.1 vorgestellte Splitting-Verfahren wird dabei nicht näher eingegangen, da dies schon in [13] bewertet wurde.

Um aussagekräftige Ergebnisse erzielen zu können, wurden für alle direkten Lösungsverfahren, das heißt das Lösen des Gleichungssystems (3.5) und der Sylvester-Gleichungen (4.4), (4.7) beziehungsweise (4.12), jeweils 50 Testdurchläufe durchgeführt.

Eine Ausnahme bildet hierbei das iterative und vorkonditionierte GMRES-Verfahren.

Bei dieser Methode wurde die Toleranzschranke auf $tol = 10^{-10}$ gesetzt und falls diese nicht erreicht wird, bricht der Algorithmus bei einer Iterationszahl von $maxit = 5000$ ab.

In diesem Fall wurden aufgrund der enormen Rechenzeiten nur 5 Testdurchläufe gemacht. Dies ist damit zu begründen, dass der Aufwand und Nutzen in keiner Relation stehen, da die Zeit, die zur Modellreduktion und zum Lösen des reduzierten Modells benötigt wird, die Lösungszeit des Originalmodells übersteigt.

Für alle getesteten Beispiele wurden der absolute und relative Fehler

$$err_{abs} = \|y(t) - y_r(t)\|_2,$$

$$err_{rel} = \left\| \frac{y(t) - y_r(t)}{y(t)} \right\|_2$$

ermittelt und grafisch dargestellt. Dabei bezeichnet $y(t)$ den Ausgang des Originalmodells zum Zeitpunkt t und $y_r(t)$ den des reduzierten Modells zum Zeitpunkt t . Somit beziehen sich die absoluten und relativen Fehler auf das gesamte Zeitintervall, das t durchläuft. Die Abhängigkeit zu t wird daher in den grafischen Darstellungen weggelassen.

In diesem Kapitel wird jedoch auf die Grafiken des absoluten Fehlers verzichtet, da die Verläufe des relativen Fehlers denen des absoluten Fehlers ähneln. Diese Abbildungen sind im Anhang B zu finden.

Desweiteren wird in den nachfolgenden Abschnitten die Ausführungszeit, das heißt die gesamte Zeit, die zur Modellreduktion und zum Lösen des reduzierten Modells benötigt wird, der Zeit zum Lösen des Originalsystems grafisch gegenübergestellt, um eine Aussage über die Effizienz der Verfahren treffen zu können.

Um einen Eindruck von den Ausgangsvektoren $y(t)$ in dem getesteten Zeitintervall t zu erhalten, wird diese Kurve, die Trajektorie genannt wird, ebenfalls für das originale Modell aller Beispiele dargestellt. Die Trajektorien der reduzierten Modelle sind in nahezu

allen Fällen nicht von der Originaltrajektorie zu unterscheiden. Aus diesem Grund wird die Kurve der Ausgangsvektoren der reduzierten Modelle nur in einigen Fällen für das größte gewählte m grafisch dargestellt.

Die Ausgangsvektoren $y(t)$ und $y_r(t)$ wurden in dieser Arbeit mit Hilfe des Rückwärts-Euler-Verfahrens berechnet. Detaillierte Informationen zu dieser impliziten numerischen Methode zum Lösen von Anfangswertproblemen sind in [11] in den Kapiteln 1 und 2 dargestellt.

Für alle Beispiele wurden die Berechnungen unter denselben Bedingungen durchgeführt. Diese sind:

- Startvektor $x_0 = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{R}^n$,
- Startzeitpunkt $t_0 = 0$,
- Zeitintervall $t \in [0, 1]$,
- Zeitschrittweite $\tau = 0.001$,
- Eingang $u(t) = \begin{cases} 0 & , t \in [0, 0.1) \\ \sin(\pi(10t - 1.5)) & , t \in [0.1, 0.2) \\ 1 & , t \in [0.2, 1] \end{cases}$.

Die folgenden Grafiken beziehen sich immer auf das in der Überschrift genannte Lösungsverfahren, aus dem die Projektionsmatrix zur Reduktion des Originalmodells gewonnen wird.

In allen Grafiken haben die getesteten Hermite-, Laguerre-, Legendre- und beide Chebychev-Polynome dieselbe Farbe, können aber aufgrund der Übersichtlichkeit andere Strukturen (durchgezogenen, gestrichelte, gepunktete Linie oder auch Marker) besitzen.

In der Abbildung, die die Ausführungszeit veranschaulicht, ist die von der reduzierten Ordnung m unabhängige Zeit zum Lösen des Originalsystems als durchgezogene Linie dargestellt.

Die in Kapitel 4 vorgestellten Lösungsstrategien werden in den Titeln der Grafiken wie folgt abgekürzt:

- Backslash: direktes Lösen des linearen Gleichungssystems (3.5),
- Sylvester-Gleichung: Lösen der zum linearen Gleichungssystem (3.5) äquivalenten Sylvester-Gleichung (4.4),
- unabhängige Variante 1: Lösen der Anfangswert-unabhängigen Sylvester-Gleichung (4.7) (Variante 1),
- zusatzfreie Variante 2: Lösen der von der Zusatzbedingung unabhängigen Sylvester-Gleichung (4.12) (Variante 2),

- GMRES mit Vorkonditionierer 1: Lösen des linearen Gleichungssystems (3.5) mit Hilfe des vorkonditionierten GMRES (Vorkonditionierer: Block-Dreiecksmatrix (4.1)),
- GMRES mit Vorkonditionierer 2: Lösen des linearen Gleichungssystems (3.5) mit Hilfe des vorkonditionierten GMRES (Vorkonditionierer: Kronecker-Produkt (4.2)).

5.2 Auswertung

5.2.1 Beispiel 1

Das erste Beispiel ist allgemein gehalten und stammt aus [13], bei dem die Matrizen und Vektoren wie folgt gewählt wurden:

$$A = \begin{bmatrix} A_1 & 0 & 0 & 0 \\ 0 & A_2 & 0 & 0 \\ 0 & 0 & A_3 & 0 \\ 0 & 0 & 0 & A_4 \end{bmatrix} \in \mathbb{R}^{1006 \times 1006},$$

wobei

$$A_1 = \begin{bmatrix} -1 & 100 \\ -100 & -1 \end{bmatrix} \in \mathbb{R}^{2 \times 2}, \quad A_2 = \begin{bmatrix} -1 & 200 \\ -200 & -1 \end{bmatrix} \in \mathbb{R}^{2 \times 2},$$

$$A_3 = \begin{bmatrix} -1 & 400 \\ -400 & -1 \end{bmatrix} \in \mathbb{R}^{2 \times 2}, \quad A_4 = \begin{bmatrix} -1 & 0 & \cdots & 0 \\ 0 & -2 & \ddots & \vdots \\ \vdots & \ddots & -999 & 0 \\ 0 & \cdots & 0 & -1000 \end{bmatrix} \in \mathbb{R}^{1000 \times 1000},$$

$$B^T = C = [\underbrace{6 \dots 6}_6 \underbrace{1 \dots 1}_{1000}] \in \mathbb{R}^{1 \times 1006}.$$

Dieses Zustandsraummodell hat demzufolge die Dimension 1006 und $E = I_{1006}$ ist als Einheitsmatrix gegeben.

Beim Betrachten der Abbildung 5.1 für den relativen Fehler fallen sofort drei Verfahren auf, bei denen der relative Fehler selbst für das größte gewählte $m = 40$ vergleichsweise groß ist. Daher werden diese Methoden als ungeeigneter eingestuft. Dies sind die vom Anfangswert unabhängige Variante 1 und das iterative GMRES-Verfahren für beide gewählten Vorkonditionierer.

Eine Erklärung für ein so schlechtes Ergebnis dieser Varianten kann in der Konditionierung der H -Matrix beziehungsweise der \hat{H} -Matrix der linearen Gleichungssysteme

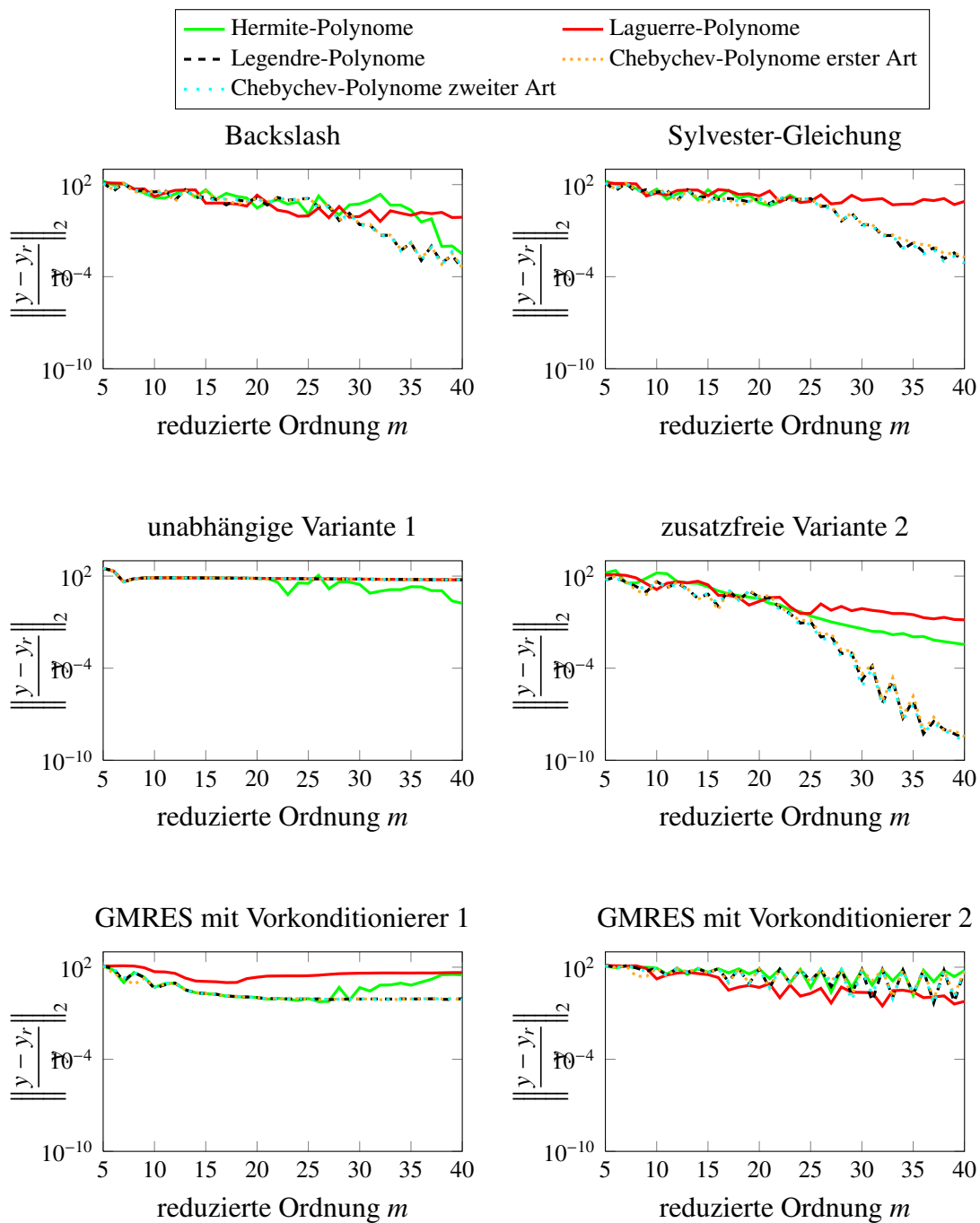
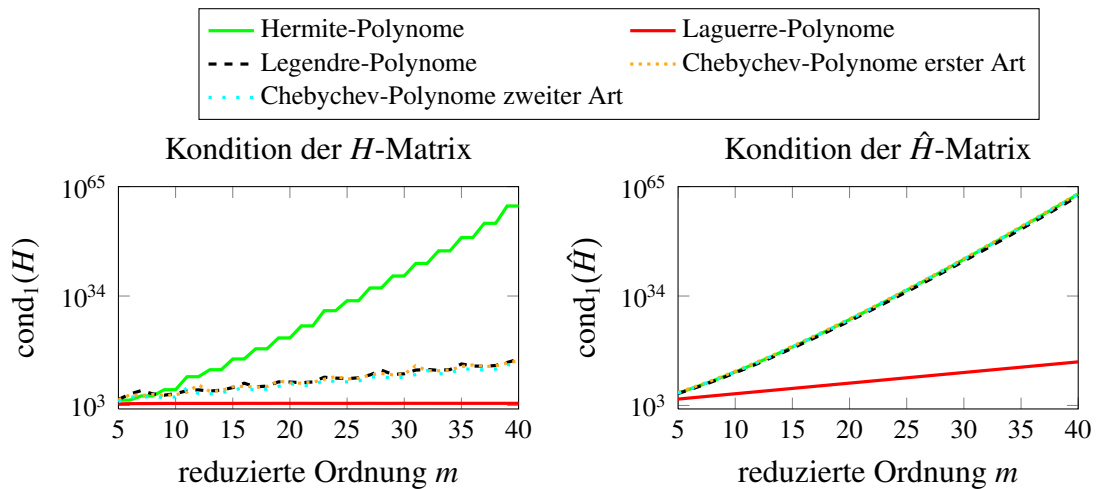


Abbildung 5.1: Relativer Fehler durch Reduktion mit Hilfe vorgestellter Verfahren (Beispiel 1)

Abbildung 5.2: Kondition der H - und \hat{H} -Matrix für Beispiel 1

(3.5) beziehungsweise (4.5) gefunden werden, die in Abbildung 5.2 für die Hermite-, Laguerre-, Legendre- und beide Chebyshev-Polynome dargestellt sind. Dabei wurde die Kondition mit Hilfe der Spaltensummennorm berechnet.

Die \hat{H} -Matrix hat für alle ausgewerteten orthogonalen Polynome eine so schlechte Kondition, dass das lineare Gleichungssystem nicht mehr eindeutig gelöst und daher das Originalmodell nicht gut approximiert wird.

Das GMRES-Verfahren setzt hingegen eine invertierbare Matrix H voraus, um eine Lösung gut approximieren zu können. Jedoch ist für alle Polynome, mit Ausnahme der Laguerre-Polynome, die H -Matrix schlecht konditioniert, sodass diese möglicherweise nicht mehr invertierbar ist und damit das Gleichungssystem nicht eindeutig gelöst werden kann. Beim Betrachten der Ergebnisse mit den Laguerre-Polynomen fällt auf, dass diese den Resultaten ähneln, die als Basis dasselbe oder ein Gleichungssystem mit äquivalenter Lösung besitzen, also das direkte Lösen des Gleichungssystems (3.5) oder das Lösen einer äquivalenten Sylvester-Gleichung (4.4). Allerdings sind diese relativen Fehler zu groß, um das Originalmodell gut zu approximieren.

Eine Besonderheit, die in der Grafik für das GMRES-Verfahren mit dem Kronecker-Produkt-Vorkonditionierer ins Auge sticht, sind die Zickzack-Verläufe des relativen Fehlers für alle ausgewerteten orthogonalen Polynome, mit Ausnahme der Laguerre-Polynome. Diese sind damit zu erklären, dass der Vorkonditionierer für ein gerades m nicht invertierbar ist und daher der relative Fehler für alle geraden m nahezu konstant ist. Diese Auffälligkeit ist ebenso im Verlauf der Ausführungszeit in Abbildung 5.4 sichtbar. Der Grund dafür ist, dass das GMRES-Verfahren für alle geraden m mit dem Startvektor $x_0 = [0 \ \dots \ 0]^T$ als Lösung abbricht und daher viel weniger Zeit benötigt wird. Für alle ungeraden m ist ein leicht fallender Verlauf des relativen Fehlers mit diesem

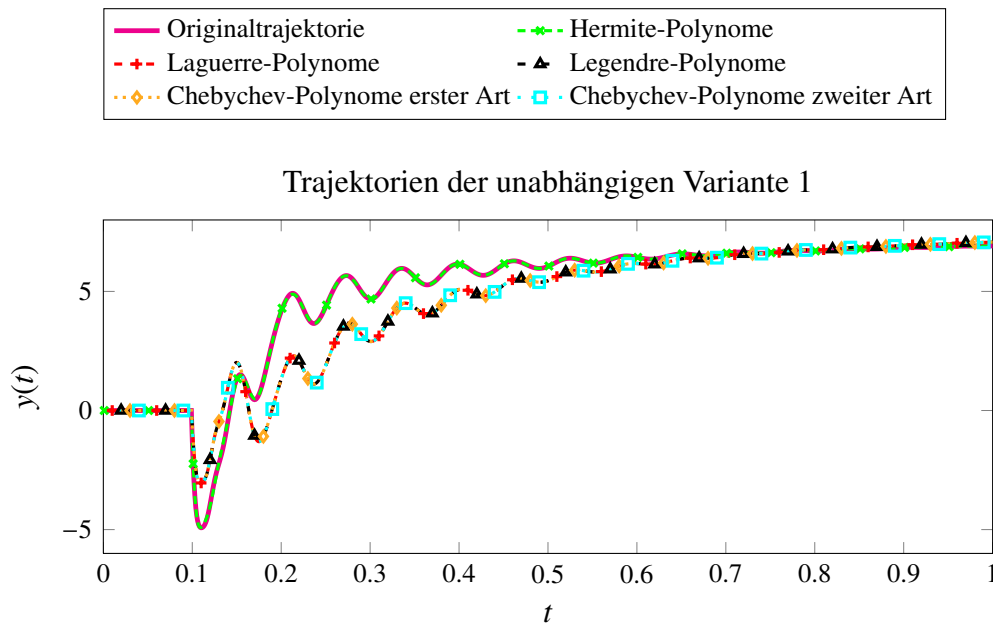


Abbildung 5.3: Trajektorien für die Anfangswert-unabhängige Variante 1 im Vergleich zur Originaltrajektorie für Beispiel 1 ($m = 40$)

vorkonditionierten GMRES-Verfahren zu erkennen.

Um exemplarisch zu sehen, wo die relativen Fehler bei diesen Verfahren auftreten, sind in Abbildung 5.3 die Lösungstrajektorien, die mit der Anfangswert-unabhängigen Variante 1 berechnet wurden, grafisch dargestellt. Die Kurven zu den reduzierten Ausgangsvektoren im Vergleich zur Originaltrajektorie für beide Varianten des GMRES-Verfahrens sind im Anhang in Abbildung 2 dargestellt.

Bei der Anfangswert-unabhängigen Variante 1 fällt auf, dass der Ausgangsvektor, der beim Reduzieren mit Hilfe der Hermite-Polynome berechnet wurde, keinen augenscheinlichen Fehler zur Originaltrajektorie aufweist und daher dieser Fehler nur in Abbildung 5.1 ersichtlich wird. Für die verbleibenden Polynome ist der Fehler zur Originaltrajektorie zu erkennen. Die Verläufe der reduzierten Ausgangsvektoren im Zeitintervall $t \in [0, 1]$ ähneln zwar denen des originalen Ausgangsvektor, jedoch erreichen sie nie dieselbe Amplitude, sondern liegen deutlich darunter.

Bei dem vorkonditionierten GMRES-Verfahren ist dieses Verhalten bei beiden Varianten genau umgekehrt. Dort erreicht die durch die Hermite-Polynome erzeugte Trajektorie nicht die gewünschte Amplitude und ist im Falle des ersten Vorkonditionierers sogar ab $t = 0.3$ nahezu konstant. Für die verbleibenden Polynome ist im Falle dieses Vorkonditionierers ein Fehler nur ab $t = 0.6$ erkennbar, da dort der reduzierte Ausgangsvektor knapp unter dem Originalen liegt. Für den zweiten Vorkonditionierer ist für diese Po-

lynome ein solcher Fehler für $m = 39$ nicht einmal erkennbar. Die reduzierte Ordnung $m = 39$ wurde für die Grafik des zweiten Vorkonditionierers so gewählt, da dieser für ein gerades m unbrauchbar ist.

Dennoch ist das GMRES-Verfahren für diese beiden Vorkonditionierer aufgrund des immensen Zeitaufwandes ungeeignet.

Bei den verbleibenden Grafiken der relativen Fehler wird wie erwartet bestätigt, dass das direkte Lösen des Gleichungssystems (3.5) und das Lösen der zu (3.5) äquivalenten Sylvester-Gleichung (4.4) einen ähnlichen Verlauf des relativen Fehlers besitzen, da die Lösung beider Gleichungen dieselbe ist. Die minimalen Abweichungen sind durch Rundungsfehler zu erklären, die vor allem beim direkten Lösen aufgrund der Größe der H -Matrix zu erwarten sind. Jedoch kann für ein $m > 24$ die äquivalente Sylvester-Gleichung mit Hilfe der Hermite-Polynome nicht mehr eindeutig gelöst werden, sodass diese relativen Fehler nicht mehr dargestellt sind. Dies ist erneut durch die schlecht konditionierte Matrix zu erklären, die sich aus der Sylvester-Gleichung ableiten lässt. Die Kondition ist ab einem $m > 24$ mindestens 10^{42} und die Matrix somit nicht mehr invertierbar.

Beim direkten Lösen mit Hilfe der Hermite-Polynome ist der relative Fehler für $m = 40$ zwar sehr klein, jedoch fällt er nicht wie der Fehler der Legendre- und beide Chebychev-Polynome, sondern wächst zwischenzeitlich an. Daher sind für diese beiden Verfahren die Legendre- und beide Chebychev-Polynome am geeignetsten, um das Originalmodell gut zu approximieren. Diese erreichen für $m = 40$ einen relativen Fehler der Größenordnung 10^{-4} .

Das beste Verfahren für dieses Beispiel ist jedoch die von der Zusatzbedingung unabhängige Variante 2. Für das größte gewählte m sind die relativen Fehler, die durch Verwendung der Hermite- und Laguerre-Polynome entstehen, vergleichbar mit denen des direkten Löser des linearen Gleichungssystems. Allerdings fallen die relativen Fehler im Gegensatz zum direkten Löser nahezu monoton für wachsendes m .

Am besten eignen sich jedoch die Legendre- und beide Chebychev-Polynome, die für $m = 40$ einen relativen Fehler der Größenordnung 10^{-9} haben.

In den letzten drei vorgestellten Verfahren sind erneut Zickzack-Verläufe bei den Legendre- und Chebychev-Polynomen erkennbar. Die Ursache dafür wird im Anschluss an die Auswertung der einzelnen Beispiele in Abschnitt 5.3 erläutert.

Beim Betrachten der Abbildung 5.4, die die gesamte Reduktions- und Lösungszeit verglichen zur Lösungszeit des Originalsystems darstellt, fällt zunächst positiv auf, dass die ausgewerteten Verfahren deutlich schneller sind als das Originalmodell zu lösen. Eine Ausnahme bildet das vorkonditionierte GMRES-Verfahren. Dieses Verfahren ist nur für den Vorkonditionierer, der sich aus einem Kronecker-Produkt ergibt, für ein gerades

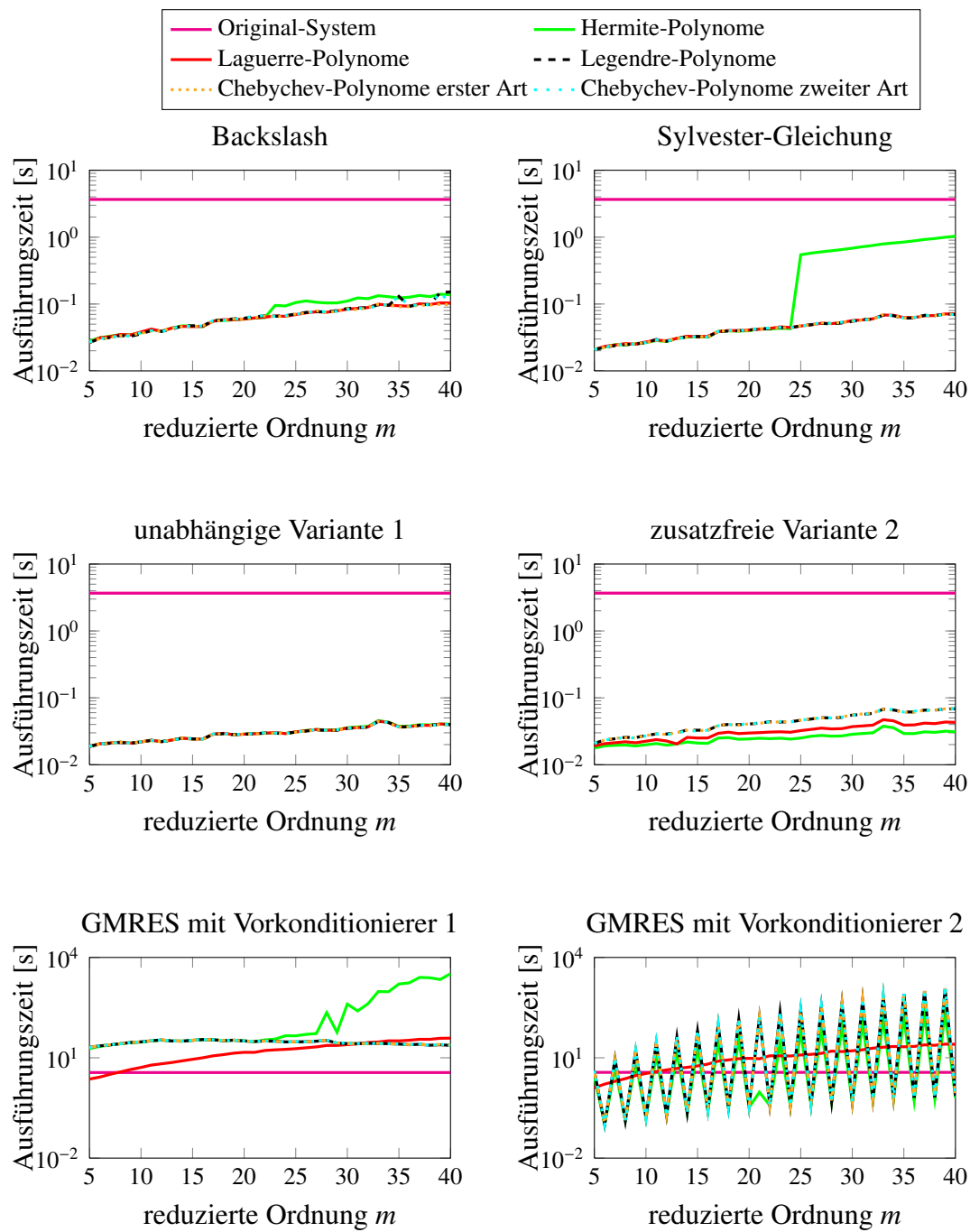


Abbildung 5.4: Vergleich der Gesamtzeiten durch Reduktion mit Hilfe vorgestellter Verfahren (Beispiel 1)

m schneller. Allerdings ist in diesem Fall, wie schon zuvor erörtert, der relative Fehler konstant hoch und daher auch diese Variante unbrauchbar.

Beim Lösen der Sylvester-Gleichung (4.4) zeigt sich, dass die Ausführungszeit für die Hermite-Polynome ab einem $m > 24$ deutlich länger dauert. Dies ist, wie zuvor beschrieben, auf die mangelnde Invertierbarkeit zurückzuführen. Ansonsten benötigt dieses Verfahren zum Lösen dieser Sylvester-Gleichung nahezu genauso lange wie das Lösen der vom Anfangswert unabhängigen Sylvester-Gleichung (4.7).

Bei der Bestimmung der Projektionsmatrix durch das Lösen der zusatzfreien Sylvester-Gleichung (4.12) wird für die Legendre- und beide Chebychev-Polynome genauso viel Zeit wie beim Lösen der anderen beiden Sylvester-Gleichungen benötigt. Auffällig ist bei dieser Variante, dass die Hermite- und Laguerre-Polynome wesentlich weniger Zeit gegenüber allen anderen vorgestellten Methoden benötigen.

Insgesamt ist die Methode des direkten Lösens zwar etwas langsamer als die restlichen direkten Lösungsstrategien, die eine Sylvester-Gleichung lösen, dennoch ist dieses Verfahren für die größte gewählte reduzierte Ordnung $m = 40$ 36-mal schneller als das Originalmodell zu lösen. Im Vergleich dazu ist das Lösen der äquivalenten Sylvester-Gleichung beziehungsweise der zusatzfreien Sylvester-Gleichung nochmals 1.5-mal schneller als das direkte Lösen, im Falle der vom Anfangswert unabhängigen Sylvester-Gleichung sogar 2.5-mal schneller.

Da an diesem ersten Beispiel deutlich wurde, dass nur das direkte Lösen, das Lösen der dazu äquivalenten Sylvester-Gleichung und der von der Zusatzbedingung unabhängigen Sylvester-Gleichung erfolgversprechend für eine Modellreduktion sind, werden nur diese Verfahren für die kommenden Beispiele betrachtet.

Da das direkte Lösen ähnliche relative und absolute Fehler wie das Lösen der dazu äquivalenten Sylvester-Gleichung liefert, wird für die kommenden Beispiele auf diese Grafiken ebenfalls im Hauptteil dieser Arbeit verzichtet. Die einzigen signifikanten Unterschiede gibt es in diesen Grafiken nur im Falle der Hermite-Polynome, da vor allem beim direkten Lösen des linearen Gleichungssystems (3.5) Rundungsfehlern aufgrund der sehr schlecht konditionierten Matrix H auftreten. Die weiteren Verläufe der absoluten und relativen Fehler für die verbleibenden Polynome sind nahezu identisch. Die zugehörigen grafischen Darstellungen der absoluten und relativen Fehler sind daher im Anhang B dieser Arbeit zu finden.

5.2.2 Beispiel 2

Das zweite Beispiel handelt von dem Problem des optimalen Kühlens eines unendlich langen Stahlprofils.

Dabei wird ein verallgemeinertes Zustandsraummodell aus einer halbdiskretisierten Wärmeleitungsgleichung aufgebaut. Eine genaue Beschreibung dieses Modells ist in [16]

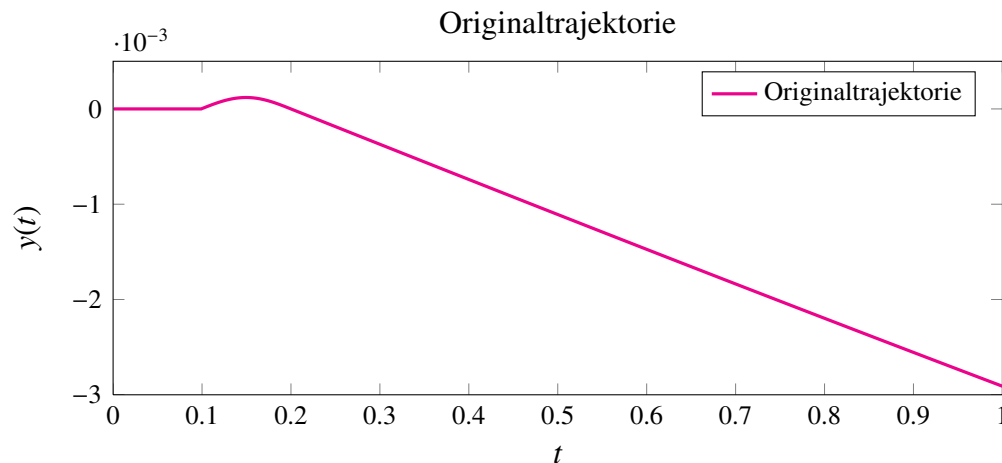


Abbildung 5.5: Originaltrajektorie für Beispiel 2

enthalten.

Das Zustandsraummodell, das in dieser Arbeit betrachtet wurde, hat die Dimension 1357 und ist als MIMO-System gegeben. Um ein SISO-System zu erhalten werden die Zeilen der B -Matrix beziehungsweise die Spalten der C -Matrix aufsummiert, in MATLAB[®]-Notation

$$B = \text{sum}(B, 2),$$

$$C = \text{sum}(C).$$

In diesem Beispiel ist E keine Identitätsmatrix.

Die Kurve des Ausgangsvektors, die durch Lösen des originalen Modells in der Zeit $t \in [0, 1]$ gegeben ist, wird in Abbildung 5.5 dargestellt.

In Abbildung 5.6, die den relativen Fehler des Ausgangsvektor für dieses Beispiel darstellt, fällt zunächst auf, dass für beide Lösungsstrategien nur eine kleine Ordnung m zum Reduzieren benötigt wird, um eine sehr gute Approximation an $y(t)$ im Zeitraum $t \in [0, 1]$ zu erzielen.

Die relativen Fehler sind für beide Verfahren besonders im Falle der Legendre- und bei der Chebyshev-Polynome sehr klein und liegen für $m = 20$ in der Größenordnung von 10^{-12} .

Auch für die Hermite-Polynome werden sehr gute Annäherungen an die Originaltrajektorie erzielt. Der relative Fehler hat für das Lösen mit der Sylvester-Gleichung eine Ordnung von 10^{-7} , im Falle des Lösen mit der zusatzfreien Sylvester-Gleichung sogar eine Ordnung von 10^{-12} für das größte gewählte m . Besonders auffällig ist dabei, dass dieser durch die zusatzfreie Variante 2 erreichte relative Fehler sogar schon ab $m = 13$ für die Hermite-Polynome sehr gering ist und auch die Legendre- beziehungs-

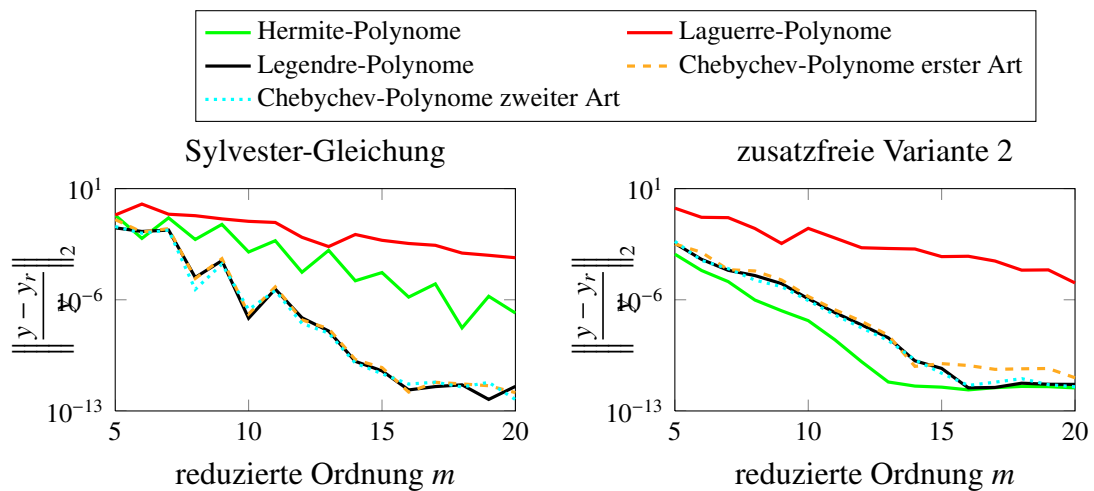


Abbildung 5.6: Relativer Fehler durch die verbleibenden Lösungsstrategien (Beispiel 2)

weise beide Chebychev-Polynome ab $m = 16$ einen nahezu konstanten relativen Fehler der Größenordnung 10^{-11} besitzen.

Nur die Laguerre-Polynome weisen wie schon in Beispiel 1 in beiden Fällen höhere relative Fehler auf, die sich im Bereich von 10^{-4} für das Lösen der äquivalent umgeformten Sylvester-Gleichung beziehungsweise 10^{-5} für die zusatzfreie Variante 2 befinden.

Eine Auffälligkeit in Abbildung 5.6, die in der Grafik der Sylvester-Gleichung auftritt, ist ein Zick-Zack-Verlauf des relativen Fehler für die Hermite-, Legendre- und beide Chebychev-Polynome. Eine Erklärung für dieses Phänomenen wird in Abschnitt 5.3 angegeben.

Die Zeit, die zum Reduzieren und Lösen des reduzierten Modells benötigt wird, ist in Abbildung 5.7 dargestellt und liegt wie in Beispiel 1 für beide auf Sylvester-Gleichungen basierende Varianten deutlich unter der Zeit, die zum Lösen des Originalmodells gebraucht wird.

Dabei fällt besonders für die zusatzfreie Variante 2 auf, dass das reduzierte System mit den Hermite- und Laguerre-Polynomen mindestens doppelt so schnell aufgestellt und gelöst wird wie mit den verbleibenden Polynomen.

Für beide dargestellten Varianten ist das Lösen mit den Legendre- und beiden Chebychev-Polynomen ungefähr genauso schnell, jedoch verglichen zum direkten Lösen des Gleichungssystem (3.5) mit denselben Polynomen bis zu 5-mal schneller. Besonders unter Verwendung der Hermite-Polynome wird für das direkte Lösen besonders viel Zeit benötigt, die sogar über die Lösungszeit des Originalmodells hinausgeht. Im Falle dieser orthogonalen Polynome ist das Verfahren, das eine äquivalente Sylvester-Gleichung

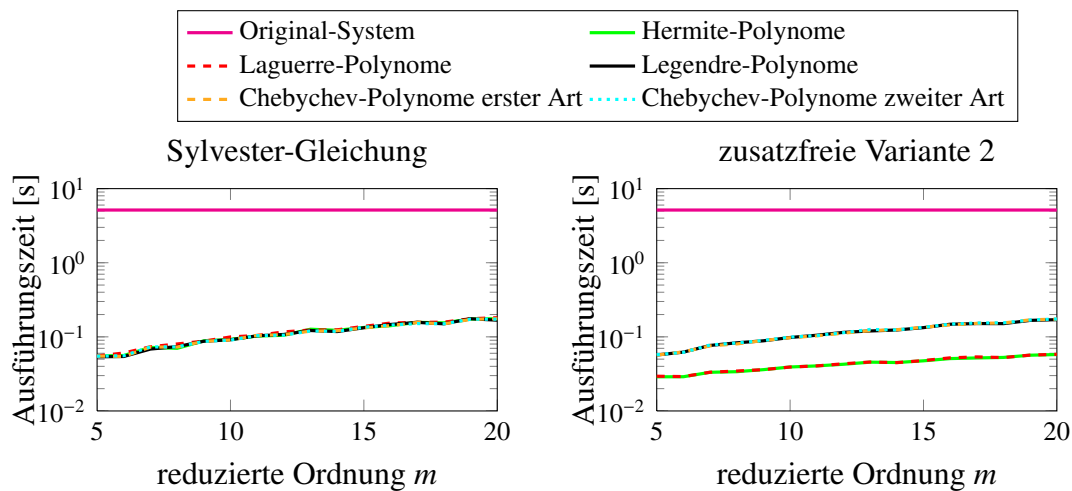


Abbildung 5.7: Vergleich der Gesamtzeit verbleibender Lösungsstrategien (Beispiel 2)

löst, für eine reduzierte Ordnung von $m = 20$ sogar 22-mal schneller.

Eine gegenüberstellende Grafik der benötigten Reduktions- und Lösungszeiten für eine Modellreduktion durch das direkte Lösen eines linearen Gleichungssystems und dem Lösen der dazu äquivalenten Sylvester-Gleichung ist im Anhang in Abbildung 5 zu finden.

Verglichen mit der Lösungszeit des originalen Modells ist das Reduzieren und das damit verbundene Lösen des reduzierten Modells der Dimension 20 für die Legendre- und beide Chebyshev-Polynome für beide in Abbildung 5.7 dargestellten Varianten 30-mal schneller. Eine noch größere Zeitersparnis gibt es im Falle der Hermite- und Laguerre-Polynome für die zusatzfreie Variante 2. Mit Hilfe dieser Lösungsstrategie kann eine Reduktion und Simulation sogar 89-mal schneller durchgeführt werden.

Mit dem Wissen, dass eine sehr gute Approximation mit Hilfe der zusatzfreien Variante 2 schon ab $m = 13$ für die Hermite-Polynome beziehungsweise ab $m = 16$ für die Legendre- und beide Chebyshev-Polynome erzielt werden kann, ist diese Variante der Modellreduktion sogar 113-mal beziehungsweise 35-mal schneller gegenüber dem Lösen des originalen Modells.

5.2.3 Beispiel 3

Das dritte Beispiel beschreibt eine diskretisierte Wärmeleitungsgleichung auf dem Einheitsquadrat $\Omega = (0, 1) \times (0, 1)$

$$\dot{u} \approx \Delta u - f_1(x, y) \frac{\partial}{\partial x} u - f_2(x, y) \frac{\partial}{\partial y} u + b,$$

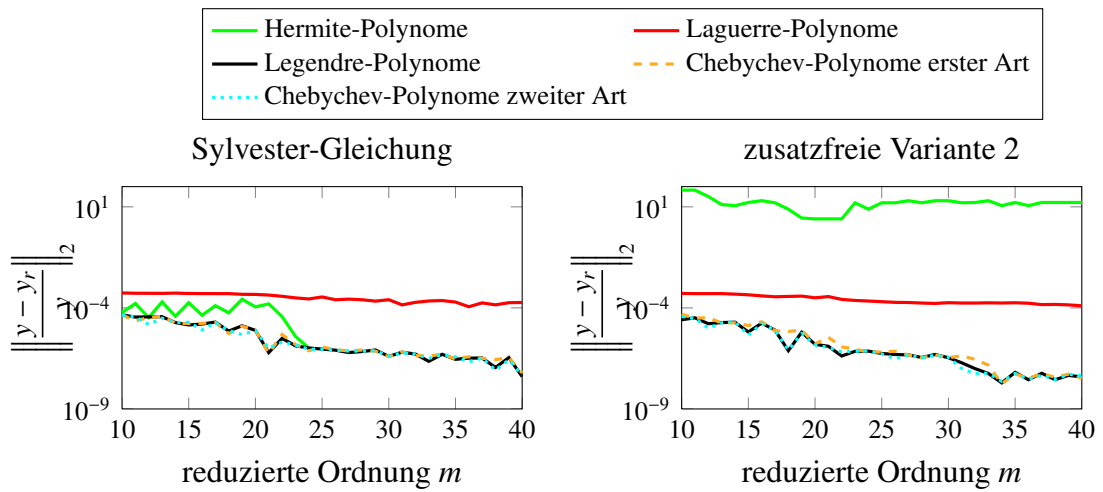


Abbildung 5.8: Relativer Fehler durch die verbleibenden Lösungsstrategien (Beispiel 3)

die ausführlicher in [12] beschrieben ist.

Hierbei ist der Vektor $f = [f_1 \ f_2]^T$ die Konvektion, die in dieser Arbeit als $f = [60 \ 90]^T$ gewählt ist, um eine unsymmetrische Matrix A zu erzeugen.

Die Dimension des daraus resultierenden Zustandsraummodells ist 900, wobei $B^T = C = [1 \ \dots \ 1] \in \mathbb{R}^{1 \times 900}$ und $E = I_{900}$ die Einheitsmatrix ist.

In der Abbildung 5.8 der relativen Fehler für dieses Beispiel sind bei beiden Verfahren ähnliche Verläufe für die Laguerre-, Legendre- und beide Chebychev-Polynome zu erkennen. Dabei sind die relativen Fehler der Laguerre-Polynome in beiden dargestellten Varianten nahezu konstant und liegen in einer Größenordnung von 10^{-4} .

Für die Legendre- und beide Chebychev-Polynome ist für eine wachsende reduzierte Ordnung m eine Verbesserung des relativen Fehlers ersichtlich, welcher bei beiden Lösungsstrategien für $m = 40$ eine Größenordnung von 10^{-8} hat.

Jedoch gibt es bei den Hermite-Polynomen große Unterschiede. Bei der Lösung der aus dem linearen Gleichungssystem (3.5) entwickelten Sylvester-Gleichung, können die Hermite-Polynome nur bis $m = 24$ bewertet werden, da die Sylvester-Gleichung für ein größeres m nicht mehr eindeutig gelöst werden kann. In diesem Falle liegt der relative Fehler bei 10^{-6} . Wird die im Anhang B befindliche Abbildung 7 hinzugezogen, bei der die relativen Fehler der beiden äquivalenten Lösungsstrategien verglichen werden, so wird deutlich, dass die Hermite-Polynome auch für $m > 24$ den Ausgangsvektor sehr gut approximieren. Dies ist im Fall der zusatzfreien Variante 2 nicht der Fall. Hier bleibt der relative Fehler auch für wachsende Werte von m nahezu konstant hoch.

Um zu zeigen, wieso der relative Fehler für die Hermite-Polynome bei dieser Variante

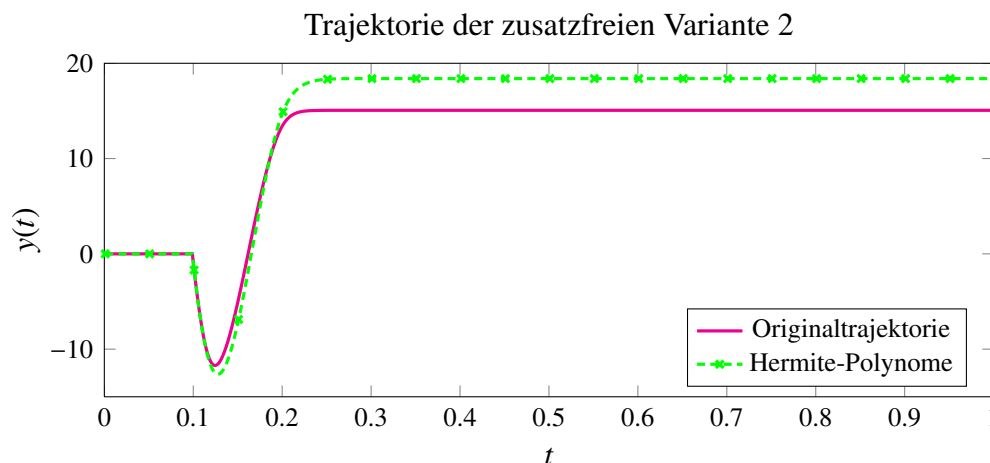


Abbildung 5.9: Trajektorie für die zusatzfreie Variante 2 (Hermite-Polynome) im Vergleich zur Originaltrajektorie für Beispiel 3 ($m = 40$)

so hoch ist, ist die Kurve des reduzierten Ausgangsvektors $y(t)$ der Originaltrajektorie in Abbildung 5.9 für einen Zeitraum $t \in [0, 1]$ gegenübergestellt.

Auf die Darstellung der restlichen orthogonalen Polynome wird verzichtet, da selbst für die Laguerre-Polynome kein Fehler mit bloßem Auge zu erkennen ist.

In Abbildung 5.9 ist ersichtlich, dass die Trajektorie der Hermite-Polynome zwar einen ähnlichen Verlauf wie die Originaltrajektorie aufweist, die Amplitude jedoch stärker ist und daher der große relative Fehler entsteht.

In beiden Grafiken in Abbildung 5.8 wird wieder ein Zick-Zack-Verlauf für die Legendre- und beide Chebyshev-Polynome deutlich. Dieser ist auch für die Hermite-Polynome in der Darstellung des relativen Fehlers für die äquivalente Sylvester-Gleichung erkennbar und wird im Anschluss an die Beispiele erläutert.

Beim Betrachten der Abbildung 5.10 fällt, wie schon in Beispiel 2, auf, dass bei beiden dargestellten Varianten der Modellreduktion unter Verwendung der Legendre- und beider Chebyshev-Polynome die Ausführungszeiten nahezu identisch sind. Die einzigen Unterschiede sind bei den Hermite- und Laguerre-Polynomen erkennbar, die für $m = 40$ in der zusatzfreien Variante 32-mal beziehungsweise 1.5-mal schneller sind im Vergleich zu den Legendre- und Chebyshev-Polynomen.

Eine riesige Ersparnis in der Zeit wird vor allem aber in der im Anhang B befindlichen Abbildung 8 deutlich, in der die benötigte Rechenzeit der Reduktionsmethoden mit Hilfe des direkten Lösens und dem Lösen der dazu äquivalenten Sylvester-Gleichung gegenübergestellt wird. Die Methode des direkten Lösens mit Hilfe der Legendre- und beider Chebyshev-Polynome benötigt dabei sogar mehr Zeit als das Lösen des origi-

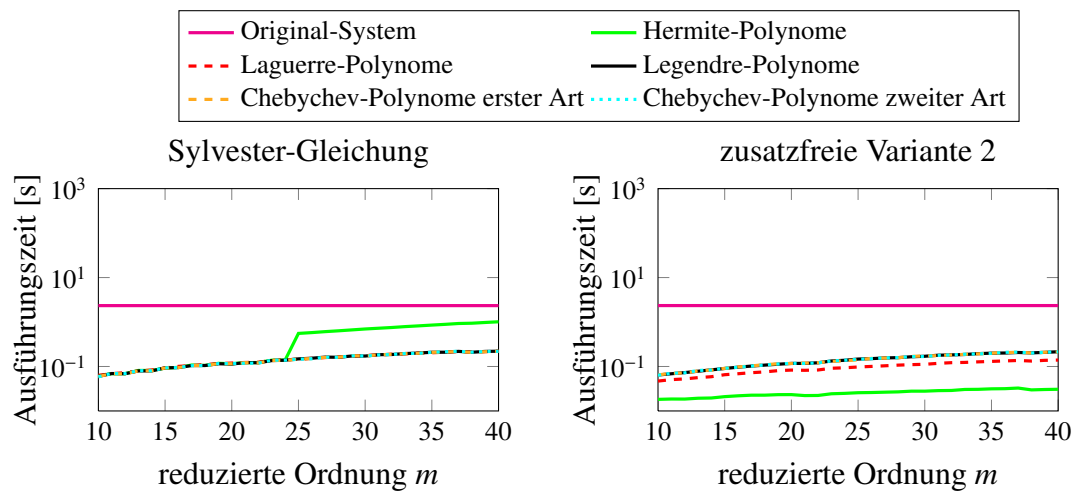


Abbildung 5.10: Vergleich der Gesamtzeit verbleibender Lösungsstrategien (Beispiel 3)

nen Modells, genau so wie bei Verwendung der Hermite-Polynome ab $m > 24$. Nur für die Laguerre-Polynome liegt die gesamte Rechenzeit bei diesem Verfahren unterhalb der Lösungszeit des originalen Modells. Daher ist das Lösen mit der äquivalenten Sylvester-Gleichung für $m = 40$ für die Legendre- und beide Chebychev-Polynome zwischen 146- und 1400-mal, für die Hermite-Polynome 7-mal und für die Laguerre-Polynome 11-mal schneller.

Insgesamt sind beide in Abbildung 5.10 dargestellten Methoden deutlich schneller als das Lösen des originalen Modells. Dabei sind die Legendre- und beide Chebychev-Polynome, die den kleinsten relativen Fehler erzielen, für eine reduzierte Ordnung $m = 40$ ungefähr 10-mal schneller.

5.2.4 Beispiel 4

Dieses Beispiel stellt ein proportional gedämpftes Schwingungssystem mit 500 Massen dar, dessen Zustandsraummodell als ein Modell zweiter Ordnung

$$\begin{aligned} M\ddot{x}(t) + D\dot{x}(t) + Kx(t) &= Bu(t) \\ y(t) &= Cx(t) \end{aligned} \quad (5.1)$$

gegeben ist und in ein LTI-System

$$\begin{aligned} \begin{bmatrix} I & 0 \\ 0 & M \end{bmatrix} \begin{bmatrix} \dot{x}(t) \\ \ddot{x}(t) \end{bmatrix} &= \begin{bmatrix} 0 & I \\ -K & -D \end{bmatrix} \begin{bmatrix} x(t) \\ \dot{x}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ B \end{bmatrix} u(t) \\ y(t) &= \begin{bmatrix} C & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ \dot{x}(t) \end{bmatrix} \end{aligned} \quad (5.2)$$

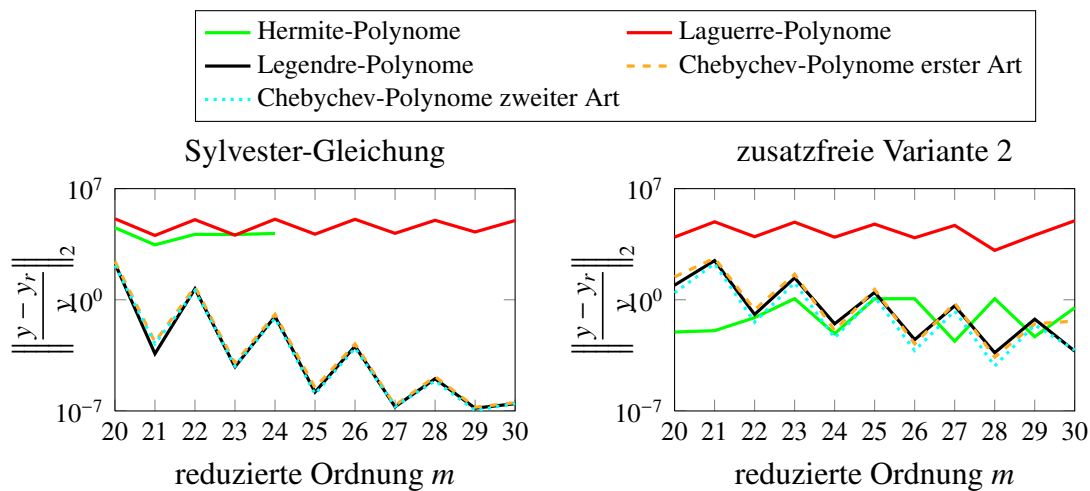


Abbildung 5.11: Relativer Fehler durch die verbleibenden Lösungsstrategien (Beispiel 4)

transformiert wird. Die Ordnung dieses Systems ist 1000.

Die Matrizen M , D und K werden als Masse-, Dämpfungs- beziehungsweise Steifigkeitsmatrix bezeichnet.

Das Besondere an diesem Modell ist, dass die Matrizen M und D mit zufälligen Werten gefüllt sind.

M und D sind dabei Diagonalmatrizen und K eine Tridiagonalmatrix. Auch in diesem Beispiel ist $B^T = C = \begin{bmatrix} 1 & \dots & 1 \end{bmatrix} \in \mathbb{R}^{1 \times 500}$.

In der Abbildung 5.11 des relativen Fehlers fällt als Erstes in beiden Grafiken der regelmäßige Zick-Zack-Verlauf von den Laguerre-, Legendre- und beiden Chebychev-Polynomen auf. Die Hermite-Polynome zeigen hingegen ein unregelmäßiges Muster. Wie bei den anderen Beispielen schon erwähnt, wird dies im Abschnitt 5.3 erklärt.

In Abbildung 5.11 bestätigt sich in beiden Grafiken, dass die Laguerre-Polynome nicht geeignet sind, um den Ausgangsvektor im gewünschten Zeitraum gut zu approximieren. Der relative Fehler ändert sich jeweils für gerades beziehungsweise ungerades m kaum und liegt für $m = 30$ bei 10^5 .

In ähnlicher Größenordnung liegt bei der Bestimmung der Projektionsmatrix durch die äquivalente Sylvester-Gleichung mit Hilfe der Hermite-Polynome der relative Fehler, der in Abbildung 5.11 nur bis $m = 24$ dargestellt ist. Im Anhang wird in Abbildung 10 jedoch deutlich, dass auch diese orthogonalen Polynome die Originaltrajektorie nicht gut approximieren und der Fehler für das größte gewählte m leicht über dem der Laguerre-Polynome liegt.

Für dieselben orthogonalen Polynome ist jedoch der relative Fehler im Falle der von der

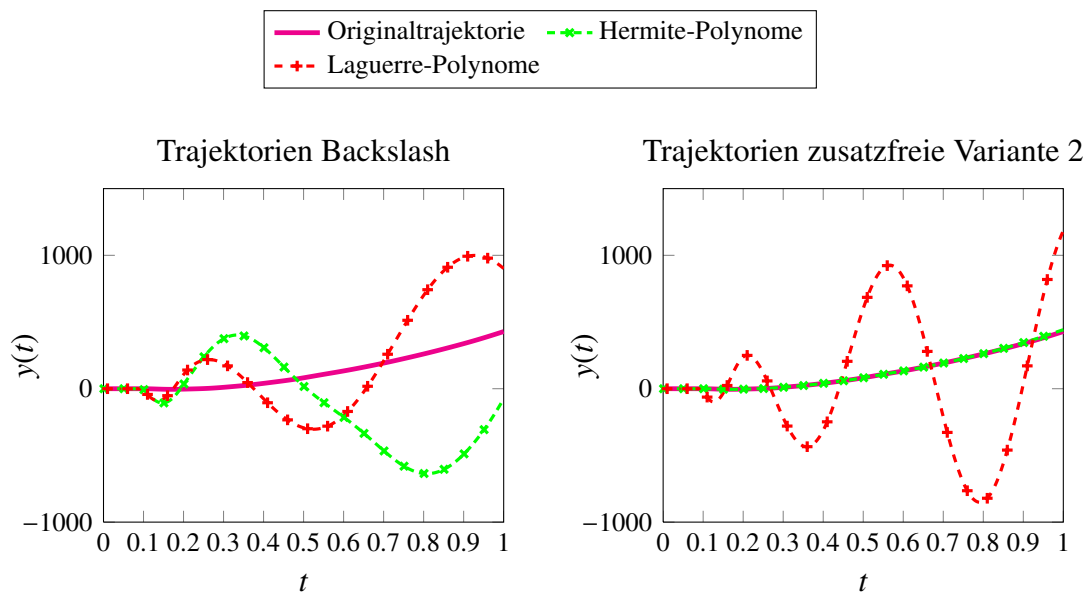


Abbildung 5.12: Trajektorien für Backslash und zusatzfreie Variante 2 (Hermite-, Laguerre-Polynome) im Vergleich zur Originaltrajektorie für Beispiel 4 ($m = 30$)

Zusatzbedingung unabhängigen Variante 2 viel kleiner und liegt für $m = 30$ bei 10^{-1} und beträgt bei einer kleineren reduzierten Ordnung $m = 29$ sogar nur 10^{-3} .

Am besten schneiden aber in beiden der in Abbildung 5.11 dargestellten Varianten die Legendre- und beide Chebyshev-Polynome ab. Für diese orthogonalen Polynome liegt der relative Fehler für $m = 30$ im Falle des Lösens der zu (3.5) äquivalenten Sylvester-Gleichung bei 10^{-7} beziehungsweise bei der zusatzfreien Variante 2 bei 10^{-4} .

In Abbildung 5.12 sind die Kurven der Ausgangsvektoren für das Reduzieren durch direktes Lösen des linearen Gleichungssystems (3.5) und der zusatzfreien Sylvester-Gleichung (4.12) im Zeitintervall $t \in [0, 1]$ für eine reduzierte Ordnung $m = 30$ der Originaltrajektorie gegenübergestellt. Hierbei wurden die Trajektorien des reduzierten Systems, dessen Projektionsmatrix durch das direkte Lösen erzeugt wurde, dargestellt. In diesem Fall kann auch der Fehler, der durch die Hermite-Polynome verursacht wurde, bewertet werden. Dies ist damit zu begründen, dass die Lösungen, die durch das direkte Lösen oder durch die umgeformte Sylvester-Gleichung berechnet werden, für die verbleibenden orthogonalen Polynome äquivalent sind, was auch im Vergleich der relativen Fehler im Anhang in Abbildung 10 ersichtlich ist.

In Abbildung 5.12 ist erkennbar, dass die Kurven der reduzierten Ausgangsvektoren im Falle des direkten Lösens stark von der Originaltrajektorie abweichen. Während die Kurve des originalen Ausgangsvektors monoton wächst, schwingen die Trajektorien der dargestellten orthogonalen Polynome.

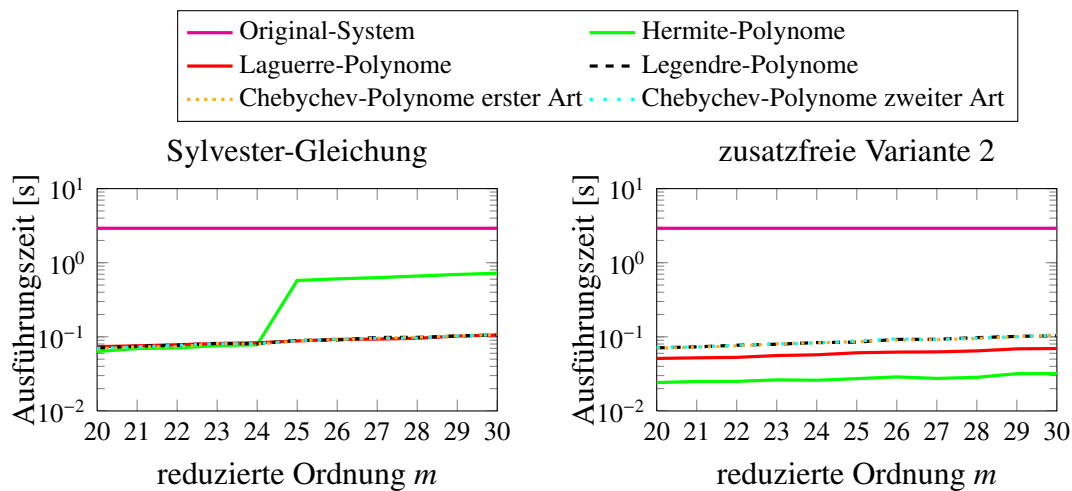


Abbildung 5.13: Vergleich der Gesamtzeit verbleibender Lösungsstrategien (Beispiel 4)

Ein ähnlicher Verlauf ist ebenso bei den Laguerre-Polynomen für die zusatzfreie Variante 2 zu erkennen. Nur die Trajektorie der Hermite-Polynome verläuft nahezu auf der Originaltrajektorie, weicht aber ab $t = 0.9$ leicht davon ab.

In Abbildung 5.13 liegt, wie auch in den Beispielen zuvor, die Zeit, die zum Reduzieren des originalen Modells und zum Lösen dieses Reduzierten benötigt wird, deutlich unter der Lösungszeit des Originalmodells.

Dabei sind in beiden dargestellten Varianten der Modellreduktion, die die Legendre- und beide Chebychev-Polynome nutzen, die Ausführungszeiten nahezu identisch.

Ebenso ist wie in den anderen Beispielen die zusatzfreie Variante 2 mit den Hermite- beziehungsweise Laguerre-Polynomen schneller als die konkurrierende Methode. In diesem Falle ist für $m = 30$ das Lösen mit Hilfe der Hermite-Polynome 22-mal und durch die Laguerre-Polynome ungefähr 1.5-mal schneller verglichen zur Reduktionsmethode, die die äquivalente Sylvester-Gleichung löst.

Eine deutliche Zeitersparnis durch diese beiden in Abbildung 5.13 dargestellten Varianten ist im Vergleich zur Methode des direkten Lösens im Anhang B in Abbildung 11 zu erkennen.

Hierbei dauert das Reduzieren und Lösen unter Verwendung der Legendre- und beider Chebychev-Polynome ab $m = 22$ mindestens genauso lange wie das Lösen des Originalmodells. Verglichen zu der äquivalenten Variante, die eine Sylvester-Gleichung löst, ist diese Methode für $m = 30$ sogar 66-mal langsamer.

Nur für die Hermite- und Laguerre-Polynome ergibt sich eine Zeitersparnis. Allerdings sind diese orthogonalen Polynome wie zuvor erörtert für dieses Beispiel nicht zur Reduktion geeignet, was in der linken Grafik in Abbildung 5.12 erkennbar ist.

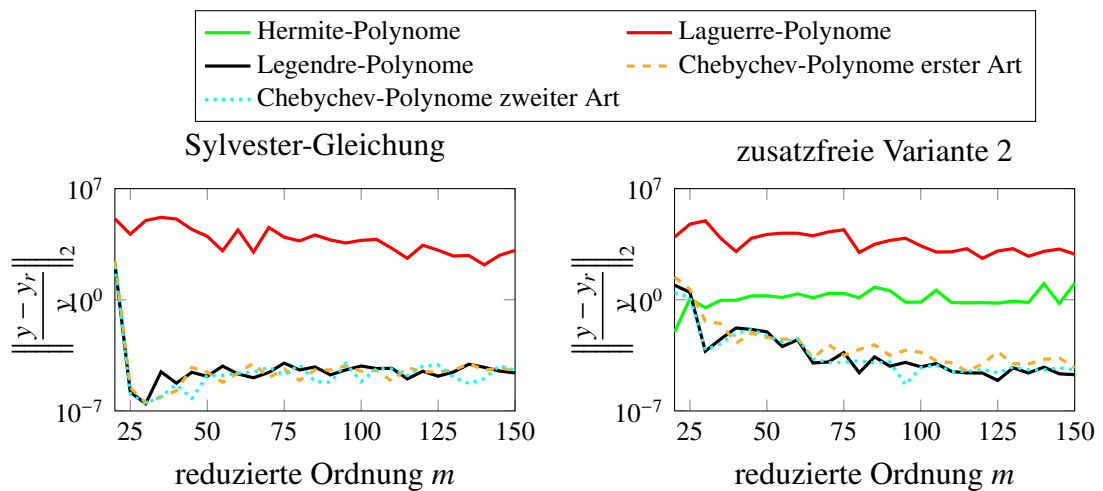


Abbildung 5.14: Relativer Fehler für $m = 20, \dots, 150$ durch die verbleibenden Lösungsstrategien (Beispiel 4)

Zusammenfassend sind auch in diesem Beispiel die in Abbildung 5.13 dargestellten Reduktionsvarianten für alle orthogonalen Polynome wesentlich schneller gegenüber dem Lösen des Originalmodells.

Dabei sind beide Varianten, die die Legendre- und beide Chebyshev-Polynome verwenden, für $m = 30$ ungefähr 27-mal schneller und liefern somit eine schnelle und sehr gute Approximation des Originalmodells.

Die ebenso gut approximierende zusatzfreie Variante 2, die die Hermite-Polynome verwendet, ist für das größte gewählte m sogar 91-mal schneller.

Für dieses Beispiel gibt es eine weitere Abbildung 5.14, die zeigt, dass es nicht unbedingt notwendig ist, die reduzierte Ordnung m weiter zu vergrößern.

In Abbildung 5.14 ist der relative Fehler für größere Werte von m für das Beispiel 4 dargestellt. Für diese Abbildung wurde der relative Fehler jedoch nur alle 5 Schritte berechnet, um eine Tendenz aufzuzeigen.

Vor allem für das Reduktionsverfahren, das die Projektionsmatrix durch die äquivalente Sylvester-Gleichung berechnet, wird deutlich, dass außer im Falle der Laguerre-Polynome eine größere Wahl der reduzierten Ordnung m den relativen Fehler nicht verbessert.

Anders sieht dies jedoch bei der zusatzfreien Variante 2 aus. In der dazu gehörigen Grafik ist ersichtlich, dass bei einer Wahl der Legendre- oder Chebyshev-Polynome eine kontinuierliche Verbesserung des relativen Fehlers auftritt. Auch bei der Verwendung der Laguerre-Polynome wird eine leichte Verbesserung des relativen Fehlers mit

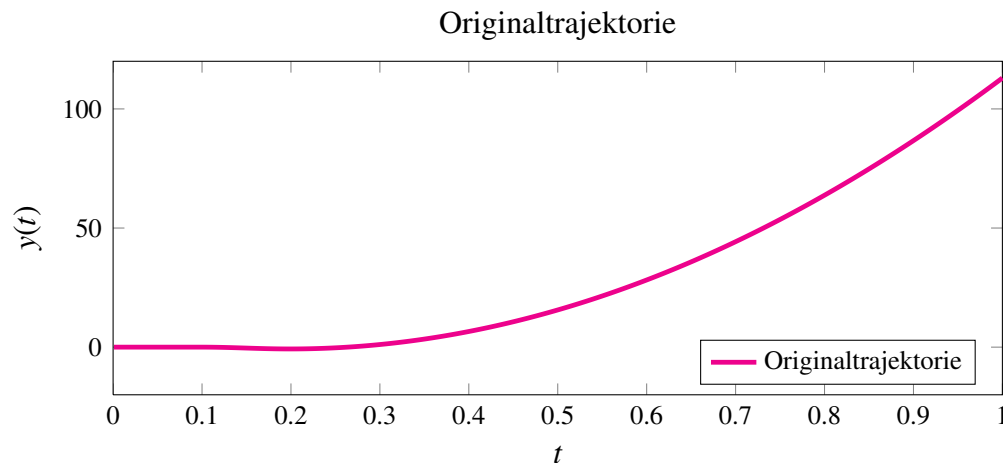


Abbildung 5.15: Originaltrajektorie für Beispiel 5

wachsendem m ersichtlich. Dieser Trend ist bei den Hermite-Polynomen leider nicht zu erkennen. Eine Erklärung dafür ist im Anschluss zu finden.

Im Anhang B dieser Arbeit sind für das vierte Beispiels in den Abbildungen 9,10 und 11 die absoluten Fehler und die vergleichenden Darstellungen des relativen Fehlers beziehungsweise der Ausführungszeit zu finden. Ebenso sind die zu Abbildung 5.14 korrespondierende Grafiken zu dem absoluten Fehler in Abbildung 12 und der Ausführungszeit in Abbildung 13 dargestellt.

In der zuletzt genannten Grafik wird deutlich, dass beide Modellreduktionsvarianten auch bei $m = 150$ schneller sind im Vergleich zur Lösung des Originalmodells. Eine Ausnahme bilden hier die Hermite-Polynome beim Lösen der Sylvester-Gleichung (4.4). Diese erwiesen sich zuvor schon als ungeeignet.

5.2.5 Beispiel 5

Dieses Beispiel beschreibt eine Koppelmasse, an der drei Ketten von Masse-Feder-Dämpfern angebracht sind, die proportional gedämpft werden.

Ein Zustandsraummodell ist wie in Beispiel 4 durch ein System zweiter Ordnung (5.1) gegeben, das in ein LTI-System (5.2) umgeformt wird.

Die Dimension des Systems ist 1202. Jede Kette besteht aus 200 Massen, die mit der Koppelmasse verbunden sind.

Die Matrix M ist eine Diagonalmatrix und erneut sind $B^T = C = \begin{bmatrix} 1 & \dots & 1 \end{bmatrix} \in \mathbb{R}^{1 \times 601}$.

Die Kurve des Ausgangsvektors ist für dieses Modells in Abbildung 5.15 über einen Zeitraum $t \in [0, 1]$ dargestellt.

In Abbildung 5.16 sind die relativen Fehler der Ausgangsvektoren $y(t)$ in einem Inter-

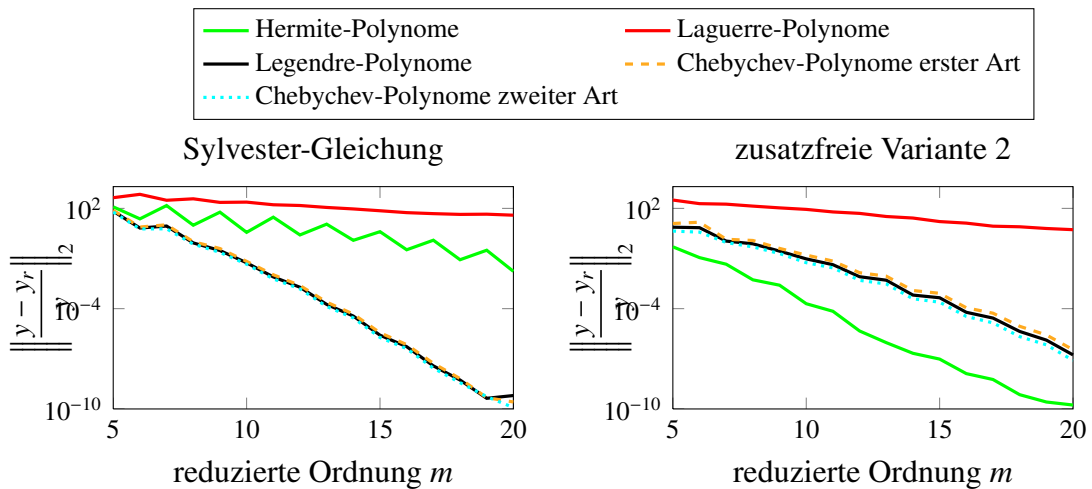


Abbildung 5.16: Relativer Fehler durch die verbleibenden Lösungsstrategien (Beispiel 5)

vall $t \in [0, 1]$ für die Modellreduktionsvarianten, die die Projektionsmatrix durch Lösen einer zu (3.5) äquivalenten beziehungsweise zusatzfreien Sylvester-Gleichung bestimmen, dargestellt.

Auch in dieser grafischen Darstellung ist erkennbar, dass die Verwendung von Laguerre-Polynomen in beiden Methoden nicht erfolgversprechend ist, da diese einen relativen Fehler der Größenordnung 10^1 aufweisen.

Anders ist dies im Falle der Legendre- und beider Chebyshev-Polynome. Wie in den Beispielen zuvor weichen die Trajektorien für $m = 20$ kaum von der Originaltrajektorie ab. Beim Lösen der Sylvester-Gleichung (4.4) liegt der relative Fehler bei 10^{-10} beziehungsweise bei 10^{-8} beim Lösen der zusatzfreien Sylvester-Gleichung (4.12).

Ähnlich wie in den Beispielen 1,2 und 4 weisen die Reduktionsmethoden unter Verwendung der Hermite-Polynome gegensätzliche Resultate für den relativen Fehler auf. Während in der linken Grafik von Abbildung 5.16 mit den Hermite-Polynomen für $m = 20$ nur ein relativer Fehler der Größenordnung 10^{-2} erreicht wird, liegt dieser Fehler in der rechten Grafik von Abbildung 5.16 bei 10^{-10} . Auffällig ist in der linken grafischen Darstellung des relativen Fehlers, dass bei diesen orthogonalen Polynomen ein Zick-Zack-Muster auftritt. Dieses Phänomen wird im nachfolgenden Abschnitt 5.3 näher untersucht.

Die Abbildung 5.17 bestätigt nur, was in den vorhergehenden Beispielen beobachtet werden konnte.

Beide dargestellten Reduktionsmethoden sind wesentlich schneller als das Lösen des Originalmodells. Dabei sind für beide Verfahren die Reduktion mittels der Legendre- und beider Chebyshev-Polynome nahezu gleich schnell.

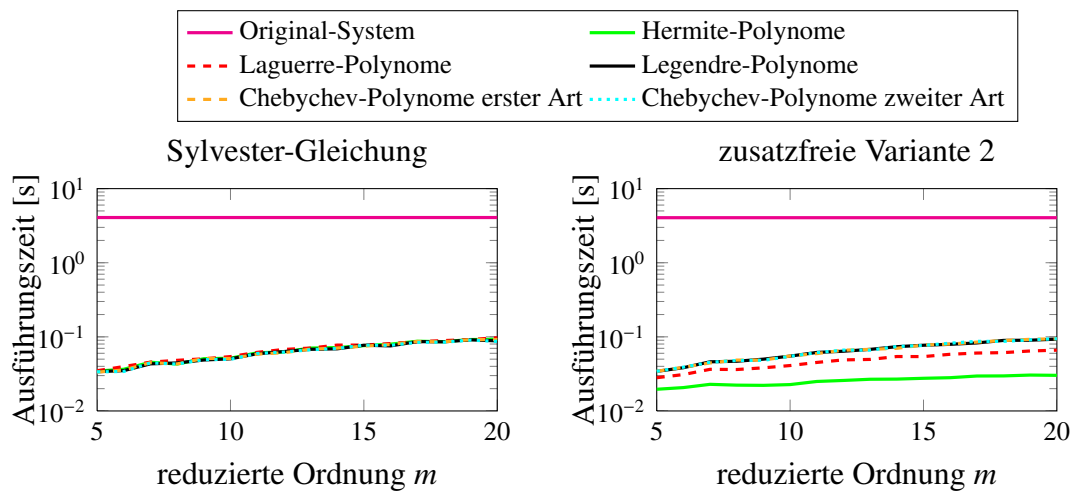


Abbildung 5.17: Vergleich der Gesamtzeit verbleibender Lösungsstrategien (Beispiel 5)

Nur bei der zusatzfreien Variante ist das Reduzieren und Lösen des reduzierten Systems unter Verwendung der Hermite- beziehungsweise Laguerre-Polynome für $m = 20$ ungefähr 3- beziehungsweise 1.5-mal schneller als das Reduzieren und Lösen mit Hilfe der Sylvester-Gleichung (4.4) für dieselben Polynome.

Die vergleichende Abbildung 16 im Anhang zeigt, dass auch die Reduktion mit Hilfe des direkten Lösens des linearen Gleichungssystems und das damit verbundene Lösen des reduzierten Systems schneller gegenüber dem Lösen des Originalmodells ist. Im Vergleich zum Lösen der äquivalenten Sylvester-Gleichung ist dieses Verfahren dennoch langsamer.

Für die größte gewählte reduzierte Ordnung $m = 20$ ist daher die Variante, die die Projektionsmatrix durch Lösen der Sylvester-Gleichung bestimmt, für die Hermite- und Laguerre-Polynome ungefähr 2-mal schneller, für die verbleibenden Polynome ungefähr 10-mal schneller.

Die in Abbildung 5.17 dargestellten Modellreduktionsvarianten sparen also im Vergleich zum Lösen des originalen Modells viel Zeit. Konkret sind beide Varianten unter Verwendung der Legendre- und Chebychev-Polynome für $m = 20$ ungefähr 44-mal schneller. Für die gut approximierenden Hermite-Polynome in der zusatzfreien Variante sogar 133-mal schneller.

5.3 Erklärung

In den Darstellungen des relativen Fehlers sind oftmals Zick-Zack-Verläufe vor allem für die Legendre- und beide Chebychev-Polynome für die zuletzt ausgewerteten Re-

duktionsverfahren erkennbar. Diese zeigen sich aber auch in den Grafiken des relativen Fehlers, die das Bestimmen der Projektionsmatrix durch Lösen der äquivalenten Sylvester-Gleichung unter Verwendung der Hermite-Polynome repräsentieren.

Diese Besonderheit soll nun mit Hilfe der in Kapitel 2 vorgestellten Dualitätstheoreme 2.20. und 2.21. erklärt werden. Diese sagen aus, dass jede Lösung $V \in \mathbb{R}^{n \times m}$ einer Sylvester-Gleichung der Form

$$AV - EVS = BL \quad (5.3)$$

einen Krylov-Unterraum aufspannt, bei der die Eigenwerte der Matrix S die Entwicklungspunkte sind.

Somit gibt es unter gewissen Voraussetzungen an die vorgestellten Sylvester-Gleichungen eine Äquivalenz zwischen der Modellreduktionsmethode des Momentenabgleichs und dem in dieser Arbeit präsentierten Reduktionsverfahren, das orthogonale Polynome verwendet.

Darüber hinaus wird in den Dualitätstheoremen in [20] in Abschnitt 2.3 beziehungsweise in Theorem 3.23 in [18] eine Aussage zu den Entwicklungspunkten getroffen. Diese interpolieren im Falle der Beobachtbarkeit des Matrixpaares (S, L) sogar die Pole des Systems und approximieren daher das Originalmodell sehr gut. Ist diese Beobachtbarkeit nicht gegeben, so sind nach Theorem 3.23 in [18] die Eigenwerte von S dennoch Interpolationspunkte des Originalmodells, die aber das System nicht notwendigerweise gut interpolieren.

Um nun eine Aussage zur Äquivalenz zwischen dem Momentenabgleich und den in dieser Arbeit vorgestellten Sylvester-Gleichungen (4.4) beziehungsweise (4.12)

$$\begin{aligned} AV\tilde{E} + EV\tilde{A} &= \tilde{R}, \\ AV\mathcal{E} + EV &= \mathcal{R} \end{aligned}$$

treffen zu können, muss zunächst die Invertierbarkeit der Matrizen \tilde{E} und \mathcal{E} untersucht werden. Im Hauptteil dieser Arbeit wird auf das explizite Aufstellen der inversen Matrizen verzichtet. Diese sind allerdings im Anhang C zu finden.

Unter Verwendung der Hermite-Polynome ist die \tilde{E} -Matrix nur im Falle eines ungerade gewählten m invertierbar, da andernfalls aufgrund von $g_{m-1}(t_0) = 0$ und der besonderen Struktur dieser Matrix eine Nullzeile entsteht und die Matrix somit nicht mehr invertierbar ist. Die \mathcal{E} -Matrix ist für diese orthogonalen Polynome gemäß (2.33) in keinem Fall invertierbar, da diese eine strikte obere Dreiecksmatrix ist.

Werden die Laguerre-Polynome zur Reduktion verwendet, so können beide Matrizen unter den verwendeten Bedingungen in jedem Fall invertiert werden. Unter anderen Voraussetzungen ist die Bedingung an die \tilde{E} -Matrix nur, dass

$$\sum_{i=0}^{m-1} g_i(t_0) \neq 0$$

gilt. Die \mathcal{E} -Matrix ist nach (2.34) unter beliebigen Voraussetzungen invertierbar, da diese Matrix eine obere Dreiecksmatrix ist, deren Hauptdiagonaleinträge alle -1 sind.

Aufgrund derselben Besetztheitsstruktur können die $\tilde{\mathcal{E}}$ - und \mathcal{E} -Matrizen unter Verwendung der Legendre- und beider Chebychev-Polynome zusammengefasst werden.

$\tilde{\mathcal{E}}$ ist unter den verwendeten Bedingungen nur für ein ungerades m beziehungsweise \mathcal{E} unter Verwendung beliebiger Bedingungen nur im Falle eines geraden m invertierbar. Die Irregularität von $\tilde{\mathcal{E}}$ ist auf die Rahmenbedingungen zurückzuführen, denn $g_{2i+1}(t_0) = 0$ für $i = 1, \dots, \frac{m-2}{2}$. Dadurch wird im Falle eines geraden m durch Linearkombination eine Nullzeile in einer geraden Spalte erzeugt. Bei einem ungeraden m in der \mathcal{E} -Matrix ist die erste Zeile eine Linearkombination der anderen ungeraden Zeilen. Bei der Verwendung anderer Ausgangsbedingungen ergeben sich weitere Bedingungen an die $\tilde{\mathcal{E}}$ -Matrix. Für die Legendre-Polynome muss für ein gerades m die Bedingung

$$\begin{aligned} & \frac{g_1(t_0)}{7 \cdot 11 \cdot \dots \cdot (2(m-1) + 1)} + \frac{g_3(t_0)}{3 \cdot 11 \cdot 15 \cdot \dots \cdot (2(m-1) + 1)} + \dots \\ & \dots + \frac{g_{m-3}(t_0)}{3 \cdot 7 \cdot \dots \cdot (2(m-5) + 1) \cdot (2(m-1) + 1)} + \frac{g_{m-1}(t_0)}{3 \cdot 7 \cdot \dots \cdot (2(m-1) + 1)} \neq 0 \end{aligned}$$

beziehungsweise für ein ungerades m

$$\begin{aligned} & \frac{1}{5 \cdot 9 \cdot \dots \cdot (2(m-1) + 1)} + \frac{g_2(t_0)}{9 \cdot 13 \cdot \dots \cdot (2(m-1) + 1)} + \frac{g_4(t_0)}{5 \cdot 13 \cdot 17 \cdot \dots \cdot (2(m-1) + 1)} + \dots \\ & \dots + \frac{g_{m-3}(t_0)}{5 \cdot 9 \cdot \dots \cdot (2(m-5) + 1) \cdot (2(m-1) + 1)} + \frac{g_{m-1}(t_0)}{5 \cdot 9 \cdot \dots \cdot (2(m-3) + 1)} \neq 0 \end{aligned}$$

gelten.

Im Falle der Chebychev-Polynome erster Art muss für ein gerades m

$$\sum_{i=0}^{\frac{m-2}{2}} g_{2i+1}(t_0) \neq 0$$

erfüllt sein beziehungsweise andernfalls

$$\frac{1}{2} + \sum_{i=1}^{\frac{m-1}{2}} g_{2i}(t_0) \neq 0.$$

Die Chebychev-Polynome zweiter Art fordern für ein gerades m

$$\frac{g_1(t_0)}{8 \cdot 12 \cdot \dots \cdot 2m} + \frac{g_3(t_0)}{4 \cdot 12 \cdot 16 \cdot \dots \cdot 2m} + \dots$$

$$\dots + \frac{g_{m-3}(t_0)}{4 \cdot 8 \cdot \dots \cdot 2(m-4) \cdot 2m} + \frac{g_{m-1}(t_0)}{4 \cdot 8 \cdot \dots \cdot 2m} \neq 0$$

beziehungsweise für ein ungerades m

$$\begin{aligned} & \frac{2}{6 \cdot 10 \cdot \dots \cdot 2m} + \frac{g_2(t_0)}{10 \cdot 14 \cdot \dots \cdot 2m} + \frac{g_4(t_0)}{6 \cdot 14 \cdot 18 \cdot \dots \cdot 2m} + \dots \\ & \dots + \frac{g_{m-3}(t_0)}{6 \cdot 10 \cdot \dots \cdot 2(m-4) \cdot 2m} + \frac{g_{m-1}(t_0)}{6 \cdot 10 \cdot \dots \cdot 2(m-2)} \neq 0. \end{aligned}$$

Diese Bedingungen ergeben sich aus den expliziten Inversen, die im Anhang in C für alle orthogonalen Polynome zu finden sind.

Somit können die Dualitätstheoreme 2.20. und 2.21. im Folgenden auf das Reduktionsverfahren, das eine zu (3.5) äquivalente Sylvester-Gleichung löst, für ungerades m beziehungsweise auf die von der Zusatzbedingung unabhängige Variante 2 für gerades m angewendet werden, da in diesen Fällen die \tilde{E} - und \mathcal{E} -Matrizen invertiert sind.

Eine weitere Bedingung der Dualitätstheoreme ist eine Aufspaltung der rechten Seite der Sylvester-Gleichung in die Eingangsmatrix B und einem weiteren Faktor L .

Für ein ungerades m lässt sich die Sylvester-Gleichung (4.4) wie folgt umformulieren:

$$\begin{aligned} & AV\tilde{E} + EV\tilde{A} = \tilde{R} \\ \Leftrightarrow & AV - EV \cdot (-\tilde{A}E^{-1}) = \begin{bmatrix} Ax_0 & Bw_1 & \dots & Bw_{m-1} \end{bmatrix} E^{-1} \\ \Leftrightarrow & AV - EV \cdot (-\tilde{A}E^{-1}) = B \begin{bmatrix} z_0 & w_1 & \dots & w_{m-1} \end{bmatrix} E^{-1} \\ \Leftrightarrow & AV - EVS = BL, \end{aligned}$$

wobei $z_0 = \begin{bmatrix} 0 & \dots & 0 \end{bmatrix}^T \in \mathbb{R}^p$ der Nullvektor ist. Diese Umformungen sind äquivalent, da der Anfangswert $x(t_0) = x_0 = \begin{bmatrix} 0 & \dots & 0 \end{bmatrix}^T$ gewählt wurde und damit die Gleichheit

$$Ax_0 = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} = Bz_0$$

gilt.

Im Falle eines geraden m gilt für die zusatzfreie Sylvester-Gleichung (4.12) folgende Äquivalenz:

$$\begin{aligned} & AV\mathcal{E} + EV = \mathcal{R} \\ \Leftrightarrow & AV - EV \cdot (-\mathcal{E}^{-1}) = \begin{bmatrix} Bw_1 & \dots & Bw_m \end{bmatrix} \mathcal{E}^{-1} \\ \Leftrightarrow & AV - EV \cdot (-\mathcal{E}^{-1}) = B \begin{bmatrix} w_1 & \dots & w_m \end{bmatrix} \mathcal{E}^{-1} \end{aligned}$$

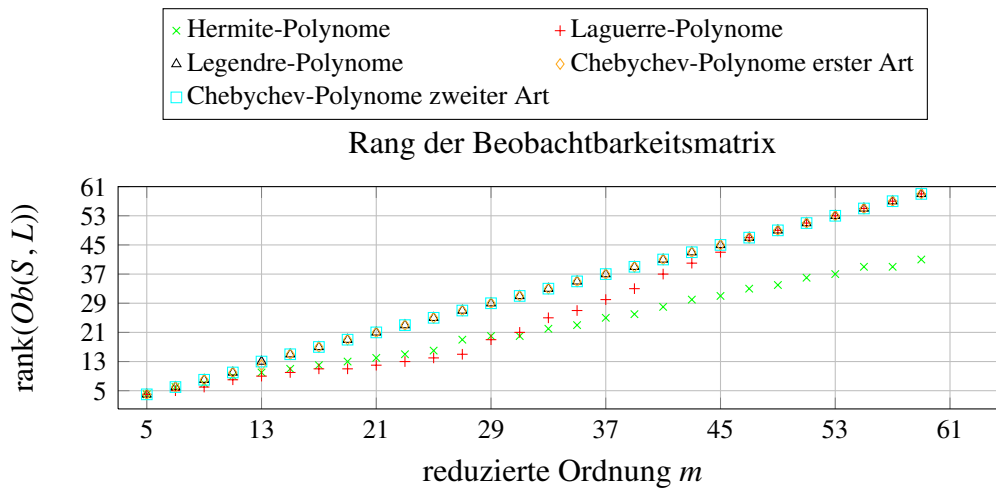


Abbildung 5.18: Rang der Beobachtbarkeitsmatrix für Sylvester-Gleichung (4.4)

$$\Leftrightarrow AV - EVS = B\mathcal{L}.$$

Da es möglich ist, die beiden betrachteten Sylvester-Gleichungen unter gewissen Voraussetzungen in die in beiden Dualitätstheoremen geforderte Form zu bringen, kann eine Äquivalenz zwischen der Modellreduktion mit orthogonalen Polynomen und dem Momentenabgleich hergestellt werden.

Um die Wahl der Entwicklungspunkte zu bewerten, muss zusätzlich die Beobachtbarkeit des Matrixpaares (S, L) beziehungsweise (S, \mathcal{L}) überprüft werden.

Die Abbildungen 5.18 und 5.19 stellen dafür den Rang der Beobachtbarkeitsmatrizen $Ob(S, L)$ beziehungsweise $Ob(S, \mathcal{L})$ in Abhängigkeit von der gewählten reduzierten Ordnung m dar.

Da für die Modellreduktionsvariante, die die zu (3.5) äquivalente Sylvester-Gleichung löst, die \tilde{E} -Matrix nur im Falle eines ungeraden m invertierbar ist, wird der Rang nur für ungerade m dargestellt.

In Abbildung 5.18 wird deutlich, dass der Rang von $Ob(S, L)$ durch die Verwendung der Hermite-Polynome in keinem Fall voll ist. Werden die Laguerre-Polynome eingesetzt, so ist der volle Rang der Beobachtbarkeitsmatrix erst ab $m = 47$ gegeben. Nur für die Legendre- und beide Chebyshev-Polynome ergibt sich für die Beobachtbarkeitsmatrix ab $m = 13$ immer ein voller Rang, sodass durch diese Polynome die besten Entwicklungspunkte gewählt werden.

In Abbildung 5.19, die die Hermite-Polynome aufgrund der Irregularität nicht darstellen, ist der Rang der Beobachtbarkeitsmatrix nur für gerade m dargestellt. Dieser ist unter Verwendung der Laguerre-Polynome erst ab $m = 58$ voll. Im Gegensatz dazu weist die Beobachtbarkeitsmatrix $Ob(S, \mathcal{L})$ der verbleibenden Polynome ab $m = 14$

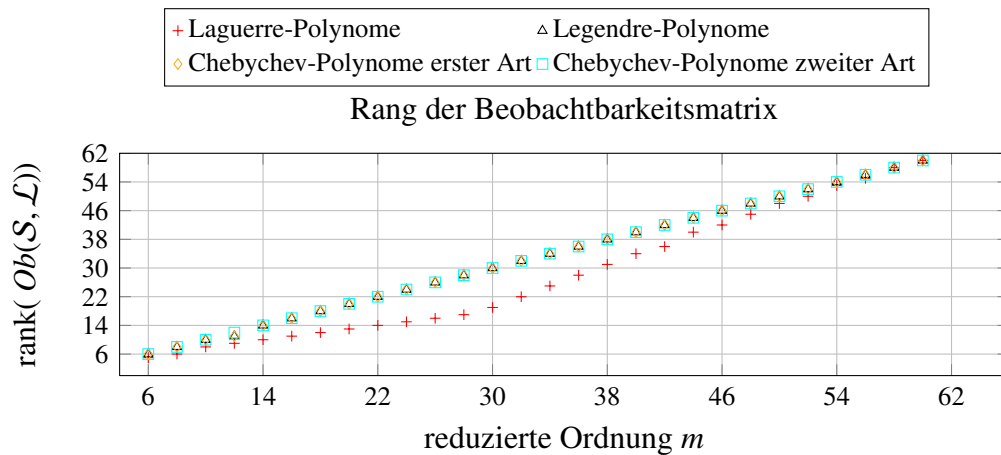


Abbildung 5.19: Rang der Beobachtbarkeitsmatrix für zusatzfreie Sylvester-Gleichung (4.12)

kein Rangdefizit mehr auf, sodass auch bei dieser Variante die besten Entwicklungspunkte durch die Legendre- und beide Chebyshev-Polynome erreicht werden.

Um einen Eindruck von den Entwicklungspunkten für die in dieser Arbeit verwendeten Polynome zu erhalten, werden diese in den Abbildungen 5.20 und 5.21 für $m = 15, 30, 45$ und 60 zusammenfassend dargestellt. Im Anhang D sind separate Grafiken zu den Entwicklungspunkten der einzelnen Polynome zu finden.

Diese Punkte sind die Eigenwerte der Matrizen $S = -\tilde{A}\tilde{E}^{-1}$ beziehungsweise $S = -\mathcal{E}^{-1}$ und wurden als verallgemeinerte Eigenwerte der Matrixbüschel $-\tilde{A} + \lambda\tilde{E}$ beziehungsweise $-I_m + \lambda\mathcal{E}$ bestimmt, um auch bei nicht invertierbarer Matrix \tilde{E} und \mathcal{E} die Entwicklungspunkte darzustellen. Diese werden im Folgenden als Eigenwerte der Paare $(-\tilde{A}, \tilde{E})$ beziehungsweise $(-I_m, \mathcal{E})$ bezeichnet. I_m ist in diesem Fall die $(m \times m)$ -Einheitsmatrix. In Abbildung 5.20 wird zunächst deutlich, dass ein Eigenwert des Paares $(-\tilde{A}, \tilde{E})$ für alle Polynome übereinstimmt, der 0-Eigenwert. Dieser ergibt sich durch die Struktur von \tilde{A} .

Bei den Eigenwerten, die durch die Hermite-Polynome erzeugt werden, fällt auf, dass es nur für $m = 15$ verschiedene Eigenwerte gibt. Dies ist auf die schlechte Konditionierung der \tilde{E} -Matrix zurückzuführen. Dadurch ist es ab $m > 24$ numerisch nicht möglich, weitere Eigenwerte zu berechnen.

Bei der Verwendung der Laguerre-Polynome liegen die Realteile der verbleibenden Eigenwerte alle bei 0.5 und nur der Imaginärteil wird mit wachsendem m vergrößert.

Beim Betrachten der Entwicklungspunkte, die durch die Legendre- und beide Chebyshev-Polynome entstehen wird deutlich, wieso diese Polynome das Originalmodell besser als die Laguerre-Polynome approximieren. Das Spektrum wird mit wachsendem m hin-

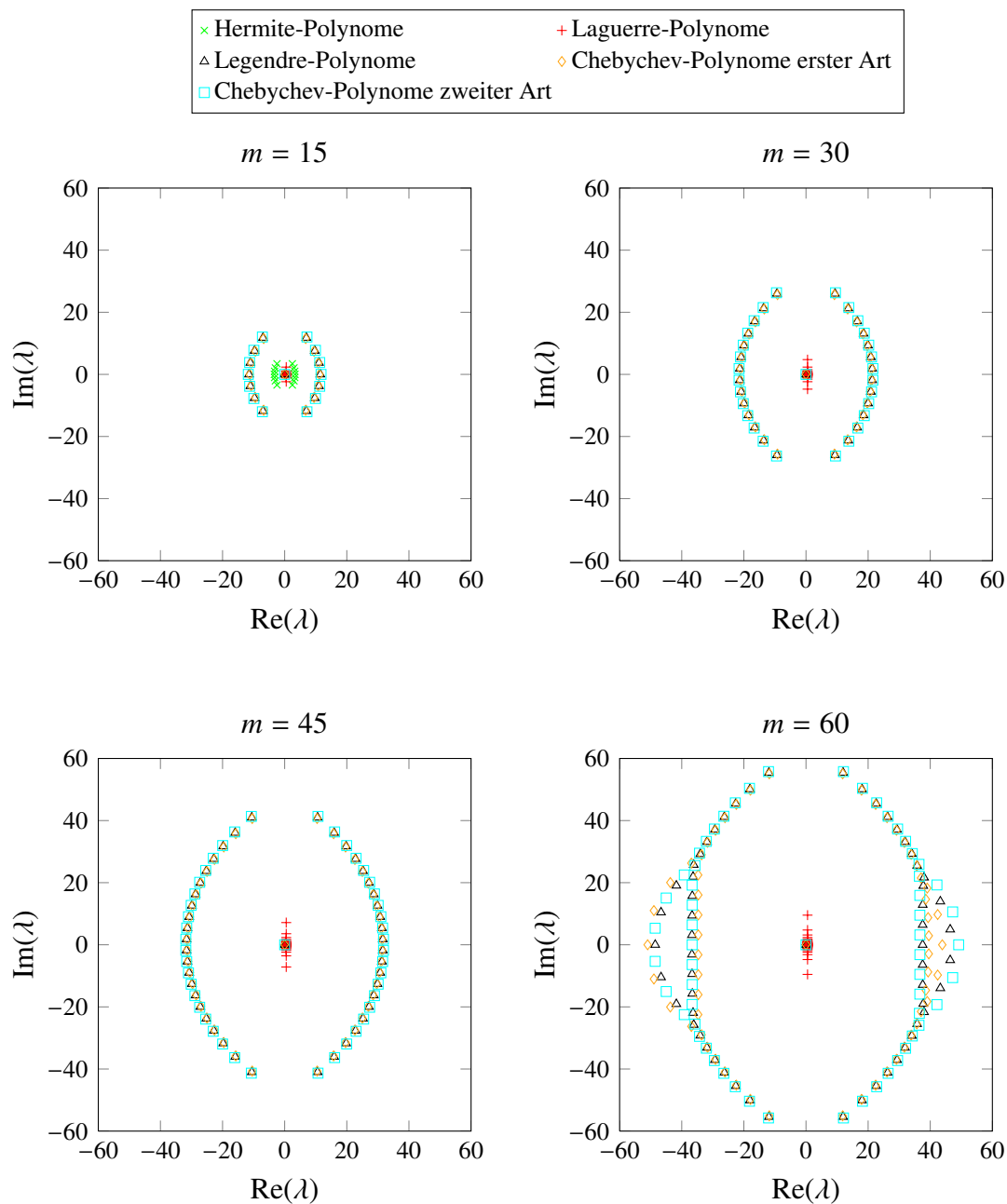


Abbildung 5.20: Endliche Eigenwerte des Matrixpaares $(-\tilde{A}, \tilde{E})$ für $m = 15, 30, 45, 60$

sichtlich des Real- und Imaginärteil erweitert und ist damit auch wesentlich besser als Menge der Entwicklungspunkte geeignet als die der Hermite-Polynome, die ab $m = 25$ nur einen 0 und $m - 1$ ∞ -Eigenwerte besitzen. Eine weitere Auffälligkeit für das Spektrum dieser drei orthogonalen Polynome ist die Symmetrie bezüglich des Real- und

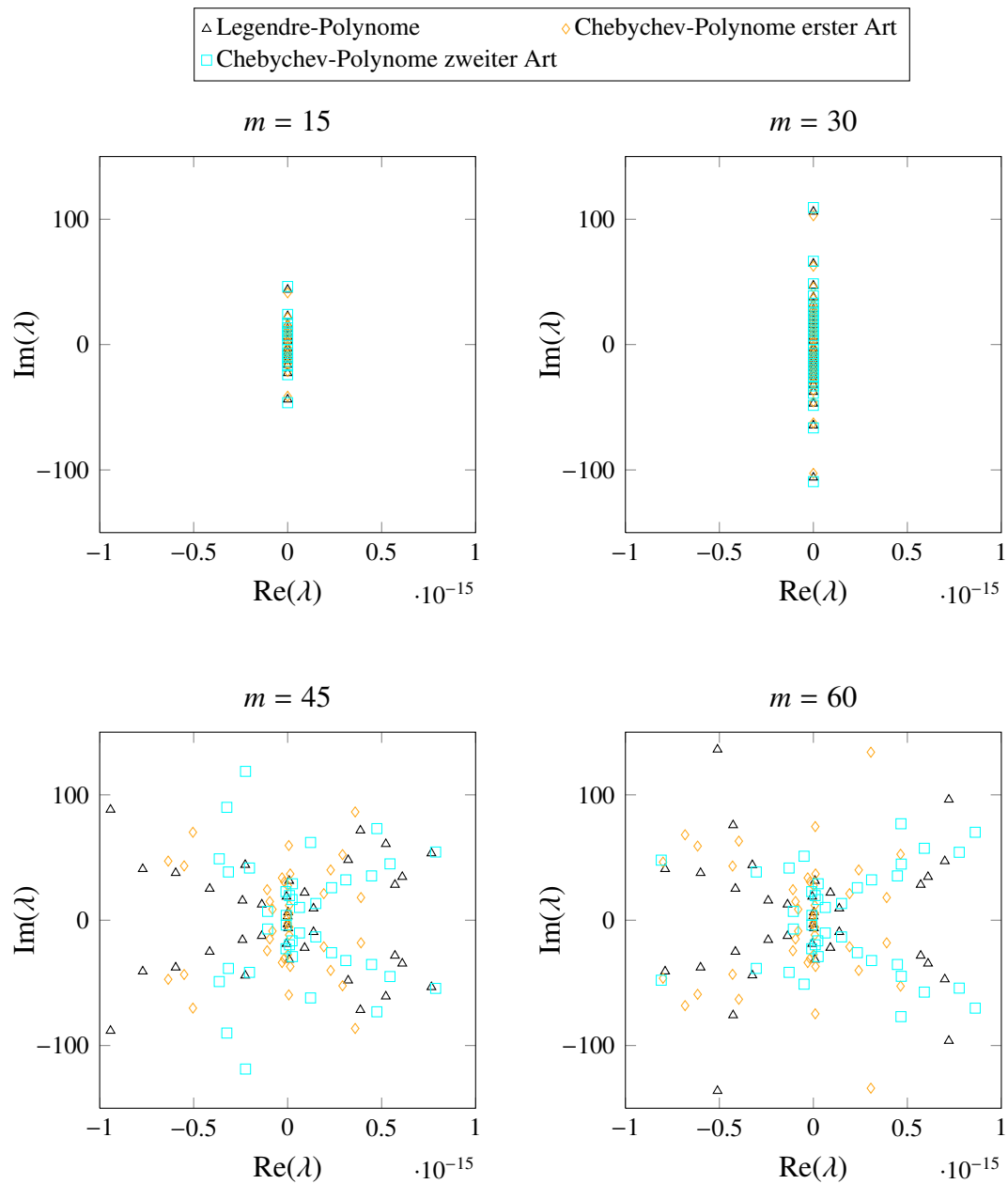


Abbildung 5.21: Endliche Eigenwerte des Matrixpaares $(-I_m, \mathcal{E})$ für $m = 15, 30, 45, 60$

Imaginärteils für $m = 15, 30$ und 45 . Für $m = 60$ ist nur eine Symmetrie bezüglich der reellen Achse gegeben.

Da die \mathcal{E} -Matrix in keinem Fall bei Verwendung der Hermite-Polynome invertierbar ist, sind deren Eigenwerte ∞ und somit nicht in Abbildung 5.21 dargestellt. Desweiteren

wurden die Spektren für die Laguerre-Polynome bewusst ausgelassen, da alle Eigenwerte 1 für beliebige m sind. Die Darstellungen für diese orthogonalen Laguerre-Polynome sind dennoch im Anhang D in Abbildung 22 zu finden.

Um die Unterschiede zwischen den einzelnen verbliebenen orthogonalen Polynomen besser bewerten zu können, wurde die reelle Achse auf einen Bereich von $[-10^{-15}, 10^{-15}]$ eingeschränkt.

In den oberen Grafiken von Abbildung 5.21 sind kaum Unterschiede zwischen den Spektren der Legendre- und beider Chebychev-Polynome erkennbar. Erst in den Darstellungen für $m = 45$ und $m = 60$ werden Streuungen der Eigenwerte um die reelle Achse sichtbar und es ist möglich die Imaginärteile der Eigenwerte besser zu unterscheiden. Hierbei fällt auf, dass die Eigenwerte eine Symmetrie hinsichtlich der reellen Achse aufweisen, was durch genaueres Betrachten der Eigenwerte für $m = 15$ und $m = 30$ bestätigt wird. Die Streuungen um die reelle Achse sind durch Rundungsfehler zu erklären.

Beim Vergleich zwischen den Spektren in den Abbildungen 5.20 und 5.21 fällt besonders auf, dass die verallgemeinerten Eigenwerte des Matrixpaares $(-\tilde{A}, \tilde{E})$ im Vergleich zu denen des Paares $(-I_m, \mathcal{E})$ einen kleineren Imaginärteil haben. Jedoch wird der Realteil dieser Eigenwerte für die Legendre- und beide Chebychev-Polynome und wachsendes m größer und ist nicht wie im Falle des Spektrums von $(-I_m, \mathcal{E})$ gleichbleibend 0.

Aufgrund der in diesem Abschnitt erlangten Erkenntnisse, ist es zunächst möglich, die Verbesserung des relativen Fehlers für sehr groß gewähltes m unter Verwendung der Laguerre-Polynome in beiden ausgewerteten Varianten zu erklären, die in Abbildung 5.14 ersichtlich ist. In der Grafik 5.19 wurde deutlich, dass die Beobachtbarkeitsmatrizen $Ob(S, L)$ beziehungsweise $Ob(S, \mathcal{L})$ ab $m = 47$ beziehungsweise $m = 58$ einen vollen Rang besitzen. Nach den Dualitätstheoremen 2.20. beziehungsweise 2.21. sind die Eigenwerte von S und \mathcal{S} aufgrund dieser Beobachtbarkeit gute Approximationen an die Pole. Somit ist der relative Fehler, der sich durch das Lösen der äquivalenten Sylvester-Gleichung für dieses Polynom ergibt, für ein sehr großes, ungerade gewähltes m klein. Im Falle der zusatzfreien Variante 2 ist eine Verbesserung des relativen Fehlers für ein sehr großes und gerade gewähltes m ersichtlich. Allerdings sind diese relativen Fehler verglichen zu denen der Legendre- und beider Chebychev-Polynome sehr groß. Durch die Dualitätstheoreme konnte dennoch für beliebige m eine Äquivalenz zwischen dem Momentenabgleich und der Modellreduktion mit Hilfe der Laguerre-Polynome im Zeitbereich nachgewiesen werden, die in [9] zuvor schon im Frequenzbereich bewiesen wurde.

Der Zick-Zack-Verlauf, den der relative Fehler für einige orthogonale Polynome beschreibt, ist wie folgt zu erklären.

Da die Beobachtbarkeitsmatrix im Falle der äquivalenten Sylvester-Gleichung unter Verwendung der Hermite-Polynome für ein ungerade gewähltes m immer ein Rangdefizit aufweist, interpoliert das Spektrum von S das Originalmodell nach Theorem 3.23 in [18] nicht bestmöglich. Wird das nächstgrößere gerade m gewählt, so ändern sich die Eigenwerte von S nicht, sondern werden nur um den ∞ -Eigenwert ergänzt, sodass dies wie in Beispiel 5 eine Verbesserung des relativen Fehlers verursacht oder wie in Beispiel 3 eine Verschlechterung. Daher entsteht der Zick-Zack-Verlauf für die Wahl der Hermite-Polynome.

Im Falle der Legendre- und beider Chebychev-Polynome ergibt sich dieses Muster wie folgt. Nach der Analyse der Beobachtbarkeit ist bekannt, dass diese orthogonalen Polynome das Originalmodell in den Polen interpolieren. Für die äquivalente Sylvester-Gleichung (4.4) wurde dies für alle ungeraden m festgestellt beziehungsweise im Falle der zusatzfreien Sylvester-Gleichung (4.12) für alle geraden m . Wird dieses m nun um 1 erhöht, sodass die \tilde{E} - und \mathcal{E} -Matrizen nicht invertierbar sind, so ergibt sich ein ∞ -Eigenwert und die übrigen Eigenwerte werden für beide Varianten leicht verändert. Jedoch kann in diesem Fall keine Aussage über die Interpolation des Originalmodells getroffen werden. Daher ist bei Wahl der äquivalenten Sylvester-Gleichung für ein ungerades m eine Verbesserung des relativen Fehlers für ein wachsendes m ersichtlich, die aber nicht zwangsläufig bei dem folgenden geraden m auftreten muss. Bei der zusatzfreien Variante 2 tritt die Verbesserung des relativen Fehlers bei jedem geraden m ein, aber für das nachfolgende ungerade m kann mit Hilfe des Dualitätstheorem 2.21. keine Aussage gemacht werden. Somit erklärt sich auch ein möglicher Zick-Zack-Verlauf bei den Legendre- und beiden Chebychev-Polynomen.

Aufgrund der in diesem Abschnitt gewonnenen Kenntnisse ist es möglich, eine Äquivalenz zwischen dem in [13] präsentierten Verfahren zur Modellreduktion auf Basis von orthogonalen Polynomen beziehungsweise der in Abschnitt 4.4.2 hergeleiteten Variante und dem Momentenabgleich herzustellen.

Satz 5.1. (Äquivalenz zwischen den auf orthogonalen Polynomen basierenden Modellreduktionsverfahren und dem Momentenabgleich)

Unter den Bedingungen aus Abschnitt 5.1 gelten folgende Äquivalenzen für die orthogonalen Legendre-Polynome und Chebychev-Polynome erster und zweiter Art.

Für eine ungerade gewählte Ordnung m ist das in Abschnitt 4.3 vorgestellte Verfahren zur Modellreduktion, welches äquivalent zu der in [13] präsentierten Methode ist, der Modellreduktion mit Hilfe des in Abschnitt 2.3 vorgestellten Momentenabgleichs gleichzusetzen.

Für eine gerade gewählte reduzierte Ordnung m ist das in Abschnitt 4.4.2 hergeleitete Modellreduktionsverfahren, das ebenfalls auf orthogonalen Polynomen basiert, äquivalent zu einer Modellreduktion durch die Methode des Momentenabgleichs.

Kapitel 6

Zusammenfassung und Ausblick

In diesem Kapitel wird ein Fazit zu den einzelnen ausgewerteten Verfahren gezogen und weitere Betrachtungen, die in dieser Arbeit nicht weiter untersucht werden konnten, in Ausblick gestellt.

6.1 Zusammenfassung

Die Ergebnisse aus Kapitel 5 zeigen, dass es sinnvoll ist, das in [13] hergeleitete und auf orthogonalen Polynomen basierende lineare Gleichungssystem durch eine äquivalent umgeformte Sylvester-Gleichung zu lösen. Die Bestimmung der Projektionsmatrix ist auf diese Weise zum einen wesentlich schneller und zum anderen weniger anfällig für Rundungsfehler. Desweiteren wird viel weniger Speicher als beim Lösen eines Gleichungssystems mit einer $(mn \times mn)$ -Matrix benötigt, bei der n sehr große Werte annehmen kann.

Allerdings eignen sich nicht alle vorgestellten orthogonalen Polynome unter den in dieser Arbeit verwendeten Bedingungen gleich gut. Die Hermite-Polynome interpolieren den Ausgangsvektor nicht gut, was durch Anwenden des Dualitätstheorems 2.21. bestätigt wird. Die Wahl der Laguerre-Polynome ist in allen getesteten Beispielen die Schlechteste. Jedoch interpolieren die Entwicklungspunkte dieser orthogonalen Polynome nach dem Dualitätstheorem 2.21. das Originalmodell ab $m = 47$ gut. Dieser Fall wurde allerdings nur in Beispiel 4 kurz aufgegriffen, da die Reduktion mit Hilfe der Legendre- und beider Chebychev-Polynome für eine wesentlich kleinere Ordnung erfolgversprechender ist, was dasselbe Theorem auch bestätigt.

Im Falle der äquivalenten Sylvester-Gleichung ist es daher ratsam, für eine Modellreduktion die Legendre- oder eines der Chebychev-Polynome und eine ungerade Ordnung zu wählen, da in diesem Fall eine gute Interpolation des Ausgangsvektor nach dem Dualitätstheorem 2.21. garantiert ist.

Das iterative GMRES-Verfahren wurde aufgrund des schlechten Verhältnisses von Aufwand und Nutzen nur im ersten Beispiel ausgetestet. Zwar ist in Abbildung 5.1 eine Abnahme des relativen Fehlers für die Legendre- und beide Chebychev-Polynome erkennbar. Jedoch übersteigt die Gesamtzeit zum Berechnen der Projektionsmatrix und dem Lösen des reduzierten Modells die Zeit zum Lösen des Originalmodells. Daher sollte dieses Verfahren nicht für eine Modellreduktion benutzt werden.

Die Variante der Modellreduktion, die eine vom Anfangswert unabhängige Sylvester-Gleichung löst, wurde ebenfalls nur in Beispiel 1 getestet, ist aber generell unbrauchbar. Auch diese Sylvester-Gleichung kann in eine Struktur gebracht werden, die für die Dualitätstheoreme 2.20. und 2.21. benötigt wird. Jedoch hat die Beobachtbarkeitmatrix, die sich in diesem Fall ergibt, aufgrund der besonderen Struktur der Matrizen \hat{A} und \hat{E} immer ein Rangdefizit für alle orthogonalen Polynome. Die Eigenwerte, die bei dieser Variante immer 0 sind, interpolieren daher die Systempole nicht gut.

Ein weiterer Indiz für die schlechte Approximation der Originaltrajektorie ist die schlechte Konditionierung der \hat{H} -Matrix des äquivalenten Gleichungssystems (4.5), die in der rechten Grafik von Abbildung 5.2 dargestellt ist.

Daher ist von dieser Variante der Modellreduktion abzuraten.

Die von der künstlichen Zusatzbedingung unabhängige Variante der Modellreduktion ist ähnlich wie das Lösen der zu (3.5) äquivalenten Sylvester-Gleichung zu bewerten. Allerdings muss auch hier zwischen den einzelnen orthogonalen Polynomen unterschieden werden.

Die Laguerre-Polynome weisen außer in Beispiel 3 die größten relativen Fehler auf. Nach dem Dualitätstheorem 2.20. interpoliert das Spektrum der Matrix S ab $m = 58$ die Systempole jedoch sehr gut. Dieser Fall wurde allerdings nur in Beispiel 4 betrachtet, da mit Hilfe der Legendre- oder Chebychev-Polynome für ein kleineres m wesentlich bessere Resultate erzielt wurden. Diese Beobachtungen werden auch durch das Dualitätstheorem 2.21. gestützt, das eine sehr gute Interpolation der Systempole für die Legendre- und beide Chebychev-Polynome ab $m = 14$ bestätigt.

Bei der Wahl der Hermite-Polynome kann dieses Theorem allerdings nicht benutzt werden. Aufgrund der in den Auswertungen der Beispiele gemachten Beobachtungen sind diese Polynome in vielen Fällen für ein klein gewähltes m sehr gut geeignet, um den Ausgangsvektor zu interpolieren.

Für diese Variante der Modellreduktion ist es daher empfehlenswert, die Legendre- oder eines der Chebychev-Polynome für ein gerades m zu verwenden, da in diesem Fall eine gute Interpolation der Systempole durch das Dualitätstheorem 2.21. gegeben ist. Es ist desweiteren auch lohnenswert für eine Reduktion mit kleiner Ordnung die Hermite-Polynome zu wählen, da für diese der relative Fehler in den meisten der getesteten Beispiele sehr klein war und die Gesamtzeit im Vergleich zu allen anderen getesteten

Modellreduktionsverfahren wesentlich geringer war.

Eines der Ergebnisse dieser Arbeit ist ein wesentlich schnelleres, aber dennoch zu dem in [13] äquivalentes Verfahren, das in Abschnitt 4.3 entwickelt werden konnte und zusätzlich weniger Speicher benötigt. In den getesteten Beispielen war diese Methode mindestens 1.5-mal schneller und konnte in Beispiel 3 sogar eine Beschleunigung um das 1400-fache verglichen zum direkten Lösen des Gleichungssystems (3.5) aufweisen. Ein weiteres Ziel dieser Arbeit, eine von der in [13] künstlich geschaffenen Zusatzbedingung unabhängige Methode herzuleiten, konnte durch die zusatzfreie Variante 2, die in Abschnitt 4.4.2 vorgestellt wurde, erreicht werden. Diese ist mindestens genauso schnell wie das zuvor erwähnte Verfahren und interpoliert das Originalmodell ebenso gut.

Mit Hilfe der Dualitätstheoreme 2.20. und 2.21. aus [20] konnte die in [9] im Frequenzbereich bewiesene Äquivalenz zwischen einer Modellreduktion mit den Laguerre-Polynomen und dem Momentenabgleich in dieser Arbeit auch im Zeitbereich nachgewiesen werden. Desweiteren ist es gelungen, die Äquivalenz zum Momentenabgleich ebenso auf die Legendre-Polynome und die Chebychev-Polynome erster und zweiter Art auszuweiten (siehe Satz 5.1.).

Die Entwicklungspunkte, die beim Momentenabgleich zum Einsatz kommen, werden bei den in den Abschnitten 4.3 und 4.4.2 präsentierten Modellreduktionsverfahren durch die Polynome sowie deren Ordnung m beeinflusst und sind daher bei gleichbleibenden Bedingungen für beliebige Modelle identisch. Bei der Methode des Momentenabgleichs hingegen können die Entwicklungspunkte frei und an das jeweilige System angepasst gewählt werden.

Daher sollten die in dieser Arbeit vorgestellten Methoden, die auf orthogonalen Polynomen basieren, nicht die erste Wahl bei einer Modellreduktion sein und höchstens dann zum Einsatz kommen, wenn andere Verfahren, wie beispielsweise das balancierte Abschneiden, versagen.

6.2 Ausblick

Da sich die Theorie dieser Arbeit ausschließlich auf LTI-Systeme bezieht, ist es möglich, diese auch auf Modelle zweiter Ordnung

$$\begin{aligned} M\ddot{x}(t) + D\dot{x}(t) + Kx(t) &= Bu(t) \\ y(t) &= Cx(t) \end{aligned}$$

auszuweiten. Ein Ansatz ist dabei das Umformulieren in ein LTI-System, welcher schon in den Beispielen 4 und 5 in den Abschnitten 5.2.4 und 5.2.5 angewandt wurde. Eine weitere Möglichkeit ist das in [21] beschriebene direkte Projizieren, bei dem ähnlich

wie in Abschnitt 3.1 vorgegangen wird.

Eine Angelegenheit, die in dieser Arbeit nicht näher betrachtet wurde, ist die Fehlerabschätzung der in Abschnitt 3.1 vorgestellten Reduktionsmethode. Diese wurde für den Zustands- und Ausgangsvektor des originalen und reduzierten Modells in [13] beziehungsweise [21] durchgeführt. Allerdings wurde dabei nie der Fehler, der durch das Abschneiden der orthogonalen Polynome bei der reduzierten Ordnung m entsteht, beachtet. Daher sollte zukünftig auch dieses Thema näher untersucht werden.

In Abschnitt 5.3 wurde in Satz 5.1. deutlich, dass die Reduktionsverfahren, die in dieser Arbeit am besten funktioniert haben, unter gewissen Voraussetzungen äquivalent zum Momentenabgleich sind. Daher wäre es interessant herauszufinden, in wie weit der Momentenabgleich im direkten Vergleich sogar besser und schneller als die beiden vorgestellten Reduktionsmethoden auf Basis von orthogonalen Polynomen ist.

Literaturverzeichnis

- [1] M. Abramowitz und I. A. Stegun, *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*, Washington, DC: United States Department of Commerce, National Bureau of Standards Applied Mathematics Series 55, 10. Auflage, pp. 771-802, 1972.
- [2] A. Antoulas, *Approximation of Large-Scale Dynamical Systems*, SIAM Publications, 2005
- [3] P. Benner, M. Köhler, J. Saak, *Sparse-Dense Sylvester Equations in \mathcal{H}_2 -Model Order Reduction*, Preprint, 2011 .
- [4] I. N. Bronstein, K. A. Semendjajew, G. Musiol und H. Mühlig, *Taschenbuch der Mathematik*, Frankfurt/Main, Verlag Harri Deutsch, 7. Auflage, 2008.
- [5] B. N. Datta, *Numerical Methods for Linear Control Systems: Design and Analysis*, Elsevier Academic Press, 2004.
- [6] K. B. Datta und B. M. Mohan, *Orthogonal Functions in System and Control*, World Scientific, Advanced Series in Electrical and Computer Engineering, Volume 9, 1995.
- [7] J.W. Demmel, *Applied Numerical Linear Algebra*, SIAM , 1997 .
- [8] R. Eid, *Time Domain Model Reduction By Moment Matching*, Dissertation, Technische Universität München, 2009 .
- [9] R. Eid, B. Salimbahrami und B. Lohmann, *Equivalence of Laguerre-Based Model Order Reduction and Moment Matching*, IEEE Transactions on Automatic Control, Volume 52, pp. 1104-1108, 2007.
- [10] G. Golub, C. van Loan, *Matrix Computations*, Johns Hopkins University Press, third edition , 1996.
- [11] M. Hermann, *Numerik gewöhnlicher Differentialgleichungen, Anfangs- und Randwertprobleme*, Oldenbourg Verlag, München und Wien , 2004.

-
- [12] M. Hund, *Adaptive Berechnung der ADI Shiftparameter in der Niedrig-Rang-Galerkin-ADI*, Bachelorarbeit, Universität Magdeburg, 2012 .
- [13] Y. L. Jiang und H. B. Chen, *Time Domain Model Order Reduction of General Orthogonal Polynomials for Linear Input-Output Systems*, IEEE Transactions on Automatic Control, Volume 57, pp. 330-343, 2012.
- [14] J. Lunze, *Regelungstechnik 2 Mehrgrößensysteme, Digitale Regelung*, Springer-Lehrbuch, Berlin, 7. Auflage, 2013 .
- [15] L. Papula, *Mathematik für Ingenieure und Naturwissenschaftler Band 1*, Friedrich Vieweg & Sohn Verlagsgesellschaft mbH, Braunschweig/Wiesbaden, 10. Auflage, 2001 .
- [16] J. Saak, *Effiziente numerische Lösung eines Optimalsteuerungsproblems für die Abkühlung von Stahlprofilen*, Diplomarbeit, Universität Bremen, 2003 .
- [17] G. Szegő, *Orthogonal Polynomials*, Providence Rhode Island, American Mathematical Society, Volume 23 ,1975.
- [18] A. Vandendorpe, *Model Reduction of Linear Systems, an Interpolation point of View*, Dissertation, Université catholique de Louvain, 2004 .
- [19] D. Werner, *Funktionalanalysis*, Springer-Lehrbuch, Berlin, 2011 .
- [20] T. Wolf, *\mathcal{H}_2 Pseudo-Optimal Model Order Reduction*, Dissertation, Technische Universität München, 2015 .
- [21] Z.H. Xiao und Y. L. Jiang, *Dimension reduction for second-order systems by general orthogonal polynomials*, Mathematical and Computer Modelling of Dynamical Systems, pp. 414-432, 2013.

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Arbeit selbstständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungskommission vorgelegt und auch nicht veröffentlicht.

Ort, Datum

Unterschrift

Anhang

A Berechnung der Differentialrekursionsformel für einzelne orthogonale Polynome

A.1 Legendre-Polynome

Zur Herleitung der Differentialrekursionsformel für die Legendre-Polynome, wird zunächst die aus der Drei-Term-Rekursion hergeleitete Gleichung

$$P_{n+1}(x) = \frac{2n+1}{n+1}xP_n(x) - \frac{n}{n+1}P_{n-1}(x), \quad n = 1, 2, \dots$$

mit $P_0(x) = 1, P_1(x) = x$

nach x differenziert

$$\begin{aligned} \frac{d}{dx}P_{n+1}(x) &= \frac{2n+1}{n+1}P_n(x) + \frac{2n+1}{n+1}x\frac{d}{dx}P_n(x) - \frac{n}{n+1}\frac{d}{dx}P_{n-1}(x) \\ \Leftrightarrow x\frac{d}{dx}P_n(x) &= \frac{n+1}{2n+1}\frac{d}{dx}P_{n+1}(x) - P_n(x) + \frac{n}{2n+1}\frac{d}{dx}P_{n-1}(x). \end{aligned} \quad (\text{A.1})$$

Weiterhin wird

$$(x^2 - 1)\frac{d}{dx}P_n(x) = nxP_n(x) - nP_{n-1}(x)$$

nach x abgeleitet

$$2x\frac{d}{dx}P_n(x) + (x^2 - 1)\frac{d^2}{dx^2}P_n(x) = nP_n(x) + nx\frac{d}{dx}P_n(x) - n\frac{d}{dx}P_{n-1}(x)$$

und anschließend zur legendreschen Differentialgleichung, in der das Legendre-Polynome $P_n(x)$ eingesetzt wurde

$$(1 - x^2)\frac{d^2}{dx^2}P_n(x) - 2x\frac{d}{dx}P_n(x) + n(n+1)P_n(x) = 0,$$

addiert, was zu der folgenden Gleichung führt:

$$\begin{aligned} n(n+1)P_n(x) &= nP_n(x) + nx \frac{d}{dx} P_n(x) - n \frac{d}{dx} P_{n-1}(x) \\ \Leftrightarrow x \frac{d}{dx} P_n(x) &= nP_n(x) + \frac{d}{dx} P_{n-1}(x). \end{aligned}$$

Eine Substitution von $x \frac{d}{dx} P_n(x)$ mit Hilfe von (A.1) führt zu

$$\begin{aligned} nP_n(x) + \frac{d}{dx} P_{n-1}(x) &= \frac{n+1}{2n+1} \frac{d}{dx} P_{n+1}(x) - P_n(x) + \frac{n}{2n+1} \frac{d}{dx} P_{n-1}(x) \\ \Leftrightarrow (n+1)P_n(x) &= \frac{n+1}{2n+1} \frac{d}{dx} P_{n+1}(x) - \frac{n+1}{2n+1} \frac{d}{dx} P_{n-1}(x) \\ \Leftrightarrow P_n(x) &= \frac{1}{2n+1} \frac{d}{dx} P_{n+1}(x) - \frac{1}{2n+1} \frac{d}{dx} P_{n-1}(x), \end{aligned}$$

einer weiteren Rekursionsformel.

A.2 Jacobi-Polynome

Um Jacobi-Polynome durch ihre Ableitungen darstellen zu können, werden zunächst die gegebenen Rekursionen

$$\begin{aligned} 2(n+1)(n+\alpha+\beta+1)(2n+\alpha+\beta)P_{n+1}^{(\alpha,\beta)}(x) &= \\ &= (2n+\alpha+\beta+1)(\alpha^2-\beta^2+(2n+\alpha+\beta)(2n+\alpha+\beta+2)x)P_n^{(\alpha,\beta)}(x) - \\ &\quad - 2(n+\alpha)(n+\beta)(2n+\alpha+\beta+2)P_{n-1}^{(\alpha,\beta)}(x) \\ &\quad \text{mit } P_0^{(\alpha,\beta)}(x) = 1, P_1^{(\alpha,\beta)}(x) = \frac{1}{2}(\alpha-\beta+(\alpha+\beta+2)x) \\ &\quad \text{für } \alpha, \beta > -1, \\ (2n+\alpha+\beta)(1-x^2) \frac{d}{dx} P_n^{(\alpha,\beta)}(x) &= n(\alpha-\beta-(2n+\alpha+\beta)x)P_n^{(\alpha,\beta)}(x) + \\ &\quad + 2(n+\alpha)(n+\beta)P_{n-1}^{(\alpha,\beta)}(x) \end{aligned}$$

nach x differenziert.

Dabei ergeben sich

$$\begin{aligned} 2(n+1)(n+\alpha+\beta+1)(2n+\alpha+\beta) \frac{d}{dx} P_{n+1}^{(\alpha,\beta)}(x) &= \\ &= (2n+\alpha+\beta)(2n+\alpha+\beta+1)(2n+\alpha+\beta+2)P_n^{(\alpha,\beta)}(x) + \\ &\quad + (2n+\alpha+\beta+1)(\alpha^2-\beta^2+(2n+\alpha+\beta)(2n+\alpha+\beta+2)x) \frac{d}{dx} P_n^{(\alpha,\beta)}(x) - \end{aligned}$$

A. Berechnung der Differentialrekursionsformel für einzelne orthogonale Polynome iii

$$\begin{aligned}
 \Leftrightarrow x \frac{d}{dx} P_n^{(\alpha, \beta)}(x) &= \frac{-2(n+\alpha)(n+\beta)(2n+\alpha+\beta+2)}{(2n+\alpha+\beta+1)(2n+\alpha+\beta+2)} \frac{d}{dx} P_{n-1}^{(\alpha, \beta)}(x) \\
 &\quad - \frac{2(n+1)(n+\alpha+\beta+1)}{(2n+\alpha+\beta+1)(2n+\alpha+\beta+2)} \frac{d}{dx} P_{n+1}^{(\alpha, \beta)}(x) - \\
 &\quad - \frac{\alpha^2 - \beta^2}{(2n+\alpha+\beta)(2n+\alpha+\beta+2)} \frac{d}{dx} P_n^{(\alpha, \beta)}(x) - P_n^{(\alpha, \beta)}(x) + \\
 &\quad + \frac{2(n+\alpha)(n+\beta)}{(2n+\alpha+\beta)(2n+\alpha+\beta+1)} \frac{d}{dx} P_{n-1}^{(\alpha, \beta)}(x)
 \end{aligned} \tag{A.2}$$

und

$$\begin{aligned}
 -2x(2n+\alpha+\beta) \frac{d}{dx} P_n^{(\alpha, \beta)}(x) + (2n+\alpha+\beta)(1-x^2) \frac{d^2}{dx^2} P_n^{(\alpha, \beta)}(x) &= \\
 &= -n(2n+\alpha+\beta) P_n^{(\alpha, \beta)}(x) + n(\alpha-\beta - (2n+\alpha+\beta)x) \frac{d}{dx} P_n^{(\alpha, \beta)}(x) + \\
 &\quad + 2(n+\alpha)(n+\beta) \frac{d}{dx} P_{n-1}^{(\alpha, \beta)}(x) \\
 \Leftrightarrow (1-x^2) \frac{d^2}{dx^2} P_n^{(\alpha, \beta)}(x) + (n-2)x \frac{d}{dx} P_n^{(\alpha, \beta)}(x) &= \\
 &= -nP_n^{(\alpha, \beta)}(x) + \frac{n(\alpha-\beta)}{2n+\alpha+\beta} \frac{d}{dx} P_n^{(\alpha, \beta)}(x) + \frac{2(n+\alpha)(n+\beta)}{2n+\alpha+\beta} \frac{d}{dx} P_{n-1}^{(\alpha, \beta)}(x) \\
 \Leftrightarrow (1-x^2) \frac{d^2}{dx^2} P_n^{(\alpha, \beta)}(x) + (n-2)x \frac{d}{dx} P_n^{(\alpha, \beta)}(x) - \frac{n(\alpha-\beta)}{2n+\alpha+\beta} \frac{d}{dx} P_n^{(\alpha, \beta)}(x) + nP_n^{(\alpha, \beta)}(x) &= \\
 &= \frac{2(n+\alpha)(n+\beta)}{2n+\alpha+\beta} \frac{d}{dx} P_{n-1}^{(\alpha, \beta)}(x).
 \end{aligned} \tag{A.3}$$

Durch das Einsetzen der Lösung $P_n^{(\alpha, \beta)}(x)$ in die Differentialgleichung

$$(1-x^2) \frac{d^2}{dx^2} y(x) + (\beta - \alpha - (\alpha + \beta + 2)x) \frac{d}{dx} y(x) + n(n + \alpha + \beta + 1) y(x) = 0$$

ergibt sich

$$(1-x^2) \frac{d^2}{dx^2} P_n^{(\alpha, \beta)}(x) + (\beta - \alpha - (\alpha + \beta + 2)x) \frac{d}{dx} P_n^{(\alpha, \beta)}(x) + n(n + \alpha + \beta + 1) P_n^{(\alpha, \beta)}(x) = 0.$$

Die Differenz dieser Gleichung mit (A.3) und eine Substitution mit Hilfe von (A.2) liefert

$$\begin{aligned}
 -\frac{2(n+\alpha)(n+\beta)}{2n+\alpha+\beta} \frac{d}{dx} P_{n-1}^{(\alpha, \beta)}(x) &= \left(\beta - \alpha + \frac{n(\alpha-\beta)}{2n+\alpha+\beta} - (n+\alpha+\beta)x \right) \frac{d}{dx} P_n^{(\alpha, \beta)}(x) + \\
 &\quad + n(n+\alpha+\beta) P_n^{(\alpha, \beta)}(x)
 \end{aligned}$$

$$\begin{aligned}
&\Leftrightarrow -\frac{2(n+\alpha)(n+\beta)}{2n+\alpha+\beta} \frac{d}{dx} P_{n-1}^{(\alpha,\beta)}(x) = -(n+\alpha+\beta)x \frac{d}{dx} P_n^{(\alpha,\beta)}(x) + \\
&\quad + \frac{(\beta-\alpha)(2n+\alpha+\beta) + n(\alpha-\beta)}{2n+\alpha+\beta} \frac{d}{dx} P_n^{(\alpha,\beta)}(x) + \\
&\quad + n(n+\alpha+\beta) P_n^{(\alpha,\beta)}(x) \\
&\Leftrightarrow x \frac{d}{dx} P_n^{(\alpha,\beta)}(x) = \frac{\beta-\alpha}{2n+\alpha+\beta} \frac{d}{dx} P_n^{(\alpha,\beta)}(x) + n P_n^{(\alpha,\beta)}(x) + \\
&\quad + \frac{2(n+\alpha)(n+\beta)}{(2n+\alpha+\beta)(n+\alpha+\beta)} \frac{d}{dx} P_{n-1}^{(\alpha,\beta)}(x) \\
&\stackrel{(A.2)}{\Leftrightarrow} -P_n^{(\alpha,\beta)}(x) + \frac{2(n+1)(n+\alpha+\beta+1)}{(2n+\alpha+\beta+1)(2n+\alpha+\beta+2)} \frac{d}{dx} P_{n+1}^{(\alpha,\beta)}(x) - \\
&\quad - \frac{\alpha^2 - \beta^2}{(2n+\alpha+\beta)(2n+\alpha+\beta+2)} \frac{d}{dx} P_n^{(\alpha,\beta)}(x) + \\
&\quad + \frac{2(n+\alpha)(n+\beta)}{(2n+\alpha+\beta)(2n+\alpha+\beta+1)} \frac{d}{dx} P_{n-1}^{(\alpha,\beta)}(x) = \\
&\quad = \frac{\beta-\alpha}{2n+\alpha+\beta} \frac{d}{dx} P_n^{(\alpha,\beta)}(x) + n P_n^{(\alpha,\beta)}(x) + \\
&\quad + \frac{2(n+\alpha)(n+\beta)}{(2n+\alpha+\beta)(n+\alpha+\beta)} \frac{d}{dx} P_{n-1}^{(\alpha,\beta)}(x) \\
&\Leftrightarrow (n+1) P_n^{(\alpha,\beta)}(x) = \frac{2(n+1)(n+\alpha+\beta+1)}{(2n+\alpha+\beta+1)(2n+\alpha+\beta+2)} \frac{d}{dx} P_{n+1}^{(\alpha,\beta)}(x) - \\
&\quad - \frac{(\alpha^2 - \beta^2) + (2n+\alpha+\beta+2)(\beta-\alpha)}{(2n+\alpha+\beta)(2n+\alpha+\beta+2)} \frac{d}{dx} P_n^{(\alpha,\beta)}(x) + \\
&\quad + \frac{2(n+\alpha)(n+\beta)(-n-1)}{(n+\alpha+\beta)(2n+\alpha+\beta)(2n+\alpha+\beta+1)} \frac{d}{dx} P_{n-1}^{(\alpha,\beta)}(x) \\
&\Leftrightarrow P_n^{(\alpha,\beta)}(x) = \frac{2(n+\alpha+\beta+1)}{(2n+\alpha+\beta+1)(2n+\alpha+\beta+2)} \frac{d}{dx} P_{n+1}^{(\alpha,\beta)}(x) + \\
&\quad + \frac{2(\alpha-\beta)}{(2n+\alpha+\beta)(2n+\alpha+\beta+2)} \frac{d}{dx} P_n^{(\alpha,\beta)}(x) - \\
&\quad - \frac{2(n+\alpha)(n+\beta)}{(n+\alpha+\beta)(2n+\alpha+\beta)(2n+\alpha+\beta+1)} \frac{d}{dx} P_{n-1}^{(\alpha,\beta)}(x),
\end{aligned}$$

eine Darstellung eines Jacobi-Polynoms mit Hilfe seiner Ableitung und die seines Vorgängers und Nachfolgers.

A.3 Chebychev-Polynome erster Art

Für die Herleitung der Differentialrekursionsformel für die Chebychev-Polynome erster Art wird zunächst die Drei-Term-Rekursion

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \quad n = 1, 2, \dots$$

mit $T_0(x) = 1, T_1(x) = x$

nach der Variablen x differenziert:

$$\begin{aligned} \frac{d}{dx} T_{n+1}(x) &= 2T_n(x) + 2x \frac{d}{dx} T_n(x) - \frac{d}{dx} T_{n-1}(x) \\ \Leftrightarrow x \frac{d}{dx} T_n(x) &= \frac{1}{2} \frac{d}{dx} T_{n+1}(x) - T_n(x) + \frac{1}{2} \frac{d}{dx} T_{n-1}(x). \end{aligned} \quad (\text{A.4})$$

Die Rekursionsvorschrift

$$(1 - x^2) \frac{d}{dx} T_n(x) = -n x T_n(x) + n T_{n-1}(x),$$

wird ebenfalls nach x abgeleitet

$$-2x \frac{d}{dx} T_n(x) + (1 - x^2) \frac{d^2}{dx^2} T_n(x) = -n T_n(x) - n x \frac{d}{dx} T_n(x) + n \frac{d}{dx} T_{n-1}(x)$$

um dies von der Chebychev-Differentialgleichung, in die das Chebychev-Polynom $T_n(x)$ als Lösung eingesetzt wird

$$(1 - x^2) \frac{d^2}{dx^2} T_n(x) - x \frac{d}{dx} T_n(x) + n^2 T_n(x) = 0$$

zu subtrahieren

$$\begin{aligned} x \frac{d}{dx} T_n(x) + n^2 T_n(x) &= n T_n(x) + n x \frac{d}{dx} T_n(x) - n \frac{d}{dx} T_{n-1}(x) \\ \Leftrightarrow (n - 1) x \frac{d}{dx} T_n(x) &= n(n - 1) T_n(x) + n \frac{d}{dx} T_{n-1}(x) \\ \Leftrightarrow x \frac{d}{dx} T_n(x) &= n T_n(x) + \frac{n}{n - 1} \frac{d}{dx} T_{n-1}(x) \\ \stackrel{(\text{A.4})}{\Leftrightarrow} n T_n(x) + \frac{n}{n - 1} \frac{d}{dx} T_{n-1}(x) &= \frac{1}{2} \frac{d}{dx} T_{n+1}(x) - T_n(x) + \frac{1}{2} \frac{d}{dx} T_{n-1}(x) \\ \Leftrightarrow (n + 1) T_n(x) &= \frac{1}{2} \frac{d}{dx} T_{n+1} - \frac{n + 1}{2(n - 1)} \frac{d}{dx} T_{n-1}(x) \\ \Leftrightarrow T_n(x) &= \frac{1}{2(n + 1)} \frac{d}{dx} T_{n+1} - \frac{1}{2(n - 1)} \frac{d}{dx} T_{n-1}(x), \end{aligned}$$

woraus sich die Differentialrekursionsformel für die Chebychev-Polynome erster Art ergeben.

A.4 Chebychev-Polynome zweiter Art

Für den Erhalt der Differentialrekursionsformel wird äquivalent wie bei den Chebychev-Polynomen erster Art vorgegangen.

Die Ableitung nach x der Drei-Term-Rekursion

$$U_{n+1}(x) = 2xU_n(x) - U_{n-1}(x), \quad n = 1, 2, \dots$$

mit $U_0(x) = 1, U_1(x) = 2x$

liefert

$$\begin{aligned} \frac{d}{dx}U_{n+1}(x) &= 2U_n(x) + 2x\frac{d}{dx}U_n(x) - \frac{d}{dx}U_{n-1}(x) \\ \Leftrightarrow x\frac{d}{dx}U_n(x) &= \frac{1}{2}\frac{d}{dx}U_{n+1}(x) - U_n(x) + \frac{1}{2}\frac{d}{dx}U_{n-1}(x) \end{aligned} \quad (\text{A.5})$$

beziehungsweise für die Differentiation der Gleichung

$$(1 - x^2)\frac{d}{dx}U_n(x) = -nxU_n(x) + (n+1)U_{n-1}(x),$$

nach x ergibt sich

$$\begin{aligned} -2x\frac{d}{dx}U_n(x) + (1 - x^2)\frac{d^2}{dx^2}U_n(x) &= -nU_n(x) - nx\frac{d}{dx}U_n(x) + (n+1)\frac{d}{dx}U_{n-1}(x) \\ \Leftrightarrow (1 - x^2)\frac{d^2}{dx^2}U_n(x) + (n-2)x\frac{d}{dx}U_n(x) &= -nU_n(x) + (n+1)\frac{d}{dx}U_{n-1}(x). \end{aligned} \quad (\text{A.6})$$

Wird $U_n(x)$ als Lösung in die Differentialgleichung, die die Chebychev-Polynome zweiter Art lösen, eingesetzt

$$(1 - x^2)\frac{d^2}{dx^2}U_n(x) - 3x\frac{d}{dx}U_n(x) + n(n+2)U_n(x) = 0$$

und das Ergebnis (A.6) davon subtrahiert, so ergibt sich

$$\begin{aligned} -(n+1)x\frac{d}{dx}U_n(x) + n(n+2)U_n(x) &= nU_n(x) - (n+1)\frac{d}{dx}U_{n-1}(x) \\ \Leftrightarrow -x\frac{d}{dx}U_n(x) &= -nU_n(x) - \frac{d}{dx}U_{n-1}(x) \\ \stackrel{(\text{A.5})}{\Leftrightarrow} -nU_n(x) - \frac{d}{dx}U_{n-1}(x) &= -\left(\frac{1}{2}\frac{d}{dx}U_{n+1}(x) - U_n(x) + \frac{1}{2}\frac{d}{dx}U_{n-1}(x)\right) \\ \Leftrightarrow (n+1)U_n(x) &= \frac{1}{2}\frac{d}{dx}U_{n+1}(x) - \frac{1}{2}\frac{d}{dx}U_{n-1}(x) \\ \Leftrightarrow U_n(x) &= \frac{1}{2(n+1)}\frac{d}{dx}U_{n+1}(x) - \frac{1}{2(n+1)}\frac{d}{dx}U_{n-1}(x), \end{aligned}$$

die Differentialrekursionsformel für die Chebychev-Polynome zweiter Art.

B Weitere Abbildungen für die Auswertung

In diesem Abschnitt folgen kommentarlos die in Kapitel 5 ausgelassenen Grafiken zu den absoluten Fehlern für alle Beispiele, sowie die Vergleiche der relativen Fehler und der Ausführungszeit zwischen dem direkten Lösen des linearen Gleichungssystems (3.5) und dem Lösen der dazu äquivalenten Sylvester-Gleichung (4.4) für die Beispiele 2, 3, 4 und 5, sowie der Vergleich der Trajektorien, die durch das GMRES-Verfahren mit beiden Vorkonditionierern bestimmt wurden, zur Originaltrajektorie von Beispiel 1.

B.1 Beispiel 1

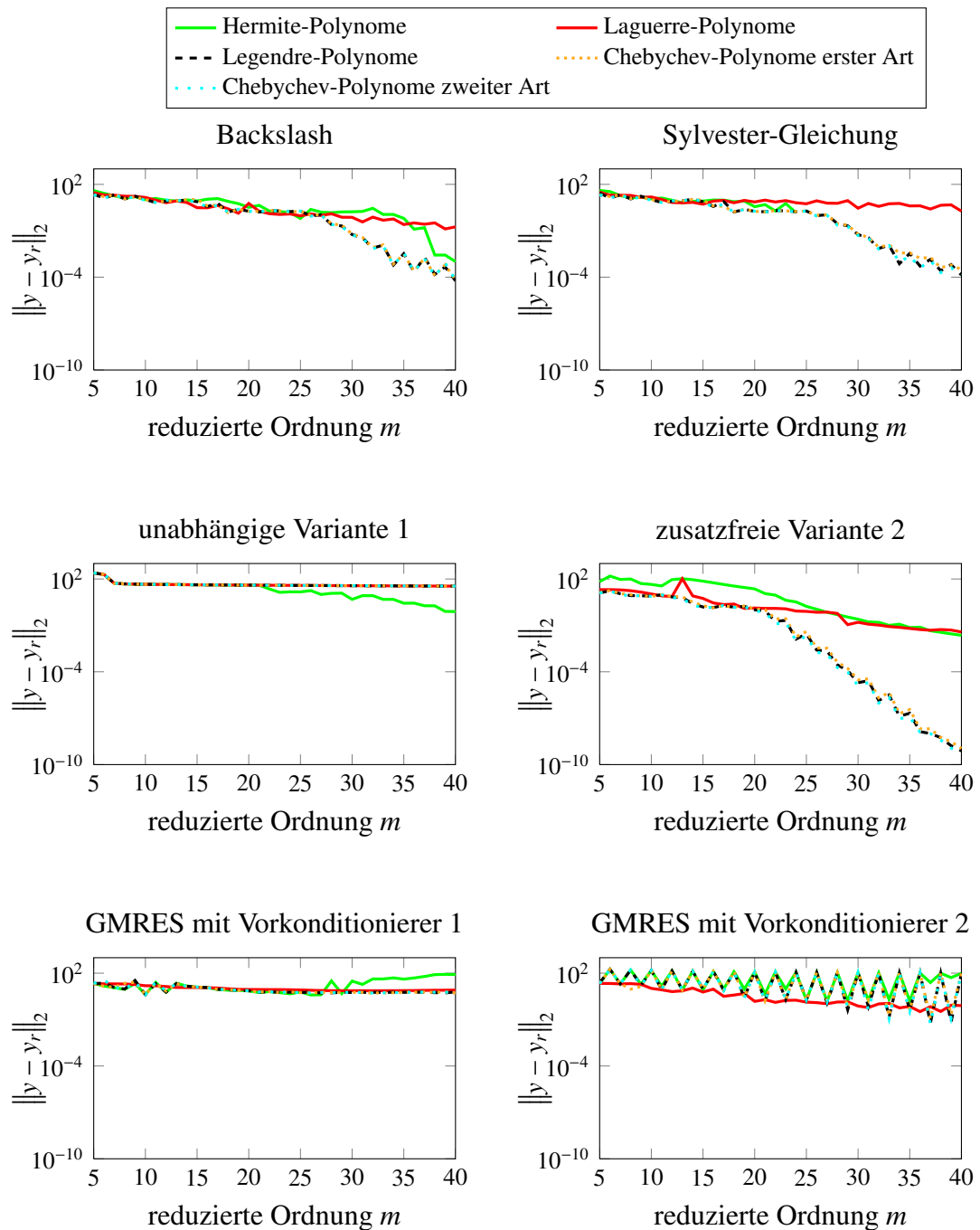


Abbildung 1: Absoluter Fehler durch Reduktion mit Hilfe vorgestellter Verfahren (Beispiel 1)

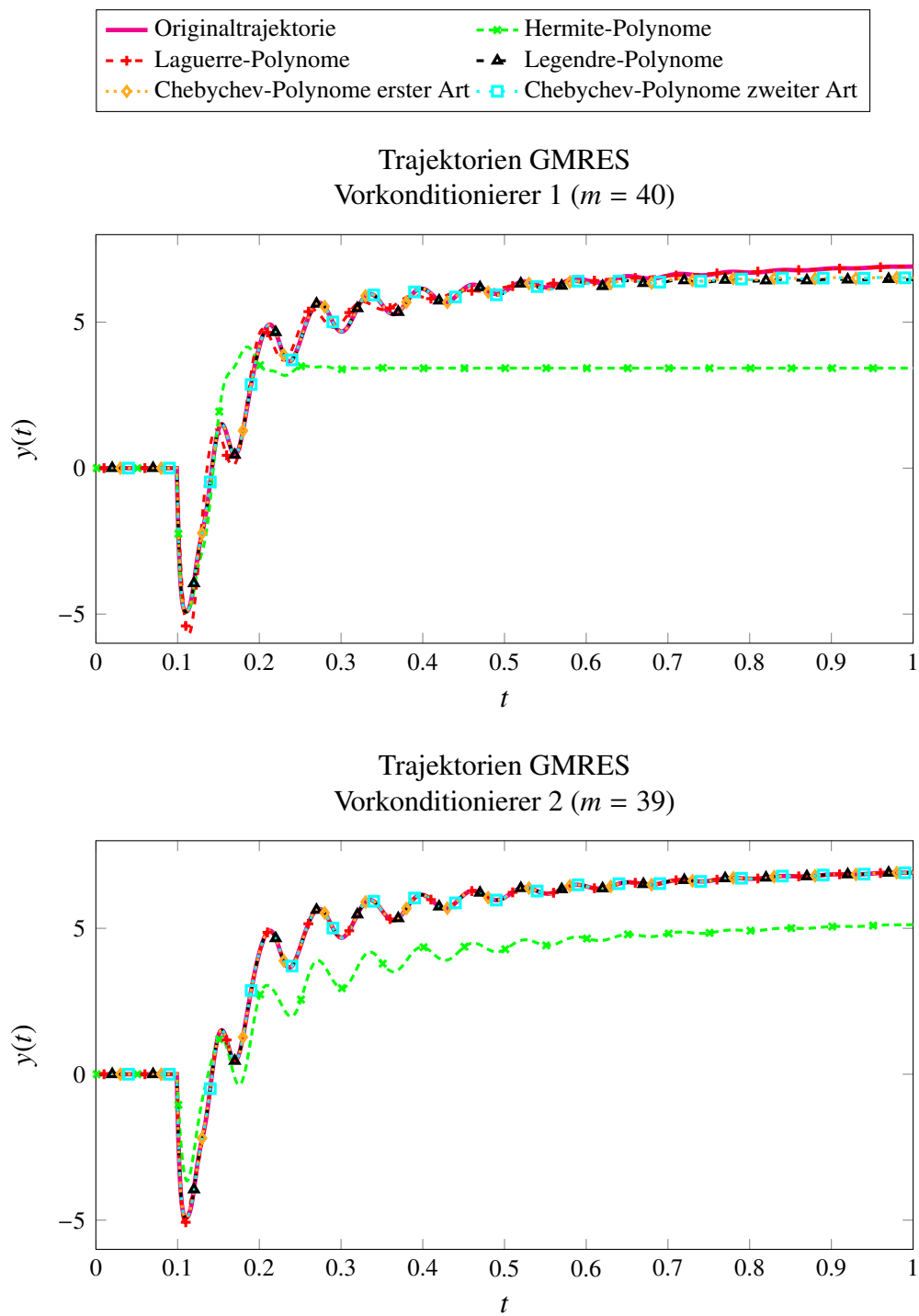


Abbildung 2: Trajektorien für das GMRES-Verfahren mit beiden Vorkonditionierern für Beispiel 1

B.2 Beispiel 2

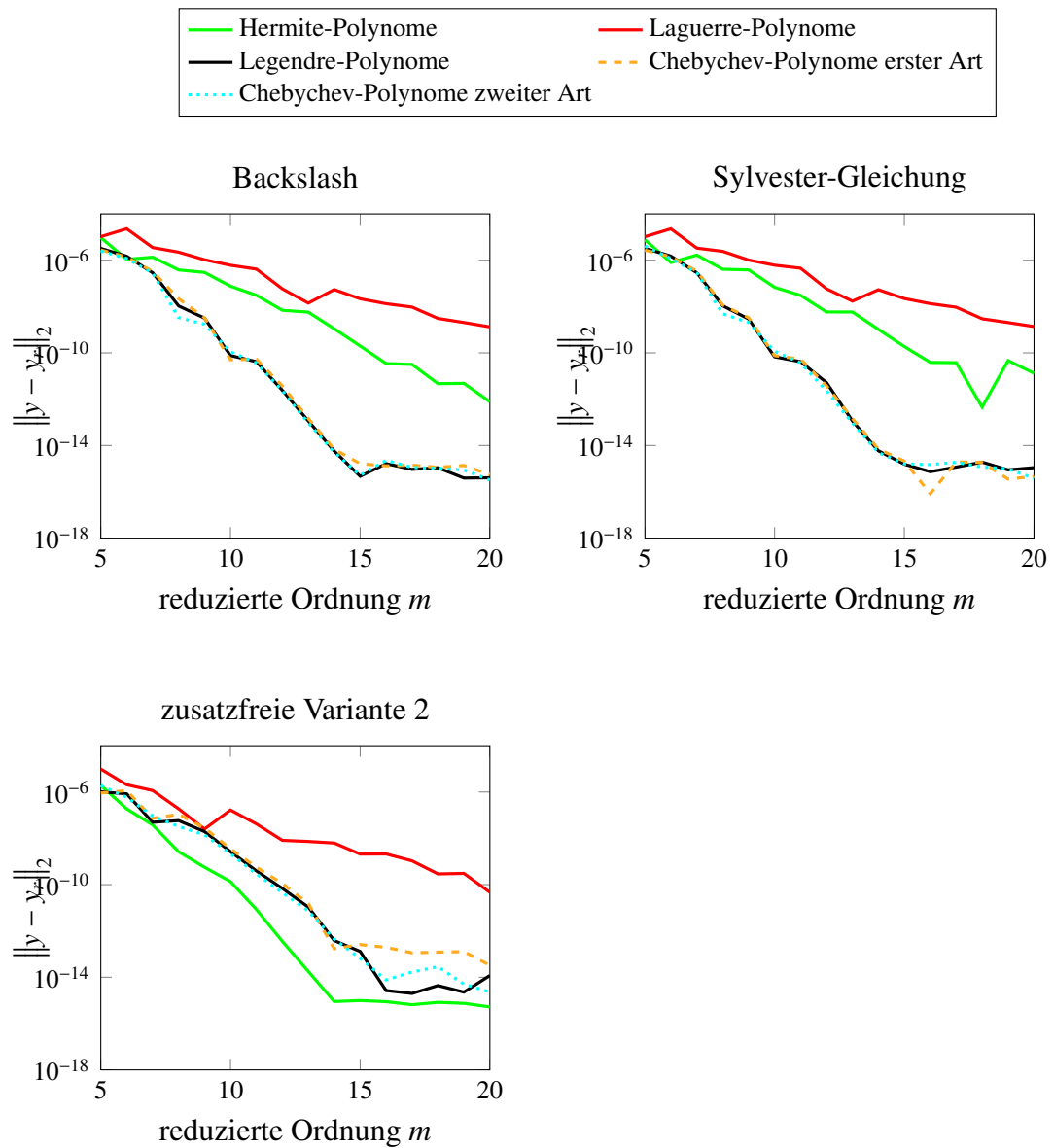


Abbildung 3: Absoluter Fehler durch Reduktion mit Hilfe der verbleibenden Lösungsstrategien (Beispiel 2)

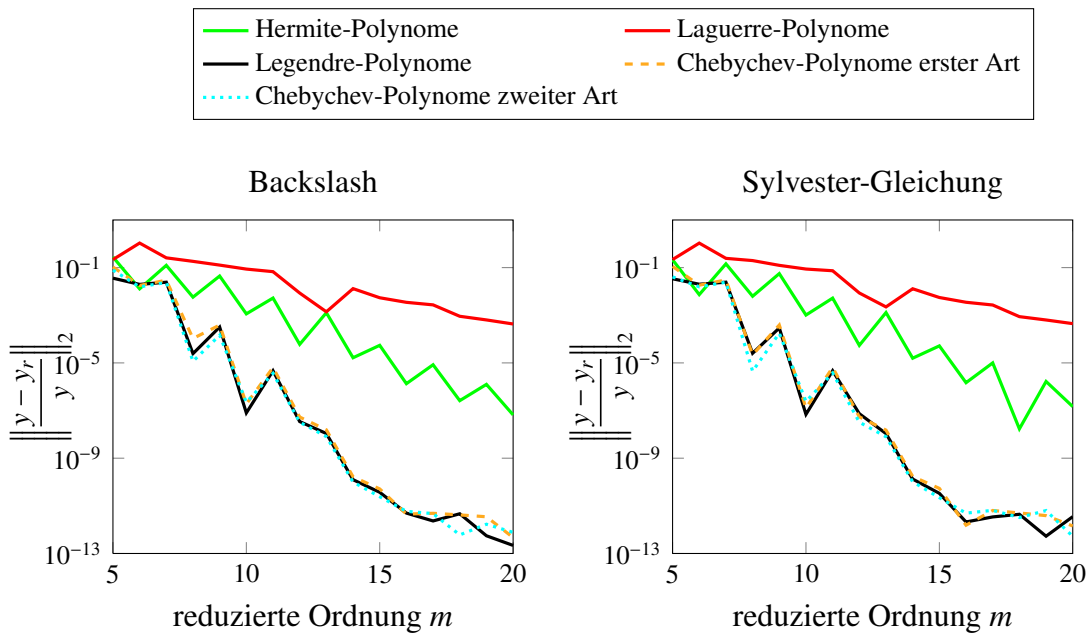


Abbildung 4: Relativer Fehler - Vergleich (Beispiel 2)

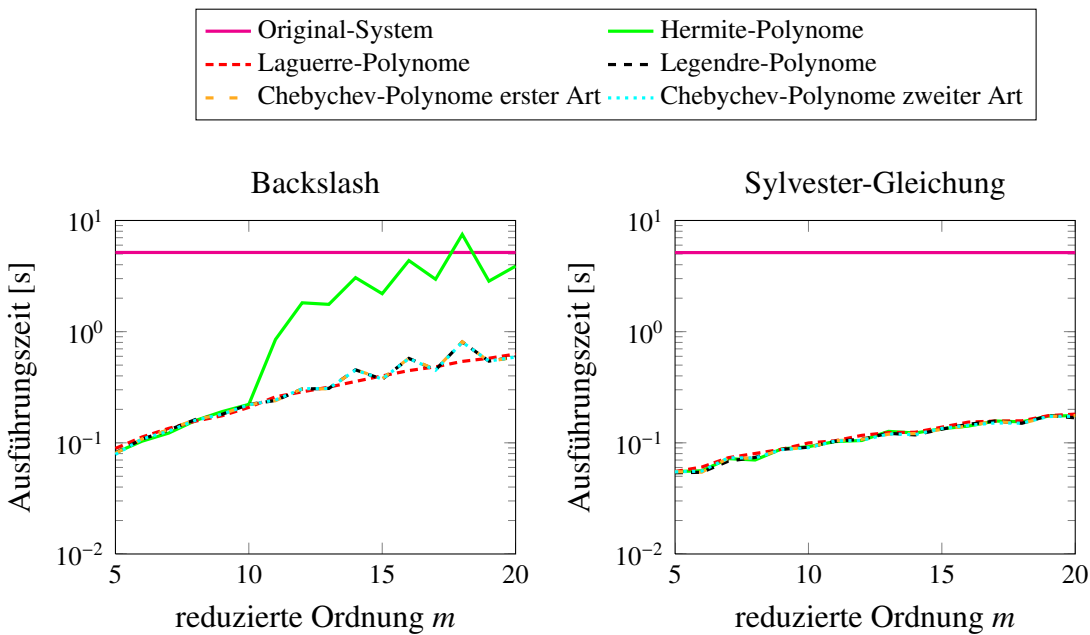


Abbildung 5: Ausführungszeit - Vergleich (Beispiel 2)

B.3 Beispiel 3

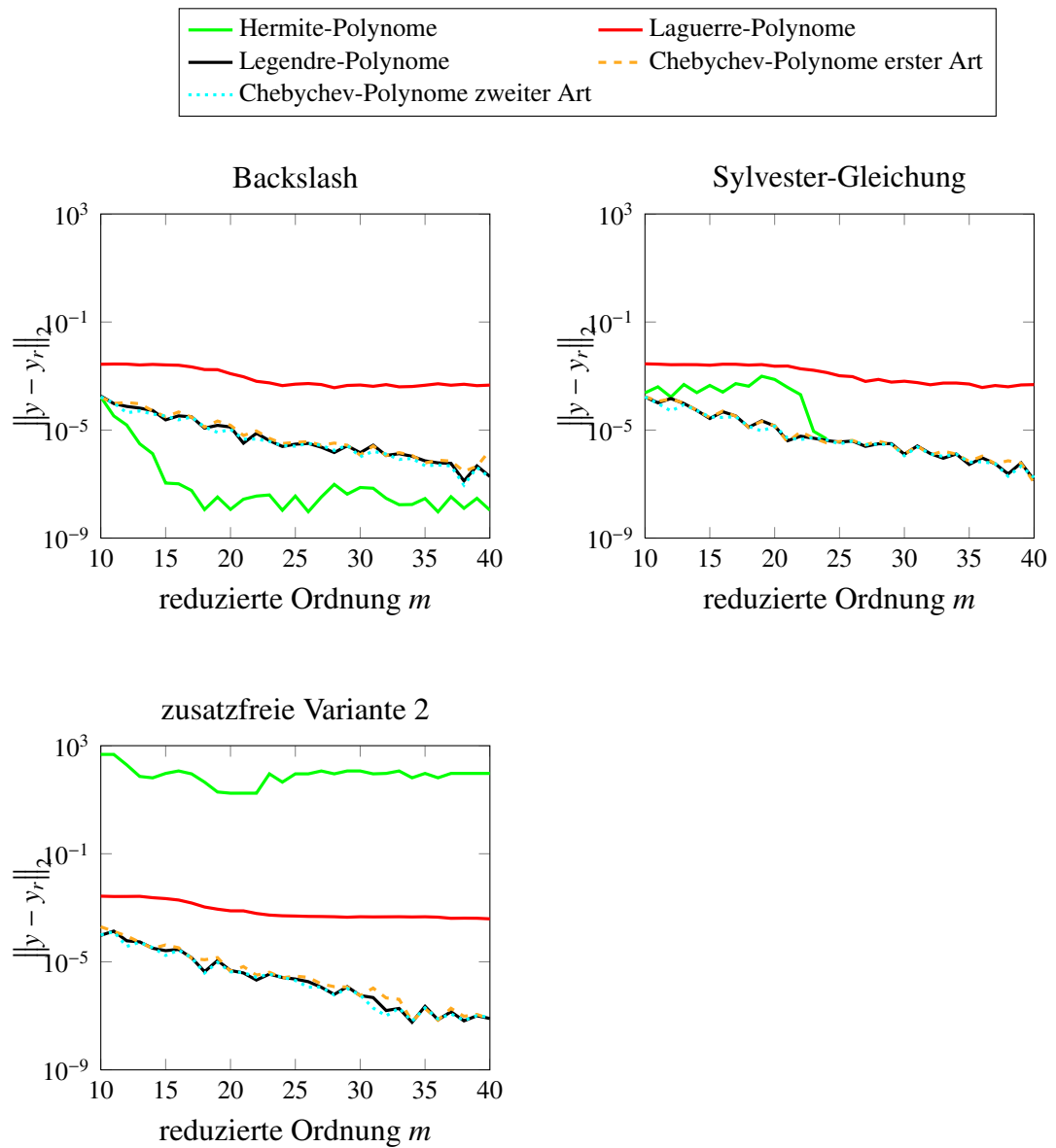


Abbildung 6: Absoluter Fehler durch Reduktion mit Hilfe der verbleibenden Lösungsstrategien (Beispiel 3)

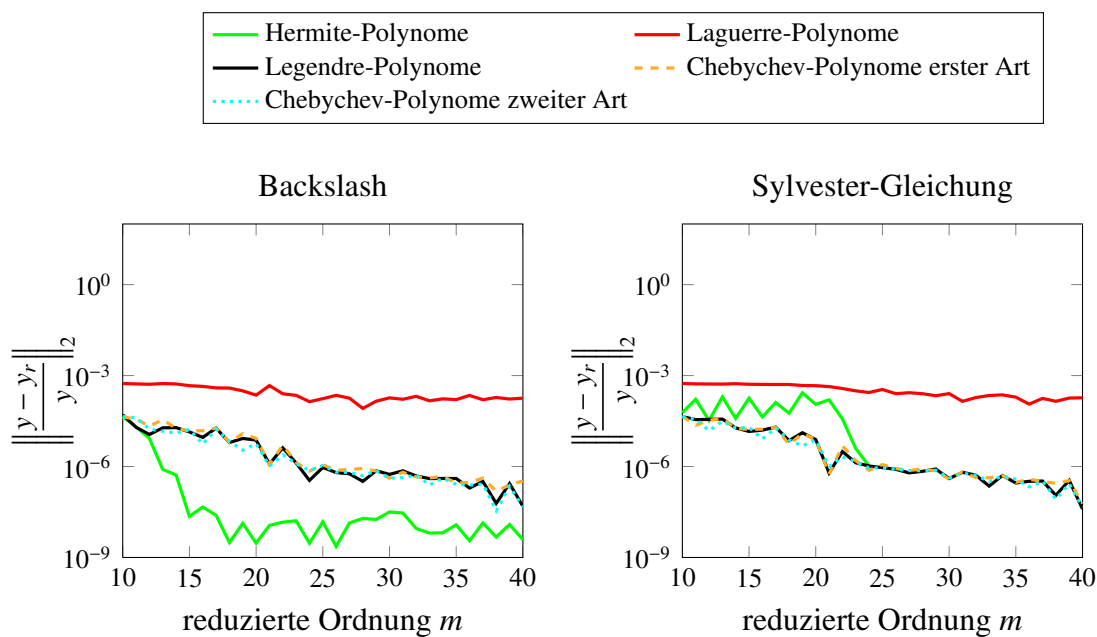


Abbildung 7: Relativer Fehler - Vergleich (Beispiel 3)

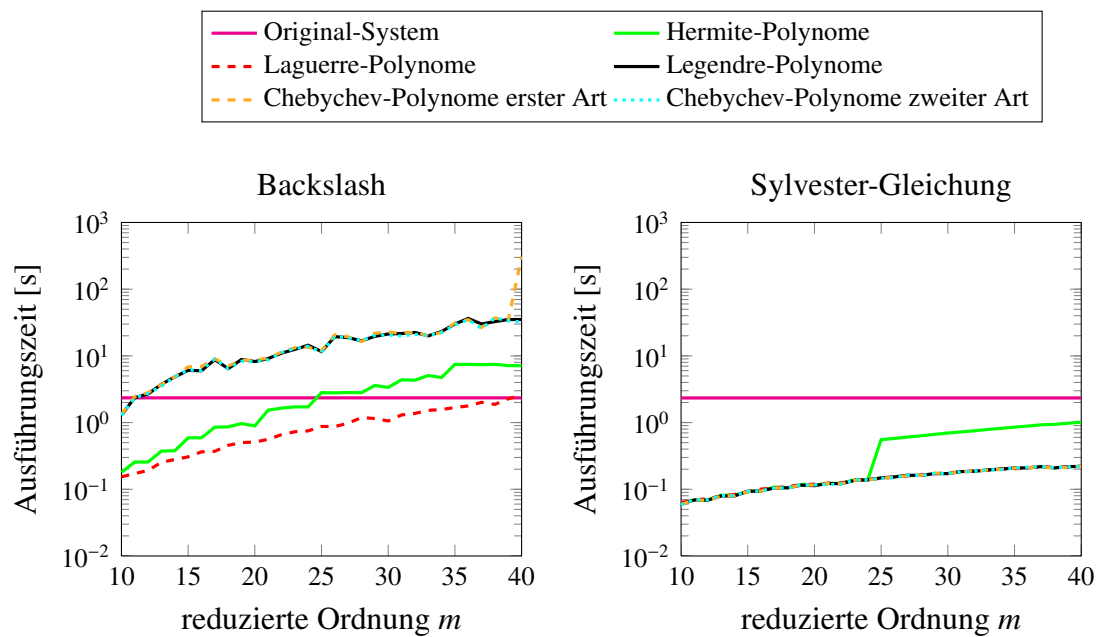


Abbildung 8: Ausführungszeit - Vergleich (Beispiel 3)

B.4 Beispiel 4

Grafiken für $m \in [20, 30]$:

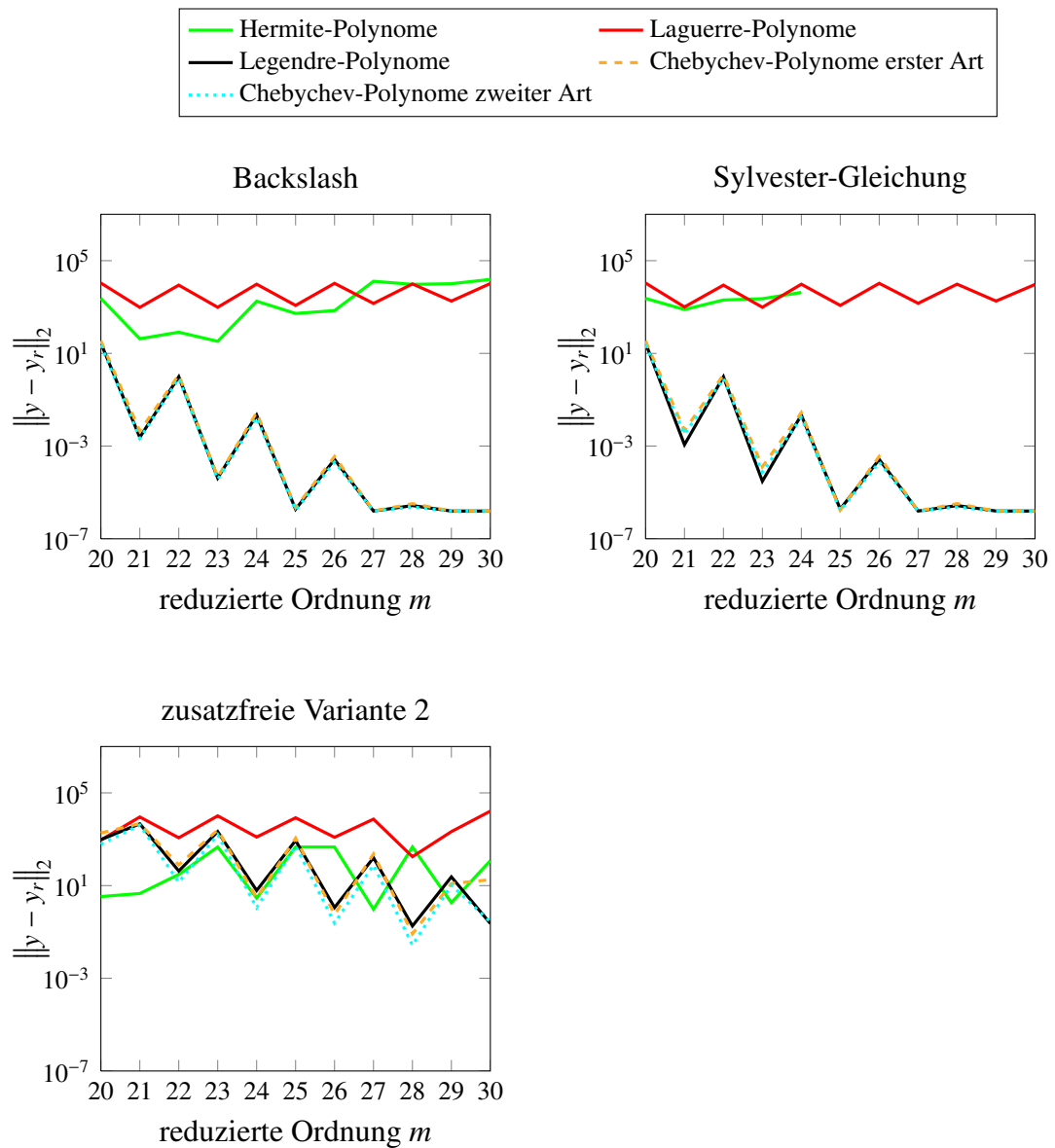


Abbildung 9: Absoluter Fehler durch Reduktion mit Hilfe der verbleibenden Lösungsstrategien (Beispiel 4)

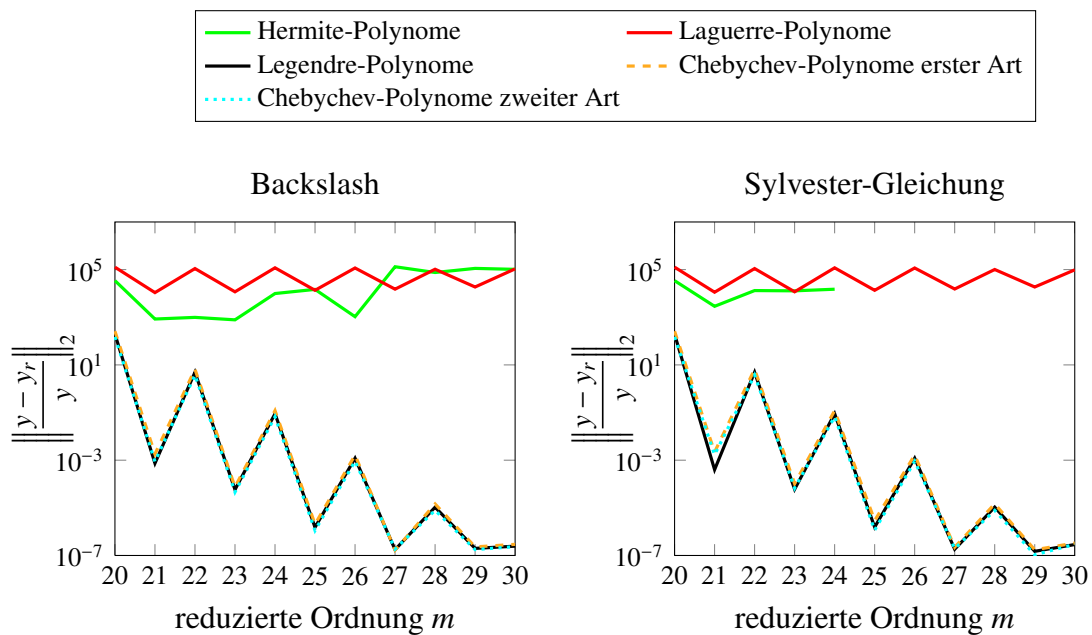


Abbildung 10: Relativer Fehler - Vergleich (Beispiel 4)

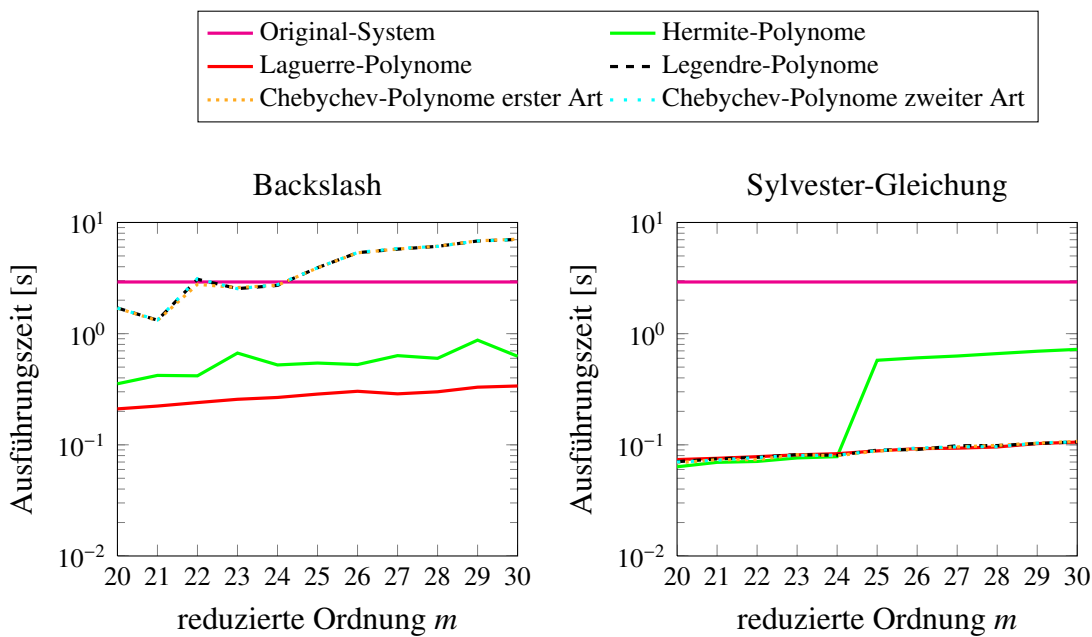


Abbildung 11: Ausführungszeit - Vergleich (Beispiel 4)

Grafiken für $m \in [20, 150]$:

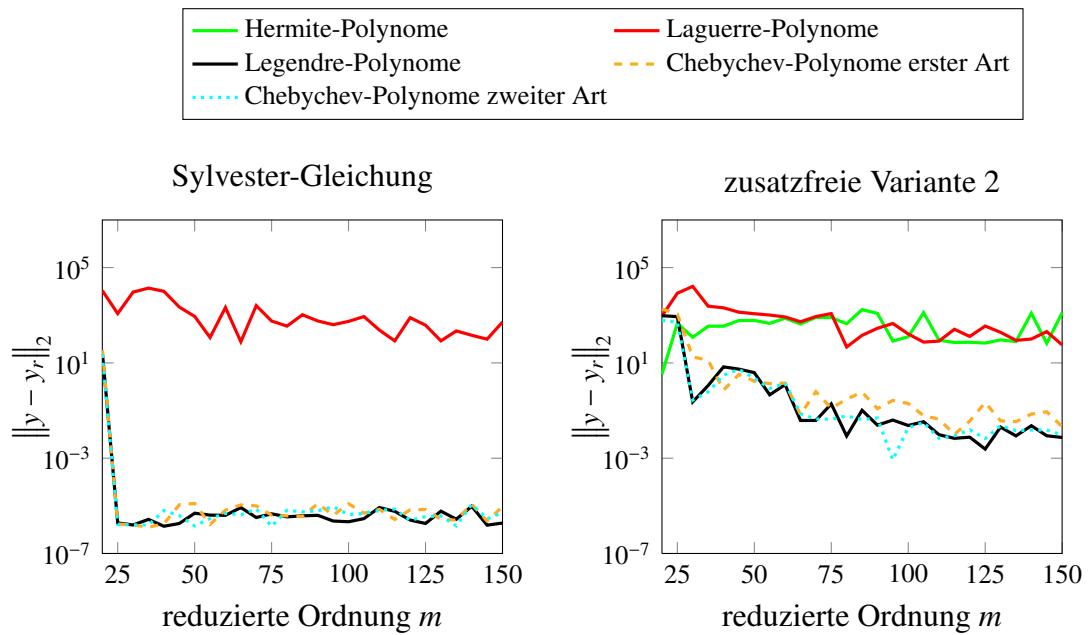


Abbildung 12: Absoluter Fehler für $m = 20, \dots, 150$ durch die verbleibenden Lösungsstrategien (Beispiel 4)

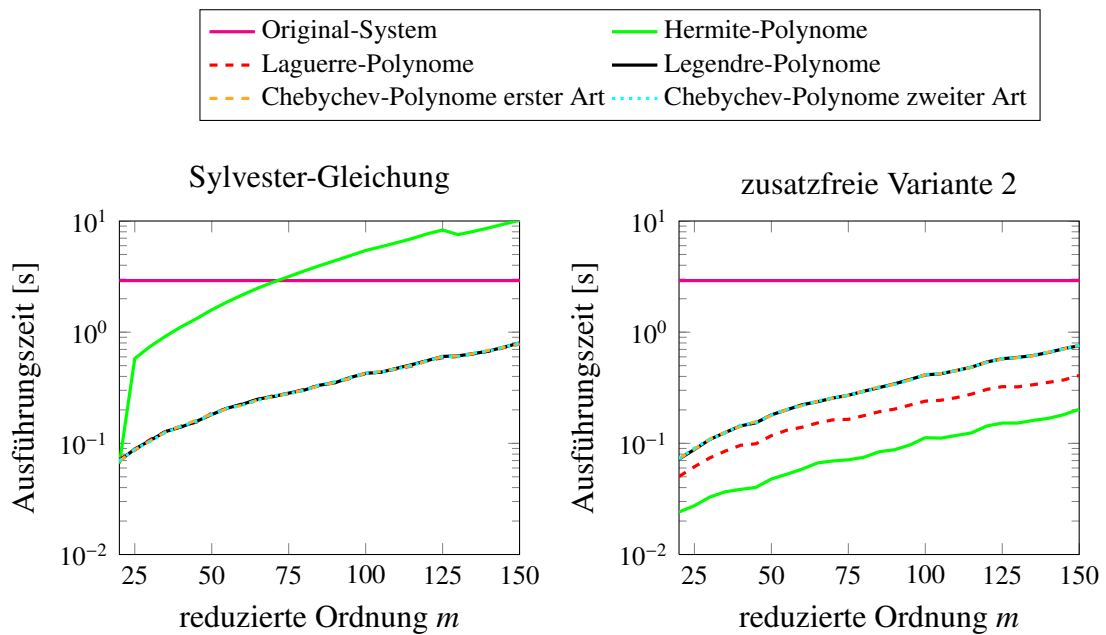


Abbildung 13: Ausführungszeit für $m = 20, \dots, 150$ durch die verbleibenden Lösungsstrategien (Beispiel 4)

B.5 Beispiel 5

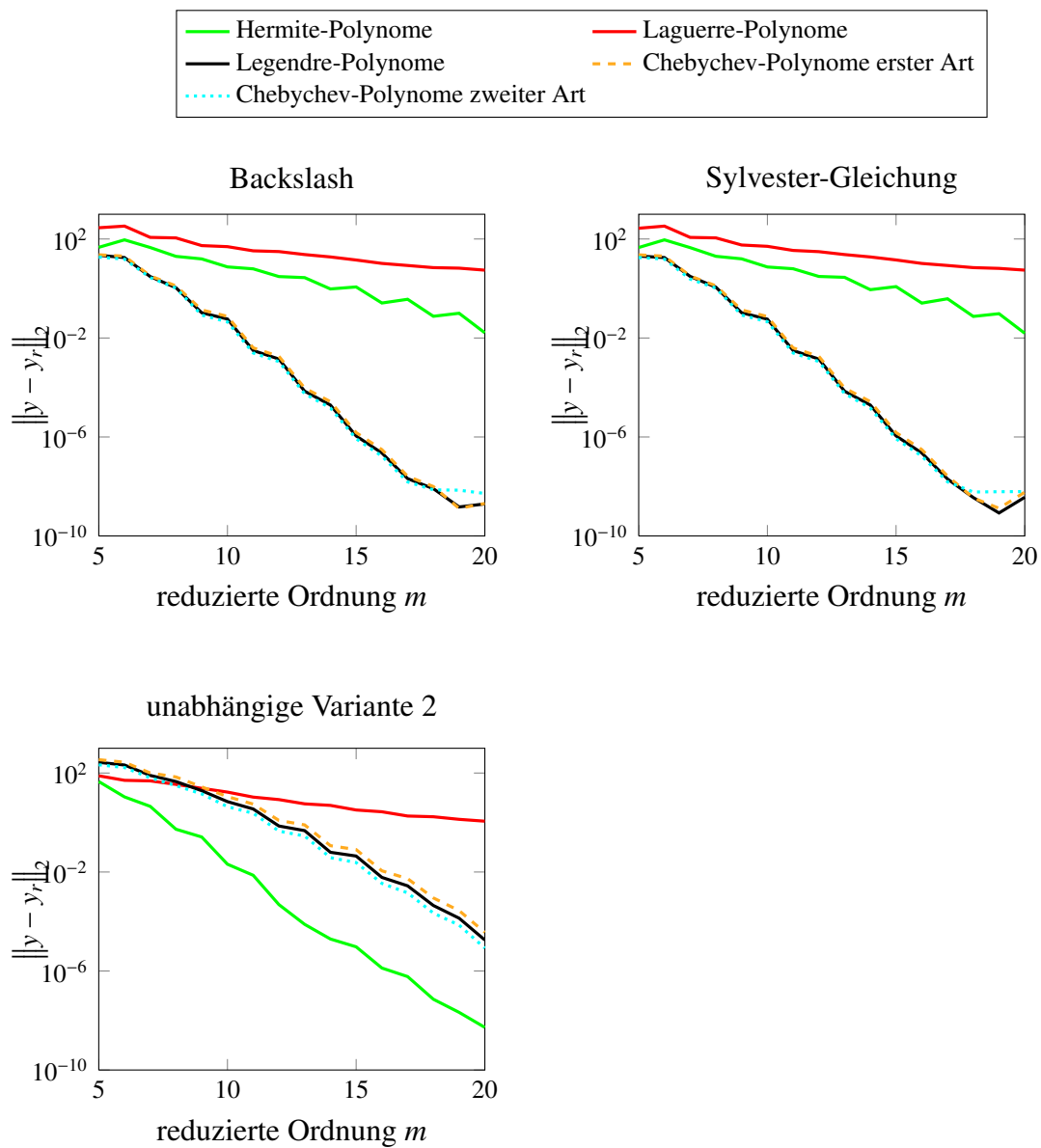


Abbildung 14: Absoluter Fehler durch Reduktion mit Hilfe der verbleibenden Lösungsstrategien (Beispiel 5)

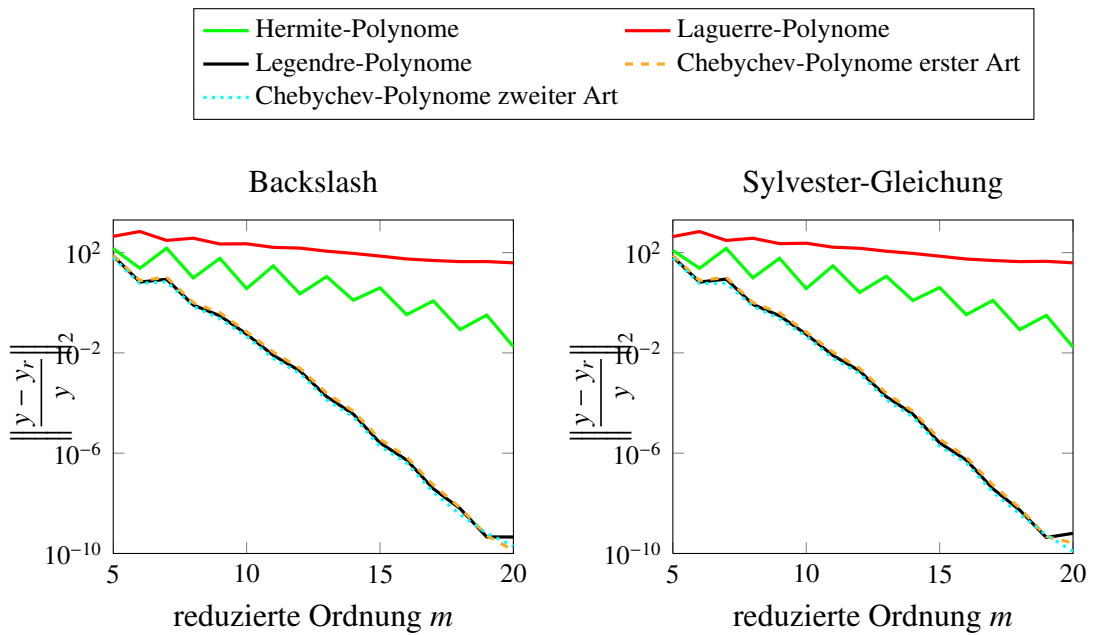


Abbildung 15: Relativer Fehler - Vergleich (Beispiel 5)

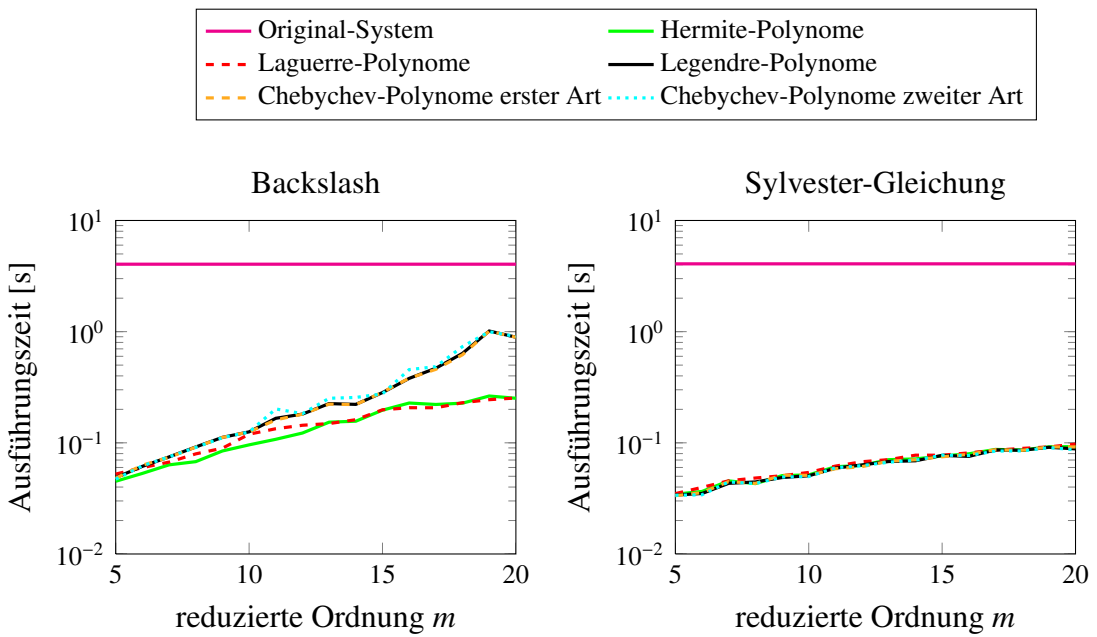


Abbildung 16: Ausführungszeit - Vergleich (Beispiel 5)

C Explizite Darstellung der Inversen von \tilde{E}

Im Folgenden werden die expliziten Inversen zu \tilde{E} und \mathcal{E} für die einzelnen orthogonale Polynome dargestellt.

C.1 Hermite-Polynome

Inverse zu \tilde{E}

$$\tilde{E}^{-1} = \begin{bmatrix} 0 & \dots & \dots & \dots & \dots & 0 & \frac{1}{g_{m-1}(t_0)} \\ -2 & \ddots & & & & \vdots & 2 \frac{g_{m-1}(t_0)}{g_0(t_0)} \\ 0 & -4 & \ddots & & & \vdots & 4 \frac{g_{m-1}(t_0)}{g_1(t_0)} \\ \vdots & \ddots & -6 & \ddots & & \vdots & 6 \frac{g_{m-1}(t_0)}{g_2(t_0)} \\ \vdots & & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & & & \ddots & -2(m-2) & 0 & \frac{2(m-2)g_{m-3}(t_0)}{g_{m-1}(t_0)} \\ 0 & \dots & \dots & \dots & 0 & -2(m-1) & \frac{2(m-1)g_{m-2}(t_0)}{g_{m-1}(t_0)} \end{bmatrix}$$

Inverse zu \mathcal{E}

\mathcal{E} ist bei Verwendung der Hermite-Polynome nicht invertierbar.

C.2 Laguerre-Polynome

Inverse zu \tilde{E}

$$\tilde{E}^{-1} = -\frac{1}{c} \begin{bmatrix} -1 & & -1 & & -1 & & \cdots \\ -(g_1(t_0) + \dots + g_{m-1}(t_0)) & g_0(t_0) & & & g_0(t_0) & & \cdots \\ -(g_2(t_0) + \dots + g_{m-1}(t_0)) & -(g_2(t_0) + \dots + g_{m-1}(t_0)) & & & g_0(t_0) + g_1(t_0) & & \cdots \\ \vdots & \vdots & & & \vdots & & \ddots \\ -(g_{m-3}(t_0) + g_{m-1}(t_0)) & -(g_{m-3}(t_0) + g_{m-1}(t_0)) & & & -(g_{m-3}(t_0) + g_{m-1}(t_0)) & & \cdots \\ -(g_{m-2}(t_0) + g_{m-1}(t_0)) & -(g_{m-2}(t_0) + g_{m-1}(t_0)) & & & -(g_{m-2}(t_0) + \dots + g_{m-1}(t_0)) & & \cdots \\ -g_{m-1}(t_0) & -g_{m-1}(t_0) & & & -g_{m-1}(t_0) & & \cdots \\ \cdots & -1 & & -1 & & -1 & \\ \cdots & g_0(t_0) & & g_0(t_0) & & g_0(t_0) & \\ \cdots & g_0(t_0) + g_1(t_0) & & g_0(t_0) + g_1(t_0) & & g_0(t_0) + g_1(t_0) & \\ \ddots & \vdots & & \vdots & & \vdots & \\ \cdots & g_0(t_0) + \dots + g_{m-4}(t_0) & & g_0(t_0) + \dots + g_{m-4}(t_0) & & g_0(t_0) + \dots + g_{m-4}(t_0) & \\ \cdots & -(g_{m-2}(t_0) + g_{m-1}(t_0)) & & g_0(t_0) + \dots + g_{m-3}(t_0) & & g_0(t_0) + \dots + g_{m-2}(t_0) & \\ \cdots & -g_{m-1}(t_0) & & -g_{m-1}(t_0) & & g_0(t_0) + \dots + g_{m-2}(t_0) & \end{bmatrix},$$

wobei $c = g_0(t_0) + \dots + g_{m-1}(t_0)$.

Inverse zu \mathcal{E}

$$\mathcal{E}^{-1} = \begin{bmatrix} -1 & \cdots & \cdots & -1 \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & -1 \end{bmatrix}$$

C.3 Struktur der Inversen für Legendre- und Chebyshev-Polynome

Für die Legendre- und Chebyshev-Polynome hat die Matrix \tilde{E} dieselbe Struktur

$$\tilde{E} = \begin{bmatrix} g_0(t_0) & -\frac{g_0(t)}{\dot{g}_1(t)} & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ g_1(t_0) & 0 & -\alpha_1 & \ddots & & & & \vdots \\ g_2(t_0) & -\gamma_2 & 0 & -\alpha_2 & \ddots & & & \vdots \\ g_3(t_0) & 0 & -\gamma_3 & 0 & -\alpha_3 & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ g_{m-3}(t_0) & \vdots & & \ddots & -\gamma_{m-3} & 0 & -\alpha_{m-3} & 0 \\ g_{m-2}(t_0) & \vdots & & & \ddots & -\gamma_{m-2} & 0 & -\alpha_{m-2} \\ g_{m-1}(t_0) & 0 & \cdots & \cdots & \cdots & 0 & -\gamma_{m-1} & 0 \end{bmatrix}.$$

Daher wurden die Inversen zu \tilde{E} zunächst in allgemeiner Form bestimmt. Für ein gerades m hat \tilde{E}^{-1} folgende Struktur:

$$\tilde{E}^{-1} = \begin{bmatrix} 0 & f_1(a_1) & 0 \\ -(-1)^{\frac{0}{2}} \frac{\dot{g}_1(t)}{g_0(t)} & f_2(a_1) & 0 \\ 0 & f_3(a_1) - \frac{1}{\alpha_1} & 0 \\ -(-1)^{\frac{2}{2}} \frac{\dot{g}_1(t) \gamma_2}{g_0(t) \alpha_2} & f_4(a_1) & -(-1)^{\frac{0}{2}} \frac{1}{\alpha_2} \\ 0 & f_5(a_1) - (-1)^{\frac{2}{2}} \frac{\gamma_3}{\alpha_1 \alpha_3} & 0 \\ \vdots & \vdots & \vdots \\ 0 & f_{m-3}(a_1) - (-1)^{\frac{m-6}{2}} \frac{\gamma_3 \gamma_5 \cdots \gamma_{m-5}}{\alpha_1 \alpha_3 \cdots \alpha_{m-5}} & 0 \\ -(-1)^{\frac{m-4}{2}} \frac{\dot{g}_1(t) \gamma_2 \gamma_4 \cdots \gamma_{m-4}}{g_0(t) \alpha_2 \cdot \alpha_4 \cdots \alpha_{m-4}} & f_{m-2}(a_1) & -(-1)^{\frac{m-6}{2}} \frac{\gamma_4 \gamma_6 \cdots \gamma_{m-4}}{\alpha_2 \cdot \alpha_4 \cdots \alpha_{m-4}} \\ 0 & f_{m-1}(a_1) - (-1)^{\frac{m-4}{2}} \frac{\gamma_3 \gamma_5 \cdots \gamma_{m-3}}{\alpha_1 \alpha_3 \cdots \alpha_{m-3}} & 0 \\ -(-1)^{\frac{m-2}{2}} \frac{\dot{g}_1(t) \gamma_2 \gamma_4 \cdots \gamma_{m-2}}{g_0(t) \alpha_2 \cdot \alpha_4 \cdots \alpha_{m-2}} & f_m(a_1) & -(-1)^{\frac{m-4}{2}} \frac{\gamma_4 \cdots \gamma_{m-2}}{\alpha_2 \alpha_4 \cdots \alpha_{m-2}} \end{bmatrix}$$

$$\begin{array}{ccccccc}
f_1(a_3) & \cdots & 0 & f_1(a_{m-3}) & 0 & f_1(a_{m-1}) \\
f_2(a_3) & \cdots & 0 & f_2(a_{m-3}) & 0 & f_2(a_{m-1}) \\
f_3(a_3) & \cdots & 0 & f_3(a_{m-3}) & 0 & f_3(a_{m-1}) \\
f_4(a_3) & \cdots & 0 & f_4(a_{m-3}) & 0 & f_4(a_{m-1}) \\
f_5(a_3) - \frac{1}{\alpha_3} & \cdots & 0 & f_5(a_{m-3}) & 0 & f_5(a_{m-1}) \\
\vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
f_{m-3}(a_3) - (-1)^{\frac{m-8}{2}} \frac{\gamma_5 \gamma_7 \cdots \gamma_{m-5}}{\alpha_3 \alpha_5 \cdots \alpha_{m-5}} & \cdots & 0 & f_{m-3}(a_{m-3}) & 0 & f_{m-3}(a_{m-1}) \\
(-1)^{\frac{m-4}{2}} f_{m-2}(a_3) & \cdots & -\frac{1}{\alpha_{m-4}} & f_{m-2}(a_{m-3}) & 0 & f_{m-2}(a_{m-1}) \\
f_{m-1}(a_3) - (-1)^{\frac{m-6}{2}} \frac{\gamma_5 \gamma_7 \cdots \gamma_{m-3}}{\alpha_3 \alpha_5 \cdots \alpha_{m-3}} & \cdots & 0 & f_{m-1}(a_{m-3}) - \frac{1}{\alpha_{m-3}} & 0 & f_{m-1}(a_{m-1}) \\
f_m(a_3) & \cdots & -(-1)^{\frac{2}{2}} \frac{\gamma_{m-2}}{\alpha_{m-4} \alpha_{m-2}} & f_m(a_{m-3}) & -\frac{1}{\alpha_{m-2}} & f_m(a_{m-1})
\end{array}$$

wobei

$$\begin{aligned}
c &= \underbrace{(-1)^{\frac{m-2}{2}} \gamma_3 \gamma_5 \cdots \gamma_{m-1}}_{=: a_1} g_1(t_0) + \underbrace{(-1)^{\frac{m-4}{2}} \alpha_1 \gamma_5 \gamma_7 \cdots \gamma_{m-1}}_{=: a_3} g_3(t_0) + \cdots \\
&\quad \cdots + \underbrace{(-1)^{\frac{2}{2}} \alpha_1 \alpha_3 \cdots \alpha_{m-5} \gamma_{m-1}}_{=: a_{m-3}} g_{m-3}(t_0) + \underbrace{\alpha_1 \alpha_3 \cdots \alpha_{m-3}}_{=: a_{m-1}} g_{m-1}(t_0), \\
f_1(x) &= \frac{x}{c}, \\
f_2(x) &= \dot{g}_1(t) f_1(x), \\
f_3(x) &= \frac{1}{\alpha_1} g_1(t_0) f_1(x), \\
f_4(x) &= \frac{1}{\alpha_2} (g_2(t_0) f_1(x) - \gamma_2 f_2(x)), \\
f_5(x) &= \frac{1}{\alpha_3} (g_3(t_0) f_1(x) - \gamma_3 f_3(x)), \\
&\quad \vdots \\
f_{m-1}(x) &= \frac{1}{\alpha_{m-3}} (g_{m-3}(t_0) f_1(x) - \gamma_{m-3} f_{m-3}(x)), \\
f_m(x) &= \frac{1}{\alpha_{m-2}} (g_{m-2}(t_0) f_1(x) - \gamma_{m-2} f_{m-2}(x)).
\end{aligned}$$

Für ein ungerades m ist die Darstellung der Inversen:

$$\tilde{E}^{-1} = \left[\begin{array}{ccc} \frac{\dot{g}_1(t)}{g_0(t)} f_1(a_0) & 0 & f_1(a_2) \\ \frac{\dot{g}_1(t)}{g_0(t)} (f_2(a_0) - 1) & 0 & f_2(a_2) \\ \frac{\dot{g}_1(t)}{g_0(t)} f_3(a_0) & -(-1)^{\frac{0}{2}} \frac{1}{\alpha_1} & f_3(a_2) \\ \frac{\dot{g}_1(t)}{g_0(t)} \left(f_4(a_0) - (-1)^{\frac{2}{2}} \frac{\gamma_2}{\alpha_2} \right) & 0 & f_4(a_2) - \frac{1}{\alpha_2} \\ \frac{\dot{g}_1(t)}{g_0(t)} f_5(a_0) & -(-1)^{\frac{2}{2}} \frac{\gamma_3}{\alpha_1 \alpha_3} & f_5(a_2) \\ \vdots & \vdots & \vdots \\ \frac{\dot{g}_1(t)}{g_0(t)} \left(f_{m-3}(a_0) - (-1)^{\frac{m-5}{2}} \frac{\gamma_2 \gamma_4 \cdots \gamma_{m-5}}{\alpha_2 \cdot \alpha_4 \cdots \alpha_{m-5}} \right) & 0 & f_{m-3}(a_2) - (-1)^{\frac{m-7}{2}} \frac{\gamma_4 \gamma_6 \cdots \gamma_{m-5}}{\alpha_2 \alpha_4 \cdots \alpha_{m-5}} \\ \frac{\dot{g}_1(t)}{g_0(t)} f_{m-2}(a_0) & -(-1)^{\frac{m-5}{2}} \frac{\gamma_3 \gamma_5 \cdots \gamma_{m-4}}{\alpha_1 \alpha_3 \cdots \alpha_{m-4}} & f_{m-2}(a_2) \\ \frac{\dot{g}_1(t)}{g_0(t)} \left(f_{m-1}(a_0) - (-1)^{\frac{m-3}{2}} \frac{\gamma_2 \gamma_4 \cdots \gamma_{m-3}}{\alpha_2 \cdot \alpha_4 \cdots \alpha_{m-3}} \right) & 0 & f_{m-1}(a_2) - (-1)^{\frac{m-5}{2}} \frac{\gamma_4 \gamma_6 \cdots \gamma_{m-3}}{\alpha_2 \alpha_4 \cdots \alpha_{m-3}} \\ \frac{\dot{g}_1(t)}{g_0(t)} f_m(a_0) & -(-1)^{\frac{m-3}{2}} \frac{\gamma_3 \gamma_5 \cdots \gamma_{m-2}}{\alpha_1 \alpha_3 \cdots \alpha_{m-2}} & f_m(a_2) \end{array} \right]$$

$$\left[\begin{array}{cccccc} 0 & \cdots & 0 & f_1(a_{m-3}) & 0 & f_1(a_{m-1}) \\ 0 & \cdots & 0 & f_2(a_{m-3}) & 0 & f_2(a_{m-1}) \\ 0 & \cdots & 0 & f_3(a_{m-3}) & 0 & f_3(a_{m-1}) \\ 0 & \cdots & 0 & f_4(a_{m-3}) & 0 & f_4(a_{m-1}) \\ -(-1)^{\frac{0}{2}} \frac{1}{\alpha_3} & \cdots & 0 & f_5(a_{m-3}) & 0 & f_5(a_{m-1}) \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & f_{m-3}(a_{m-3}) & 0 & f_{m-3}(a_{m-1}) \\ -(-1)^{\frac{m-7}{2}} \frac{\gamma_5 \gamma_7 \cdots \gamma_{m-4}}{\alpha_3 \alpha_5 \cdots \alpha_{m-4}} & \cdots & -\frac{1}{\alpha_{m-4}} & f_{m-2}(a_{m-3}) & 0 & f_{m-2}(a_{m-1}) \\ 0 & \cdots & 0 & f_{m-1}(a_{m-3}) - \frac{1}{\alpha_{m-3}} & 0 & f_{m-1}(a_{m-1}) \\ -(-1)^{\frac{m-5}{2}} \frac{\gamma_5 \gamma_7 \cdots \gamma_{m-2}}{\alpha_3 \alpha_5 \cdots \alpha_{m-2}} & \cdots & -(-1)^{\frac{2}{2}} \frac{\gamma_{m-2}}{\alpha_{m-4} \alpha_{m-2}} & f_m(a_{m-3}) & -\frac{1}{\alpha_{m-2}} & f_m(a_{m-1}) \end{array} \right],$$

wobei

$$\begin{aligned}
c = & \underbrace{(-1)^{\frac{m-1}{2}} \gamma_2 \gamma_4 \cdots \gamma_{m-1}}_{=: a_0} \dot{g}_1(t) + \underbrace{(-1)^{\frac{m-3}{2}} \gamma_4 \gamma_6 \cdots \gamma_{m-1}}_{=: a_2} g_2(t_0) + \underbrace{(-1)^{\frac{m-5}{2}} \alpha_2 \gamma_6 \gamma_8 \cdots \gamma_{m-1}}_{=: a_4} g_4(t_0) + \cdots \\
& \cdots + \underbrace{(-1)^{\frac{2}{2}} \alpha_2 \alpha_4 \cdots \alpha_{m-5} \gamma_{m-1}}_{=: a_{m-3}} g_{m-3}(t_0) + \underbrace{\alpha_2 \alpha_4 \cdots \alpha_{m-3}}_{=: a_{m-1}} g_{m-1}(t_0)
\end{aligned}$$

und $f_1(x), \dots, f_m(x)$ wie bei der Inversen zu einem geradem m definiert sind.

C.4 Legendre-Polynome

Im Falle eines geraden m sieht die explizite Inverse wie folgt aus:

Inverse zu \tilde{E}

$$\tilde{E} = \begin{bmatrix} 0 & \frac{a_1}{c} & 0 \\ -1 & \frac{a_1}{c} & 0 \\ 0 & 3g_1(t_0) \frac{a_1}{c} - 3 & 0 \\ -1 & \frac{a_1}{c} (5g_2(t_0) + 1) & -5 \\ 0 & \frac{a_1}{c} (7g_4(t_0) + 3g_1(t_0)) - 3 & 0 \\ \vdots & \vdots & \vdots \\ 0 & \frac{a_1}{c} [(2(m-5)+1)g_{m-5}(t_0) + (2(m-7)+1)g_{m-7}(t_0) + \dots + 3g_1(t_0)] - 3 & 0 \\ -1 & \frac{a_1}{c} [(2(m-4)+1)g_{m-4}(t_0) + (2(m-6)+1)g_{m-6}(t_0) + \dots + 5g_2(t_0) + 1] & -5 \\ 0 & \frac{a_1}{c} [(2(m-3)+1)g_{m-3}(t_0) + (2(m-5)+1)g_{m-5}(t_0) + \dots + 3g_1(t_0)] - 3 & 0 \\ -1 & \frac{a_1}{c} [(2(m-2)+1)g_{m-2}(t_0) + (2(m-4)+1)g_{m-4}(t_0) + \dots + 5g_2(t_0) + 1] & -5 \end{bmatrix}$$

$$\begin{array}{rcccl}
& \frac{a_3}{c} & \cdots & 0 \\
& \frac{a_3}{c} & \cdots & 0 \\
& 3g_1(t_0)\frac{a_3}{c} & \cdots & 0 \\
& \frac{a_3}{c}(5g_2(t_0) + 1) & \cdots & 0 \\
& \frac{a_3}{c}(7g_4(t_0) + 3g_1(t_0)) - 7 & \cdots & 0 \\
& \vdots & \ddots & \vdots \\
\frac{a_3}{c} [(2(m-5) + 1)g_{m-5}(t_0) + (2(m-7) + 1)g_{m-7}(t_0) + \dots + 3g_1(t_0)] - 7 & \cdots & 0 \\
\frac{a_3}{c} [(2(m-4) + 1)g_{m-4}(t_0) + (2(m-6) + 1)g_{m-6}(t_0) + \dots + 5g_2(t_0) + 1] & \cdots & -(2(m-4) + 1) \\
\frac{a_3}{c} [(2(m-3) + 1)g_{m-3}(t_0) + (2(m-5) + 1)g_{m-5}(t_0) + \dots + 3g_1(t_0)] - 7 & \cdots & 0 \\
\frac{a_3}{c} [(2(m-2) + 1)g_{m-2}(t_0) + (2(m-4) + 1)g_{m-4}(t_0) + \dots + 5g_2(t_0) + 1] & \cdots & -(2(m-4) + 1) \\
& & & \\
& \frac{a_{m-3}}{c} & & 0 \\
& \frac{a_{m-3}}{c} & & 0 \\
& 3g_1(t_0)\frac{a_{m-3}}{c} & & 0 \\
& \frac{a_{m-3}}{c}(5g_2(t_0) + 1) & & 0 \\
& \frac{a_{m-3}}{c}(7g_4(t_0) + 3g_1(t_0)) - 7 & & 0 \\
& \vdots & & \vdots \\
\frac{a_{m-3}}{c} [(2(m-5) + 1)g_{m-5}(t_0) + (2(m-7) + 1)g_{m-7}(t_0) + \dots + 3g_1(t_0)] - (2(m-3) + 1) & & 0 \\
& \frac{a_{m-3}}{c} [(2(m-4) + 1)g_{m-4}(t_0) + (2(m-6) + 1)g_{m-6}(t_0) + \dots + 5g_2(t_0) + 1] & & 0 \\
\frac{a_{m-3}}{c} [(2(m-3) + 1)g_{m-3}(t_0) + (2(m-5) + 1)g_{m-5}(t_0) + \dots + 3g_1(t_0)] - (2(m-3) + 1) & & 0 \\
& \frac{a_{m-3}}{c} [(2(m-2) + 1)g_{m-2}(t_0) + (2(m-4) + 1)g_{m-4}(t_0) + \dots + 5g_2(t_0) + 1] & & -(2(m-2) + 1)
\end{array}$$

$$\begin{array}{c}
\frac{a_{m-1}}{c} \\
\frac{c}{a_{m-1}} \\
\frac{c}{a_{m-1}} \\
3g_1(t_0) \frac{a_{m-1}}{c} \\
\frac{a_{m-1}}{c} (5g_2(t_0) + 1) \\
\frac{a_{m-1}}{c} (7g_4(t_0) + 3g_1(t_0)) - 7 \\
\vdots \\
\frac{a_{m-1}}{c} [(2(m-5) + 1)g_{m-5}(t_0) + (2(m-7) + 1)g_{m-7}(t_0) + \dots + 3g_1(t_0)] \\
\frac{a_{m-1}}{c} [(2(m-4) + 1)g_{m-4}(t_0) + (2(m-6) + 1)g_{m-6}(t_0) + \dots + 5g_2(t_0) + 1] \\
\frac{a_{m-1}}{c} [(2(m-3) + 1)g_{m-3}(t_0) + (2(m-5) + 1)g_{m-5}(t_0) + \dots + 3g_1(t_0)] \\
\frac{a_{m-1}}{c} [(2(m-2) + 1)g_{m-2}(t_0) + (2(m-4) + 1)g_{m-4}(t_0) + \dots + 5g_2(t_0) + 1]
\end{array}$$

wobei

$$\begin{aligned}
c &= a_1 g_1(t_0) + a_3 g_3(t_0) + \dots + a_{m-3} g_{m-3}(t_0) + a_{m-1} g_{m-1}(t_0), \\
a_1 &= \frac{1}{7 \cdot 11 \cdot \dots \cdot (2(m-1) + 1)}, \\
a_3 &= \frac{1}{3 \cdot 11 \cdot 15 \cdot \dots \cdot (2(m-1) + 1)}, \\
&\vdots \\
a_{m-3} &= \frac{1}{3 \cdot 7 \cdot \dots \cdot (2(m-5) + 1) \cdot (2(m-1) + 1)}, \\
a_{m-1} &= \frac{1}{3 \cdot 7 \cdot \dots \cdot (2(m-1) + 1)}.
\end{aligned}$$

Andernfalls ist die Inverse der Legendre-Polynome:

$$\tilde{E}^{-1} = \begin{array}{r} \left[\begin{array}{ccc} \frac{a_0}{c} & & 0 \\ \frac{a_0}{c} - 1 & & 0 \\ 3g_1(t_0)\frac{a_0}{c} & & -3 \\ \frac{a_0}{c}(5g_2(t_0) + 1) - 1 & & 0 \\ \frac{a_0}{c}(7g_3(t_0) + 3g_1(t_0)) & & -3 \\ \vdots & & \vdots \\ \frac{a_0}{c} [(2(m-5) + 1)g_{m-5}(t_0) + (2(m-7) + 1)g_{m-7}(t_0) + \dots + 5g_2(t_0) + 1] - 1 & & 0 \\ \frac{a_0}{c} [(2(m-4) + 1)g_{m-4}(t_0) + (2(m-6) + 1)g_{m-6}(t_0) + \dots + 7g_3(t_0) + 3g_1(t_0)] & & -3 \\ \frac{a_0}{c} [(2(m-3) + 1)g_{m-3}(t_0) + (2(m-5) + 1)g_{m-5}(t_0) + \dots + 5g_2(t_0) + 1] - 1 & & 0 \\ \frac{a_0}{c} [(2(m-2) + 1)g_{m-2}(t_0) + (2(m-4) + 1)g_{m-4}(t_0) + \dots + 7g_3(t_0) + 3g_1(t_0)] & & -3 \end{array} \right. \\ \\ \left[\begin{array}{ccc} \frac{a_2}{c} & & 0 \dots \\ \frac{a_2}{c} & & 0 \dots \\ 3g_1(t_0)\frac{a_2}{c} & & 0 \dots \\ \frac{a_2}{c}(5g_2(t_0) + 1) - 5 & & 0 \dots \\ \frac{a_2}{c}(7g_3(t_0) + 3g_1(t_0)) & & -7 \dots \\ \vdots & & \vdots \ddots \\ \frac{a_2}{c} [(2(m-5) + 1)g_{m-5}(t_0) + (2(m-7) + 1)g_{m-7}(t_0) + \dots + 5g_2(t_0) + 1] - 5 & & 0 \dots \\ \frac{a_2}{c} [(2(m-4) + 1)g_{m-4}(t_0) + (2(m-6) + 1)g_{m-6}(t_0) + \dots + 7g_3(t_0) + 3g_1(t_0)] & & -7 \dots \\ \frac{a_2}{c} [(2(m-3) + 1)g_{m-3}(t_0) + (2(m-5) + 1)g_{m-5}(t_0) + \dots + 5g_2(t_0) + 1] - 5 & & 0 \dots \\ \frac{a_2}{c} [(2(m-2) + 1)g_{m-2}(t_0) + (2(m-4) + 1)g_{m-4}(t_0) + \dots + 7g_3(t_0) + 3g_1(t_0)] & & -7 \dots \end{array} \right. \end{array}$$

$$\begin{array}{rcl}
\dots & 0 & \frac{a_{m-3}}{c} \\
\dots & 0 & \frac{a_{m-3}}{c} \\
\dots & 0 & 3g_1(t_0) \frac{a_{m-3}}{c} \\
\dots & 0 & \frac{a_{m-3}}{c} (5g_2(t_0) + 1) \\
\dots & 0 & \frac{a_{m-3}}{c} (7g_3(t_0) + 3g_1(t_0)) \\
\ddots & \vdots & \vdots \\
\dots & 0 & \frac{a_{m-3}}{c} [(2(m-5) + 1)g_{m-5}(t_0) + (2(m-7) + 1)g_{m-7}(t_0) + \dots + 5g_2(t_0) + 1] \\
\dots & -(2(m-4) + 1) & \frac{a_{m-3}}{c} [(2(m-4) + 1)g_{m-4}(t_0) + (2(m-6) + 1)g_{m-6}(t_0) + \dots + 7g_3(t_0) + 3g_1(t_0)] \\
\dots & 0 & \frac{a_{m-3}}{c} [(2(m-3) + 1)g_{m-3}(t_0) + (2(m-5) + 1)g_{m-5}(t_0) + \dots + 5g_2(t_0) + 1] - (2(m-3) + 1) \\
\dots & -(2(m-4) + 1) & \frac{a_{m-3}}{c} [(2(m-2) + 1)g_{m-2}(t_0) + (2(m-4) + 1)g_{m-4}(t_0) + \dots + 7g_3(t_0) + 3g_1(t_0)]
\end{array}$$

$$\begin{array}{rcl}
0 & \frac{a_{m-1}}{c} & \\
0 & \frac{a_{m-1}}{c} & \\
0 & 3g_1(t_0) \frac{a_{m-1}}{c} & \\
0 & \frac{a_{m-1}}{c} (5g_2(t_0) + 1) & \\
0 & \frac{a_{m-1}}{c} (7g_3(t_0) + 3g_1(t_0)) & \\
\vdots & \vdots & \\
0 & \frac{a_{m-1}}{c} [(2(m-5) + 1)g_{m-5}(t_0) + (2(m-7) + 1)g_{m-7}(t_0) + \dots + 5g_2(t_0) + 1] & \\
0 & \frac{a_{m-1}}{c} [(2(m-4) + 1)g_{m-4}(t_0) + (2(m-6) + 1)g_{m-6}(t_0) + \dots + 7g_3(t_0) + 3g_1(t_0)] & \\
0 & \frac{a_{m-1}}{c} [(2(m-3) + 1)g_{m-3}(t_0) + (2(m-5) + 1)g_{m-5}(t_0) + \dots + 5g_2(t_0) + 1] & \\
-(2(m-2) + 1) & \frac{a_{m-1}}{c} [(2(m-2) + 1)g_{m-2}(t_0) + (2(m-4) + 1)g_{m-4}(t_0) + \dots + 7g_3(t_0) + 3g_1(t_0)] &
\end{array}$$

wobei

$$\begin{aligned}
c &= a_0 + a_2g_2(t_0) + a_4g_4(t_0) + \dots + a_{m-3}g_{m-3}(t_0) + a_{m-1}g_{m-1}(t_0), \\
a_0 &= \frac{1}{5 \cdot 9 \cdot \dots \cdot (2(m-1) + 1)}, \\
a_2 &= \frac{1}{9 \cdot 13 \cdot \dots \cdot (2(m-1) + 1)},
\end{aligned}$$

$$\begin{aligned}
 a_4 &= \frac{1}{5 \cdot 13 \cdot 17 \cdot \dots \cdot (2(m-1) + 1)}, \\
 &\quad \vdots \\
 a_{m-3} &= \frac{1}{5 \cdot 9 \cdot \dots \cdot (2(m-5) + 1) \cdot (2(m-1) + 1)}, \\
 a_{m-1} &= \frac{1}{5 \cdot 9 \cdot \dots \cdot (2(m-3) + 1)}.
 \end{aligned}$$

Inverse zu \mathcal{E}

Für ein gerades m ist die Inverse zu \mathcal{E} :

$$\mathcal{E}^{-1} = \begin{bmatrix}
 0 & 5 & 0 & 9 & \dots & 0 & 2(m-2) + 1 & 0 & 2m + 1 \\
 -3 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 9 & \dots & 0 & 2(m-2) + 1 & 0 & 2m + 1 \\
 -3 & 0 & -7 & 0 & \dots & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & \dots & 0 & 2(m-2) + 1 & 0 & 2m + 1 \\
 \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
 0 & 0 & 0 & 0 & \dots & 0 & 2(m-2) + 1 & 0 & 2m + 1 \\
 -3 & 0 & -7 & 0 & \dots & -2(m-3) - 1 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 2m + 1 \\
 -3 & 0 & -7 & 0 & \dots & -2(m-3) - 1 & 0 & -2(m-1) - 1 & 0
 \end{bmatrix}.$$

C.5 Chebychev Polynome erster Art

Inverse zu \tilde{E}

Ist m eine gerade Zahl, so sieht die Inverse zum Chebychev-Polynom wie folgt aus:

$$\tilde{E} = \begin{bmatrix} 0 & \frac{a_1}{c} & 0 \\ -1 & \frac{a_1}{c} & 0 \\ 0 & 4 \left(g_1(t_0) \frac{a_1}{c} - 1 \right) & 0 \\ -3 & 6 \frac{a_1}{c} \left(g_2(t_0) + \frac{1}{2} \right) & -6 \\ 0 & 8 \left[\frac{a_1}{c} (g_3(t_0) + g_1(t_0)) - 1 \right] & 0 \\ \vdots & \vdots & \vdots \\ 0 & 2(m-4) \left[\frac{a_1}{c} (g_{m-5}(t_0) + g_{m-7}(t_0) + \dots + g_1(t_0)) - 1 \right] & 0 \\ -(m-3) & 2(m-3) \frac{a_1}{c} \left[g_{m-4}(t_0) + g_{m-6}(t_0) + \dots + g_2(t_0) + \frac{1}{2} \right] & -2(m-3) \\ 0 & 2(m-2) \left[\frac{a_1}{c} (g_{m-3}(t_0) + g_{m-5}(t_0) + \dots + g_1(t_0)) - 1 \right] & 0 \\ -(m-1) & 2(m-1) \frac{a_1}{c} \left[g_{m-2}(t_0) + g_{m-4}(t_0) + \dots + g_2(t_0) + \frac{1}{2} \right] & -2(m-1) \end{bmatrix}$$

$$\begin{bmatrix} \frac{a_1}{c} & \dots & 0 \\ \frac{a_1}{c} & \dots & 0 \\ \frac{a_1}{c} & \dots & 0 \\ 4g_1(t_0) \frac{a_1}{c} & \dots & 0 \\ 6 \frac{a_1}{c} \left(g_2(t_0) + \frac{1}{2} \right) & \dots & 0 \\ 8 \left[\frac{a_1}{c} (g_3(t_0) + g_1(t_0)) - 1 \right] & \dots & 0 \\ \vdots & \ddots & \vdots \\ 2(m-4) \left[\frac{a_1}{c} (g_{m-5}(t_0) + g_{m-7}(t_0) + \dots + g_1(t_0)) - 1 \right] & \dots & 0 \\ 2(m-3) \frac{a_1}{c} \left[g_{m-4}(t_0) + g_{m-6}(t_0) + \dots + g_2(t_0) + \frac{1}{2} \right] & \dots & -2(m-3) \\ 2(m-2) \left[\frac{a_1}{c} (g_{m-3}(t_0) + g_{m-5}(t_0) + \dots + g_1(t_0)) - 1 \right] & \dots & 0 \\ 2(m-1) \frac{a_1}{c} \left[g_{m-2}(t_0) + g_{m-4}(t_0) + \dots + g_2(t_0) + \frac{1}{2} \right] & \dots & -2(m-1) \end{bmatrix}$$

$$\begin{array}{r}
\frac{a_1}{c} \\
\frac{a_1}{c} \\
4g_1(t_0)\frac{a_1}{c} \\
6\frac{a_1}{c}\left(g_2(t_0) + \frac{1}{2}\right) \\
8\left[\frac{a_1}{c}(g_3(t_0) + g_1(t_0)) - 1\right] \\
\vdots \\
2(m-4)\left[\frac{a_1}{c}(g_{m-5}(t_0) + g_{m-7}(t_0) + \dots + g_1(t_0)) - 1\right] \\
2(m-3)\frac{a_1}{c}\left[g_{m-4}(t_0) + g_{m-6}(t_0) + \dots + g_2(t_0) + \frac{1}{2}\right] \\
2(m-2)\left[\frac{a_1}{c}(g_{m-3}(t_0) + g_{m-5}(t_0) + \dots + g_1(t_0)) - 1\right] \\
2(m-1)\frac{a_1}{c}\left[g_{m-2}(t_0) + g_{m-4}(t_0) + \dots + g_2(t_0) + \frac{1}{2}\right]
\end{array}
\begin{array}{r}
0 \\
0 \\
0 \\
0 \\
0 \\
\vdots \\
0 \\
-2(m-3) \\
0 \\
-2(m-1)
\end{array}$$

$$\left. \begin{array}{r}
\frac{a_1}{c} \\
\frac{a_1}{c} \\
4g_1(t_0)\frac{a_1}{c} \\
6\frac{a_1}{c}\left(g_2(t_0) + \frac{1}{2}\right) \\
8\left[\frac{a_1}{c}(g_3(t_0) + g_1(t_0)) - 1\right] \\
\vdots \\
2(m-4)\left[\frac{a_1}{c}(g_{m-5}(t_0) + g_{m-7}(t_0) + \dots + g_1(t_0)) - 1\right] \\
2(m-3)\frac{a_1}{c}\left[g_{m-4}(t_0) + g_{m-6}(t_0) + \dots + g_2(t_0) + \frac{1}{2}\right] \\
2(m-2)\frac{a_1}{c}(g_{m-3}(t_0) + g_{m-5}(t_0) + \dots + g_1(t_0)) \\
2(m-1)\frac{a_1}{c}\left(g_{m-2}(t_0) + g_{m-4}(t_0) + \dots + g_2(t_0) + \frac{1}{2}\right)
\end{array} \right],$$

wobei

$$\begin{aligned}
c &= a_1(g_1(t_0) + g_3(t_0) + \dots + g_{m-3}(t_0) + g_{m-1}(t_0)), \\
a_1 &= \frac{1}{4 \cdot 8 \cdot \dots \cdot 2((m-1)-1)}.
\end{aligned}$$

Im Falle eines ungeraden m hat die Inverse folgende Gestalt:

$$\tilde{E}^{-1} = \begin{array}{l} \left[\begin{array}{ccc} \frac{1}{2} \frac{a_2}{c} & & 0 \\ \frac{1}{2} \frac{a_2}{c} - 1 & & 0 \\ 2g_1(t_0) \frac{a_2}{c} & & -4 \\ 3 \left[\frac{a_2}{c} \left(g_2(t_0) + \frac{1}{2} \right) - 1 \right] & & 0 \\ 4 \frac{a_2}{c} (g_3(t_0) + g_1(t_0)) & & -8 \\ \vdots & & \vdots \\ (m-4) \left[\frac{a_2}{c} \left(g_{m-5}(t_0) + g_{m-7}(t_0) + \dots + g_2(t_0) + \frac{1}{2} \right) - 1 \right] & & 0 \\ (m-3) \frac{a_2}{c} [g_{m-4}(t_0) + g_{m-6}(t_0) + \dots + g_3(t_0) + g_1(t_0)] & & -2(m-3) \\ (m-2) \left[\frac{a_2}{c} \left(g_{m-3}(t_0) + g_{m-5}(t_0) + \dots + g_2(t_0) + \frac{1}{2} \right) - 1 \right] & & 0 \\ (m-1) \frac{a_2}{c} [g_{m-2}(t_0) + g_{m-4}(t_0) + \dots + g_3(t_0) + g_1(t_0)] & & -2(m-1) \end{array} \right. \\ \\ \left[\begin{array}{ccc} \frac{a_2}{c} & & 0 \quad \dots \\ \frac{a_2}{c} & & 0 \quad \dots \\ 4g_1(t_0) \frac{a_2}{c} & & 0 \quad \dots \\ 6 \left[\frac{a_2}{c} \left(g_2(t_0) + \frac{1}{2} \right) - 1 \right] & & 0 \quad \dots \\ 8 \frac{a_2}{c} (g_3(t_0) + g_1(t_0)) & & -8 \quad \dots \\ \vdots & & \vdots \quad \ddots \\ 2(m-4) \left[\frac{a_2}{c} \left(g_{m-5}(t_0) + g_{m-7}(t_0) + \dots + g_2(t_0) + \frac{1}{2} \right) - 1 \right] & & 0 \quad \dots \\ 2(m-3) \frac{a_2}{c} [g_{m-4}(t_0) + g_{m-6}(t_0) + \dots + g_3(t_0) + g_1(t_0)] & & -2(m-3) \quad \dots \\ 2(m-2) \left[\frac{a_2}{c} \left(g_{m-3}(t_0) + g_{m-5}(t_0) + \dots + g_2(t_0) + \frac{1}{2} \right) - 1 \right] & & 0 \quad \dots \\ 2(m-1) \frac{a_2}{c} [g_{m-2}(t_0) + g_{m-4}(t_0) + \dots + g_3(t_0) + g_1(t_0)] & & -2(m-1) \quad \dots \end{array} \right. \end{array}$$

$$\begin{array}{rcl}
\dots & 0 & \frac{a_2}{c} \\
\dots & 0 & \frac{a_2}{c} \\
\dots & 0 & 4g_1(t_0)\frac{a_2}{c} \\
\dots & 0 & 6\frac{a_2}{c}\left(g_2(t_0) + \frac{1}{2}\right) \\
\dots & 0 & 8\frac{a_2}{c}(g_3(t_0) + g_1(t_0)) \\
\dots & \vdots & \vdots \\
\dots & 0 & 2(m-4)\frac{a_2}{c}\left(g_{m-5}(t_0) + g_{m-7}(t_0) + \dots + g_2(t_0) + \frac{1}{2}\right) \\
\dots & -2(m-3) & 2(m-3)\frac{a_2}{c}[g_{m-4}(t_0) + g_{m-6}(t_0) + \dots + g_3(t_0) + g_1(t_0)] \\
\dots & 0 & 2(m-2)\left[\frac{a_2}{c}\left(g_{m-3}(t_0) + g_{m-5}(t_0) + \dots + g_2(t_0) + \frac{1}{2}\right) - 1\right] \\
\dots & -2(m-1) & 2(m-1)\frac{a_2}{c}[g_{m-2}(t_0) + g_{m-4}(t_0) + \dots + g_3(t_0) + g_1(t_0)] \\
\\
0 & & \frac{a_2}{c} \\
0 & & \frac{a_2}{c} \\
0 & & \frac{a_2}{c} \\
0 & & 4g_1(t_0)\frac{a_2}{c} \\
0 & & 6\frac{a_2}{c}\left(g_2(t_0) + \frac{1}{2}\right) \\
0 & & 8\frac{a_2}{c}(g_3(t_0) + g_1(t_0)) \\
\vdots & & \vdots \\
0 & & 2(m-4)\frac{a_2}{c}\left(g_{m-5}(t_0) + g_{m-7}(t_0) + \dots + g_2(t_0) + \frac{1}{2}\right) \\
0 & & 2(m-3)\frac{a_2}{c}[g_{m-4}(t_0) + g_{m-6}(t_0) + \dots + g_3(t_0) + g_1(t_0)] \\
0 & & 2(m-2)\frac{a_2}{c}\left(g_{m-3}(t_0) + g_{m-5}(t_0) + \dots + g_2(t_0) + \frac{1}{2}\right) \\
-2(m-1) & & 2(m-1)\frac{a_2}{c}[g_{m-2}(t_0) + g_{m-4}(t_0) + \dots + g_3(t_0) + g_1(t_0)]
\end{array} \Bigg| ,$$

wobei

$$c = a_2 \left(\frac{1}{2} + g_2(t_0) + g_4(t_0) + \dots + g_{m-3}(t_0) + g_{m-1}(t_0) \right),$$

$$a_2 = \frac{1}{6 \cdot 10 \cdot \dots \cdot 2(m-2)}.$$

Inverse zu \mathcal{E}

Für ein gerades m ist die Inverse von \mathcal{E} gegeben durch:

$$\mathcal{E}^{-1} = \begin{bmatrix} 0 & 2 & 0 & 2 & \dots & 0 & 2 & 0 & 2 \\ -4 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 6 & \dots & 0 & 6 & 0 & 6 \\ -8 & 0 & -8 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 10 & 0 & 10 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & 2(m-3) & 0 & 2(m-3) \\ -2(m-2) & 0 & -2(m-2) & 0 & \dots & -2(m-2) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 2(m-1) \\ -2m & 0 & -2m & 0 & \dots & -2m & 0 & -2m & 0 \end{bmatrix}.$$

C.6 Chebychev Polynome zweiter Art

Inverse zu \tilde{E}

Wird die reduzierte Ordnung m gerade gewählt, so hat die Inverse von \tilde{E} folgende Gestalt:

$$\tilde{E} = \begin{bmatrix} 0 & \frac{a_1}{c} & 0 \\ -2 & 2\frac{a_1}{c} & 0 \\ 0 & 4g_1(t_0)\frac{a_1}{c} - 4 & 0 \\ -2 & \frac{a_1}{c}(6g_2(t_0) + 2) & -6 \\ 0 & \frac{a_1}{c}(8g_3(t_0) + 4g_1(t_0)) - 4 & 0 \\ \vdots & \vdots & \vdots \\ 0 & \frac{a_1}{c}[2(m-4)g_{m-5}(t_0) + 2(m-6)g_{m-7}(t_0) + \dots + 4g_1(t_0)] - 4 & 0 \\ -2 & \frac{a_1}{c}[2(m-3)g_{m-4}(t_0) + 2(m-5)g_{m-6}(t_0) + \dots + 6g_2(t_0) + 2] & -6 \\ 0 & \frac{a_1}{c}[2(m-2)g_{m-3}(t_0) + 2(m-4)g_{m-5}(t_0) + \dots + 4g_1(t_0)] - 4 & 0 \\ -2 & \frac{a_1}{c}[2(m-1)g_{m-2}(t_0) + 2(m-3)g_{m-4}(t_0) + \dots + 6g_2(t_0) + 2] & -6 \end{bmatrix}.$$

$$\begin{array}{rcccc}
& \frac{a_3}{c} & & \cdots & 0 \\
& 2\frac{a_3}{c} & & \cdots & 0 \\
& 4g_1(t_0)\frac{a_3}{c} & & \cdots & 0 \\
& \frac{a_3}{c}(6g_2(t_0) + 2) & & \cdots & 0 \\
& \frac{a_3}{c}(8g_3(t_0) + 4g_1(t_0)) - 8 & & \cdots & 0 \\
& \vdots & & \ddots & \vdots \\
\frac{a_3}{c}[2(m-4)g_{m-5}(t_0) + 2(m-6)g_{m-7}(t_0) + \dots + 4g_1(t_0)] - 8 & & \cdots & & 0 \\
\frac{a_3}{c}[2(m-3)g_{m-4}(t_0) + 2(m-5)g_{m-6}(t_0) + \dots + 6g_2(t_0) + 2] & & \cdots & & -2(m-3) \\
\frac{a_3}{c}[2(m-2)g_{m-3}(t_0) + 2(m-4)g_{m-5}(t_0) + \dots + 4g_1(t_0)] - 8 & & \cdots & & 0 \\
\frac{a_3}{c}[2(m-1)g_{m-2}(t_0) + 2(m-3)g_{m-4}(t_0) + \dots + 6g_2(t_0) + 2] & & \cdots & & -2(m-3) \\
\\
& \frac{a_{m-3}}{c} & & & 0 \\
& 2\frac{a_{m-3}}{c} & & & 0 \\
& 4g_1(t_0)\frac{a_{m-3}}{c} & & & 0 \\
& \frac{a_{m-3}}{c}(6g_2(t_0) + 2) & & & 0 \\
& \frac{a_{m-3}}{c}(8g_3(t_0) + 4g_1(t_0)) & & & 0 \\
& \vdots & & & \vdots \\
\frac{a_{m-3}}{c}[2(m-4)g_{m-5}(t_0) + 2(m-6)g_{m-7}(t_0) + \dots + 4g_1(t_0)] & & & & 0 \\
\frac{a_{m-3}}{c}[2(m-3)g_{m-4}(t_0) + 2(m-5)g_{m-6}(t_0) + \dots + 6g_2(t_0) + 2] & & & & 0 \\
\frac{a_{m-3}}{c}[2(m-2)g_{m-3}(t_0) + 2(m-4)g_{m-5}(t_0) + \dots + 4g_1(t_0)] - 2(m-2) & & & & 0 \\
\frac{a_{m-3}}{c}[2(m-1)g_{m-2}(t_0) + 2(m-3)g_{m-4}(t_0) + \dots + 6g_2(t_0) + 2] & & & & -2(m-1)
\end{array}$$

$$\begin{array}{c}
\frac{a_{m-1}}{c} \\
2 \frac{a_{m-1}}{c} \\
4g_1(t_0) \frac{a_{m-1}}{c} \\
\frac{a_{m-1}}{c} (6g_2(t_0) + 2) \\
\frac{a_{m-1}}{c} (8g_3(t_0) + 4g_1(t_0)) \\
\vdots \\
\frac{a_{m-1}}{c} [2(m-4)g_{m-5}(t_0) + 2(m-6)g_{m-7}(t_0) + \dots + 4g_1(t_0)] \\
\frac{a_{m-1}}{c} [2(m-3)g_{m-4}(t_0) + 2(m-5)g_{m-6}(t_0) + \dots + 6g_2(t_0) + 2] \\
\frac{a_{m-1}}{c} [2(m-2)g_{m-3}(t_0) + 2(m-4)g_{m-5}(t_0) + \dots + 4g_1(t_0)] \\
\frac{a_{m-1}}{c} [2(m-1)g_{m-2}(t_0) + 2(m-3)g_{m-4}(t_0) + \dots + 6g_2(t_0) + 2]
\end{array}$$

wobei

$$\begin{aligned}
c &= a_1 g_1(t_0) + a_3 g_3(t_0) + \dots + a_{m-3} g_{m-3}(t_0) + a_{m-1} g_{m-1}(t_0), \\
a_1 &= \frac{1}{8 \cdot 12 \cdot \dots \cdot 2m}, \\
a_3 &= \frac{1}{4 \cdot 12 \cdot 16 \cdot \dots \cdot 2m}, \\
&\vdots \\
a_{m-3} &= \frac{1}{4 \cdot 8 \cdot \dots \cdot 2(m-4) \cdot 2m}, \\
a_{m-1} &= \frac{1}{4 \cdot 8 \cdot \dots \cdot 2m}.
\end{aligned}$$

Ist andernfalls m ungerade, so ergibt sich als Inverse von \tilde{E} :

$$\tilde{E}^{-1} = \begin{array}{r} \left[\begin{array}{ccc} 2\frac{a_0}{c} & & 0 \\ 2\left(2\frac{a_0}{c} - 1\right) & & 0 \\ 8g_1(t_0)\frac{a_0}{c} & & -4 \\ 2\left(\frac{a_0}{c}(6g_2(t_0) + 2) - 1\right) & & 0 \\ 2\frac{a_0}{c}(8g_3(t_0) + 4g_1(t_0)) & & -4 \\ \vdots & & \vdots \\ 2\left[\frac{a_0}{c}(2(m-2)g_{m-5}(t_0) + 2(m-6)g_{m-7}(t_0) + \dots + 6g_2(t_0) + 2) - 1\right] & & 0 \\ 2\frac{a_0}{c}[2(m-3)g_{m-4}(t_0) + 2(m-5)g_{m-6}(t_0) + \dots + 8g_3(t_0) + 4g_1(t_0)] & & -4 \\ 2\left[\frac{a_0}{c}(2(m-2)g_{m-3}(t_0) + 2(m-4)g_{m-5}(t_0) + \dots + 6g_2(t_0) + 2) - 1\right] & & 0 \\ 2\frac{a_0}{c}[2(m-1)g_{m-2}(t_0) + 2(m-3)g_{m-4}(t_0) + \dots + 8g_3(t_0) + 4g_1(t_0)] & & -4 \end{array} \right. \\ \\ \left[\begin{array}{ccc} \frac{a_2}{c} & & 0 \dots \\ 2\frac{a_2}{c} & & 0 \dots \\ 4g_1(t_0)\frac{a_2}{c} & & 0 \dots \\ \frac{a_2}{c}(6g_2(t_0) + 2) - 6 & & 0 \dots \\ \frac{a_2}{c}(8g_3(t_0) + 4g_1(t_0)) & & -8 \dots \\ \vdots & & \vdots \ddots \\ \frac{a_2}{c}[2(m-4)g_{m-5}(t_0) + 2(m-6)g_{m-7}(t_0) + \dots + 6g_2(t_0) + 2] - 6 & & 0 \dots \\ \frac{a_2}{c}[2(m-3)g_{m-4}(t_0) + 2(m-5)g_{m-6}(t_0) + \dots + 8g_3(t_0) + 4g_1(t_0)] & & -8 \dots \\ \frac{a_2}{c}[2(m-2)g_{m-3}(t_0) + 2(m-4)g_{m-5}(t_0) + \dots + 6g_2(t_0) + 2] - 6 & & 0 \dots \\ \frac{a_2}{c}[2(m-1)g_{m-2}(t_0) + 2(m-3)g_{m-4}(t_0) + \dots + 8g_3(t_0) + 4g_1(t_0)] & & -8 \dots \end{array} \right. \end{array}$$

$$\begin{array}{r}
\dots \quad 0 \quad \frac{a_{m-3}}{c} \\
\dots \quad 0 \quad 2 \frac{a_{m-3}}{c} \\
\dots \quad 0 \quad 4g_1(t_0) \frac{a_{m-3}}{c} \\
\dots \quad 0 \quad \frac{a_{m-3}}{c} (6g_2(t_0) + 2) \\
\dots \quad 0 \quad \frac{a_{m-3}}{c} (8g_3(t_0) + 4g_1(t_0)) \\
\dots \quad \vdots \quad \vdots \\
\dots \quad 0 \quad \frac{a_{m-3}}{c} [2(m-4)g_{m-5}(t_0) + 2(m-6)g_{m-7}(t_0) + \dots + 6g_2(t_0) + 2] \\
\dots \quad -2(m-3) \quad \frac{a_{m-3}}{c} [2(m-3)g_{m-4}(t_0) + 2(m-5)g_{m-6}(t_0) + \dots + 8g_3(t_0) + 4g_1(t_0)] \\
\dots \quad 0 \quad \frac{a_{m-3}}{c} [2(m-2)g_{m-3}(t_0) + 2(m-4)g_{m-5}(t_0) + \dots + 6g_2(t_0) + 2] - 2(m-2) \\
\dots \quad -2(m-3) \quad \frac{a_{m-3}}{c} [2(m-1)g_{m-2}(t_0) + 2(m-3)g_{m-4}(t_0) + \dots + 8g_3(t_0) + 4g_1(t_0)] \\
\\
0 \quad \frac{a_{m-1}}{c} \\
0 \quad \frac{a_{m-1}}{c} \\
0 \quad 4g_1(t_0) \frac{a_{m-1}}{c} \\
0 \quad \frac{a_{m-1}}{c} (6g_2(t_0) + 2) \\
0 \quad \frac{a_{m-1}}{c} (8g_3(t_0) + 4g_1(t_0)) \\
\vdots \quad \vdots \\
0 \quad \frac{a_{m-1}}{c} [2(m-4)g_{m-5}(t_0) + 2(m-6)g_{m-7}(t_0) + \dots + 6g_2(t_0) + 2] \\
0 \quad \frac{a_{m-1}}{c} [2(m-3)g_{m-4}(t_0) + 2(m-5)g_{m-6}(t_0) + \dots + 8g_3(t_0) + 4g_1(t_0)] \\
0 \quad \frac{a_{m-1}}{c} [2(m-2)g_{m-3}(t_0) + 2(m-4)g_{m-5}(t_0) + \dots + 6g_2(t_0) + 2] \\
-2(m-1) \quad \frac{a_{m-1}}{c} [2(m-1)g_{m-2}(t_0) + 2(m-3)g_{m-4}(t_0) + \dots + 8g_3(t_0) + 4g_1(t_0)]
\end{array} \Bigg] ,$$

wobei

$$\begin{aligned}
c &= 2a_0 + a_2g_2(t_0) + a_4g_4(t_0) + \dots + a_{m-3}g_{m-3}(t_0) + a_{m-1}g_{m-1}(t_0), \\
a_0 &= \frac{1}{6 \cdot 10 \cdot \dots \cdot 2m}, \\
a_2 &= \frac{1}{10 \cdot 14 \cdot \dots \cdot 2m},
\end{aligned}$$

$$\begin{aligned}
 a_4 &= \frac{1}{6 \cdot 14 \cdot 18 \cdot \dots \cdot 2m}, \\
 &\quad \vdots \\
 a_{m-3} &= \frac{1}{6 \cdot 10 \cdot \dots \cdot 2(m-4) \cdot 2m}, \\
 a_{m-1} &= \frac{1}{6 \cdot 10 \cdot \dots \cdot 2(m-2)}.
 \end{aligned}$$

Inverse zu \mathcal{E}

Ist m gerade, so ist die Inverse zu \mathcal{E}

$$\mathcal{E}^{-1} = \begin{bmatrix}
 0 & 6 & 0 & 10 & \dots & 0 & 2(m-1) & 0 & 2(m+1) \\
 -4 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 10 & \dots & 0 & 2(m-1) & 0 & 2(m+1) \\
 -4 & 0 & -8 & 0 & \dots & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & \dots & 0 & 2(m-1) & 0 & 2(m+1) \\
 \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
 0 & 0 & 0 & 0 & \dots & 0 & 2(m-1) & 0 & 2(m+1) \\
 -4 & 0 & -8 & 0 & \dots & -2(m-2) & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 2(m+1) \\
 -4 & 0 & -8 & 0 & \dots & -2(m-2) & 0 & -2m & 0
 \end{bmatrix}.$$

D Weitere Abbildungen zu den verallgemeinerten Eigenwerten

D.1 Verallgemeinerte Eigenwerte für die äquivalente Sylvester-Gleichung

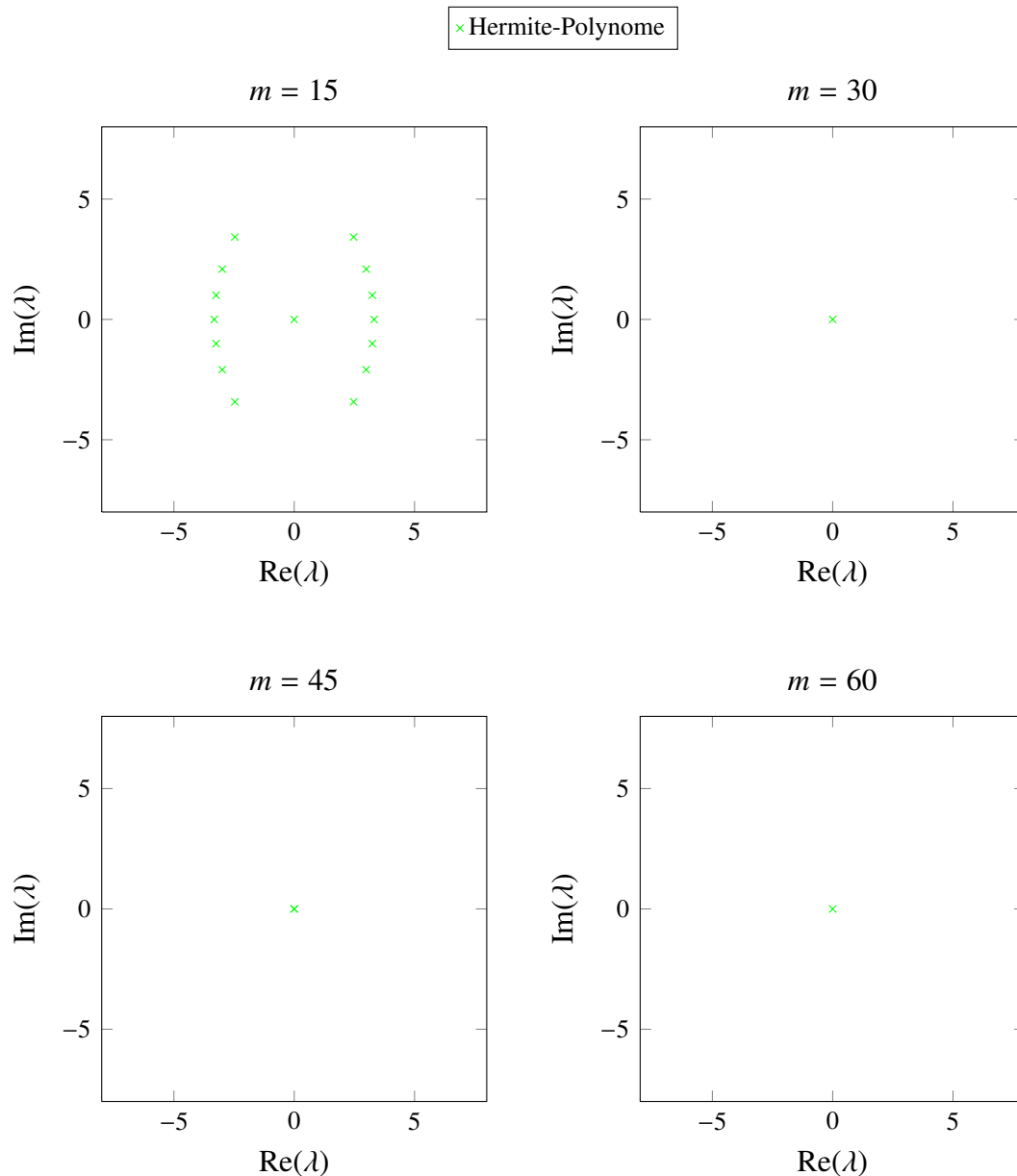


Abbildung 17: Endliche Eigenwerte des Matrixpaares $(-\tilde{A}, \tilde{E})$ für Hermite-Polynome für $m = 15, 30, 45, 60$

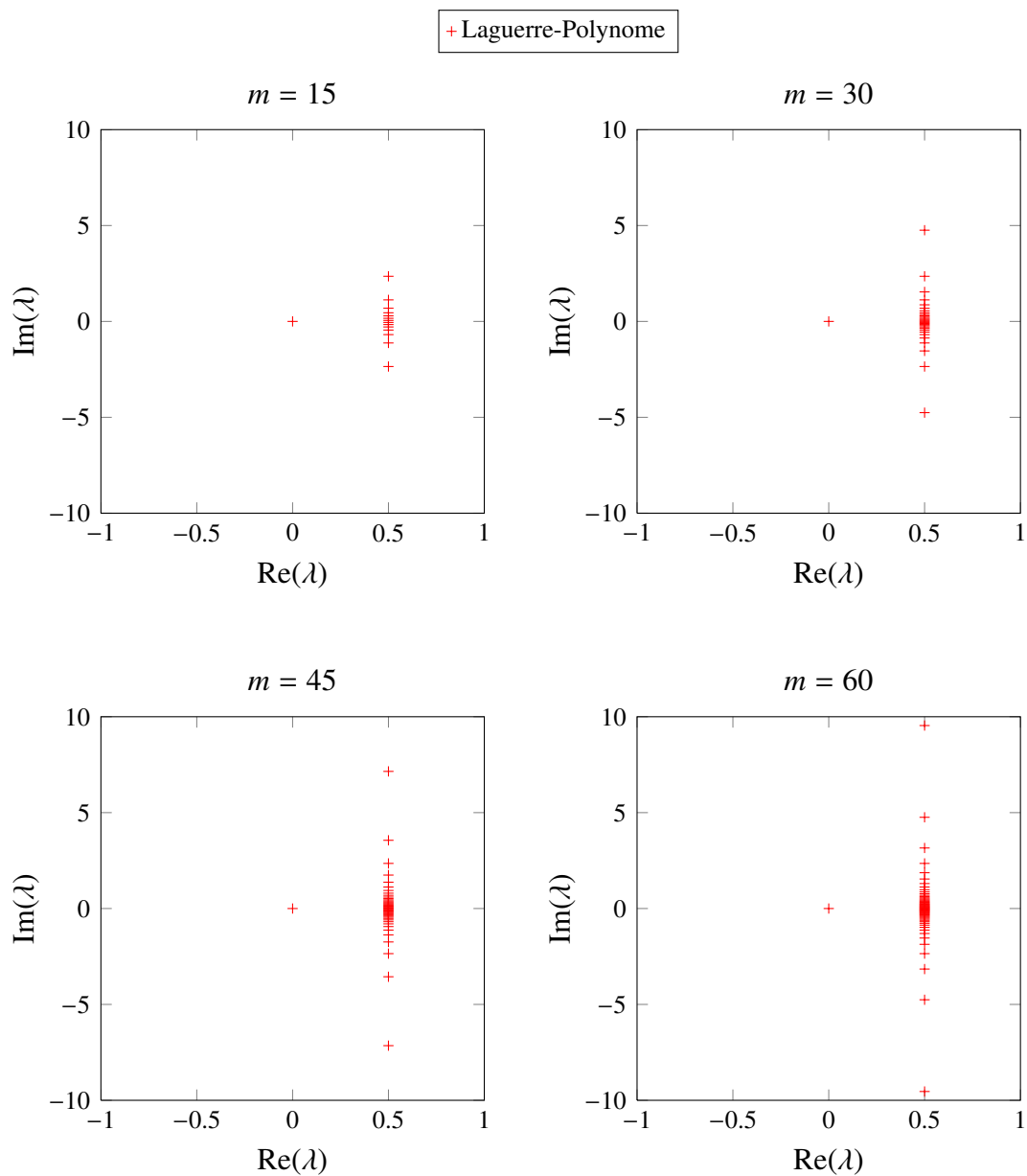


Abbildung 18: Endliche Eigenwerte des Matrixpaares $(-\tilde{A}, \tilde{E})$ für Laguerre-Polynome für $m = 15, 30, 45, 60$

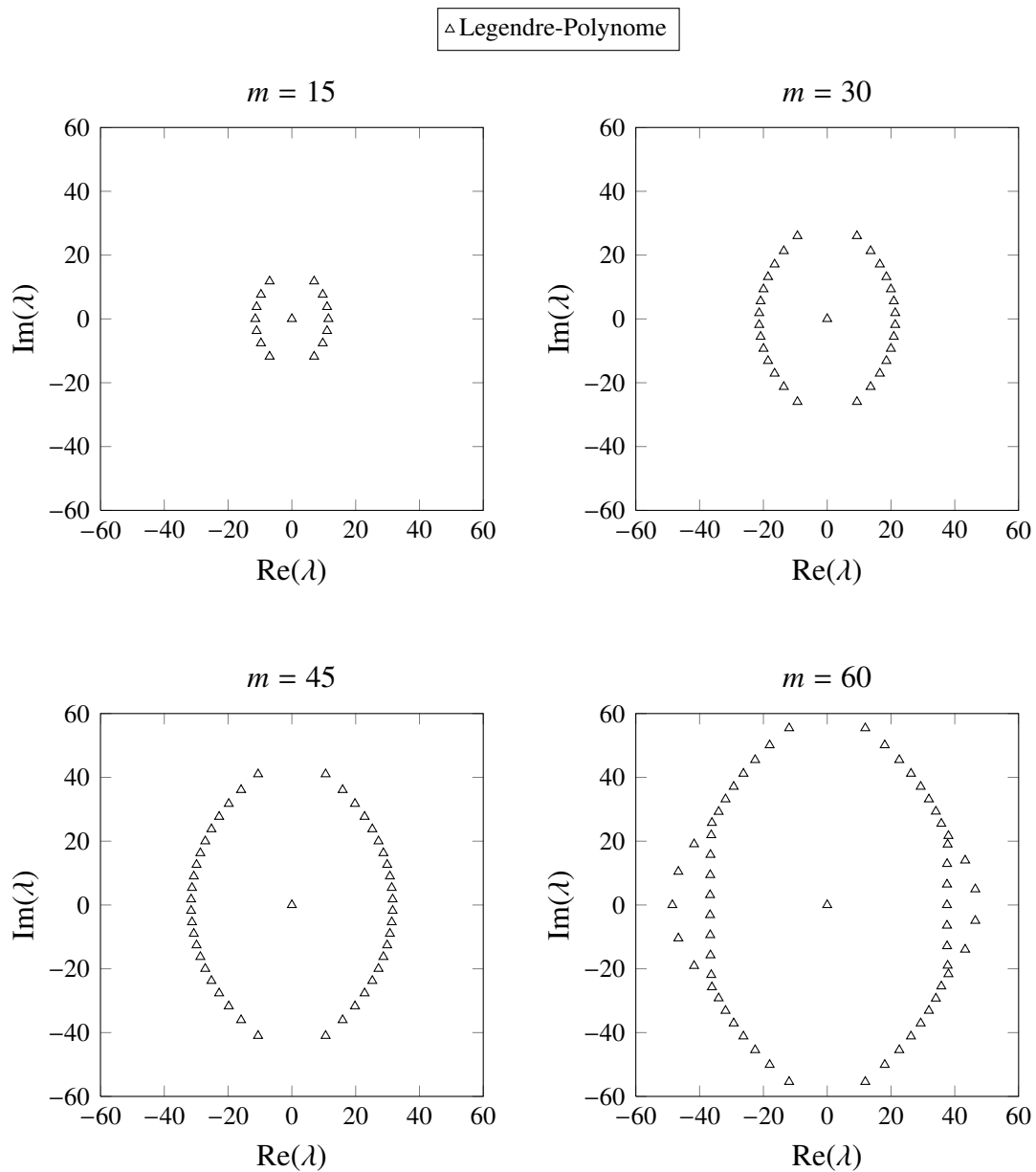


Abbildung 19: Endliche Eigenwerte des Matrixpaares $(-\tilde{A}, \tilde{E})$ für Legendre-Polynome für $m = 15, 30, 45, 60$

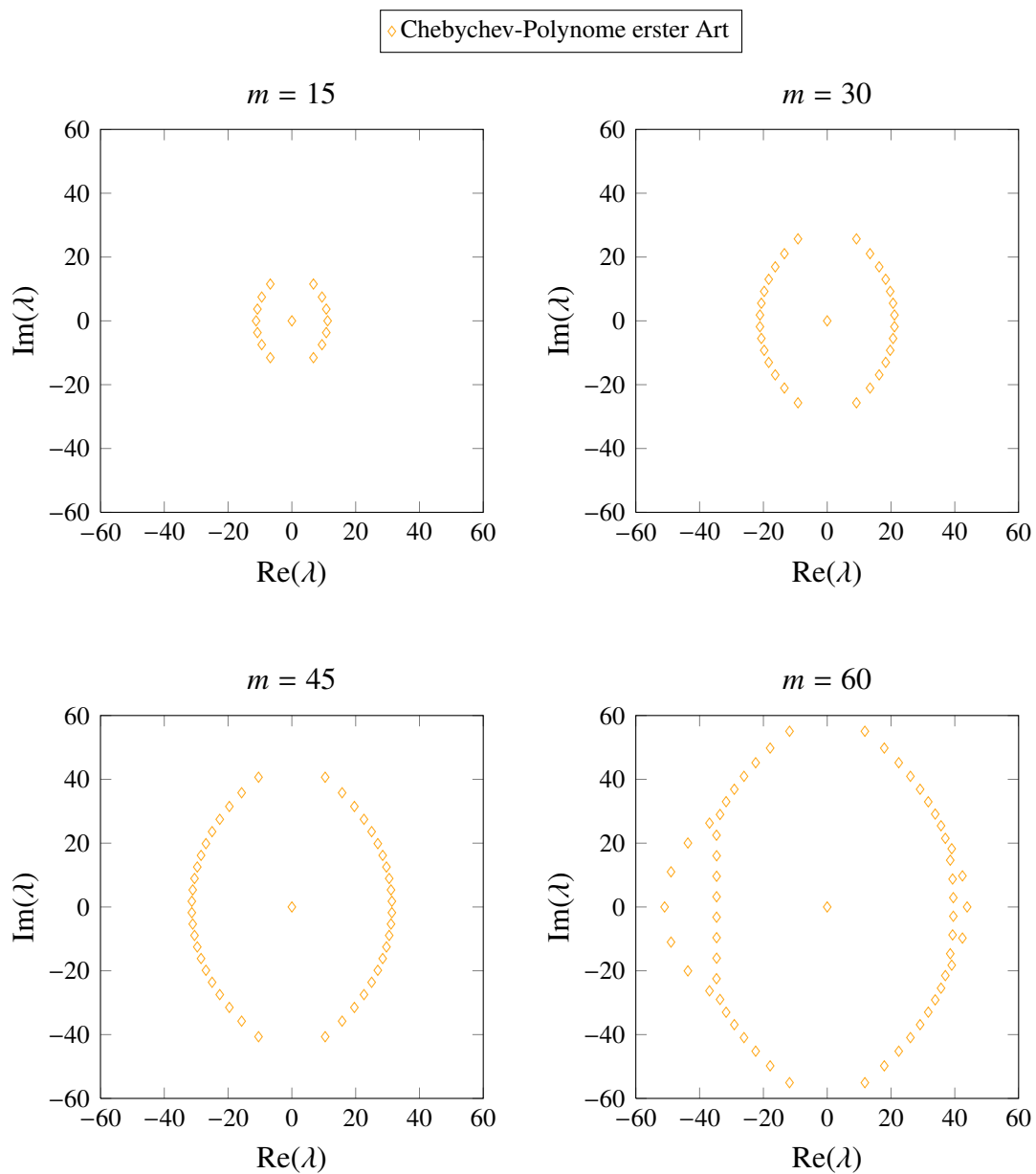


Abbildung 20: Endliche Eigenwerte des Matrixpaares $(-\tilde{A}, \tilde{E})$ für Chebychev-Polynome erster Art für $m = 15, 30, 45, 60$

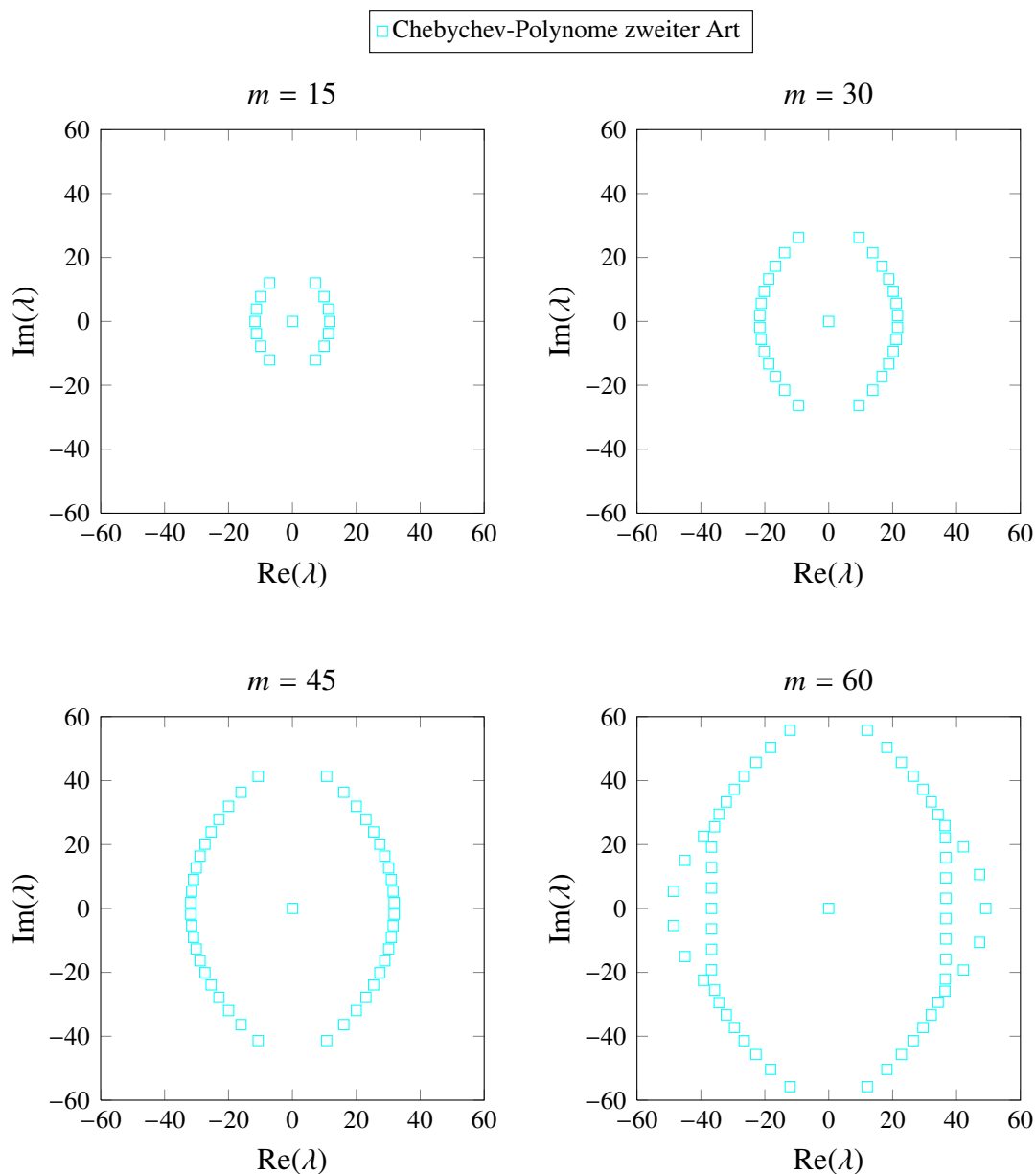


Abbildung 21: Endliche Eigenwerte des Matrixpaares $(-\tilde{A}, \tilde{E})$ für Chebychev-Polynome zweiter Art für $m = 15, 30, 45, 60$

D.2 Verallgemeinerte Eigenwerte für die zusatzfreie Variante 2

Da die \mathcal{E} -Matrix in keinem Fall invertierbar ist, sind alle verallgemeinerten Eigenwerte ∞ . Daher wird keine Grafik zu diesen Entwicklungspunkten dargestellt.

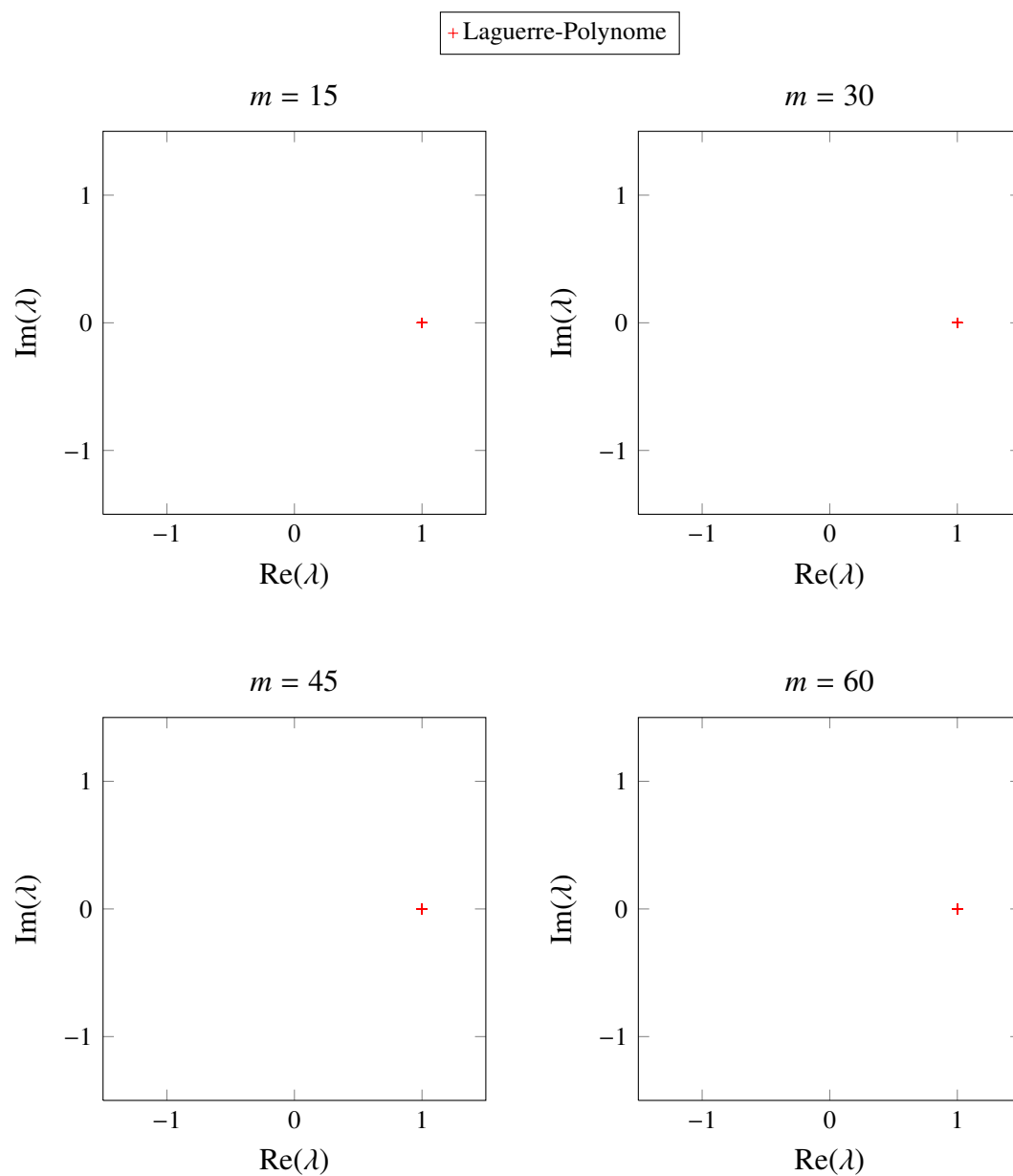


Abbildung 22: Endliche Eigenwerte des Matrixpaares $(-I_m, \mathcal{E})$ für Laguerre-Polynome für $m = 15, 30, 45, 60$

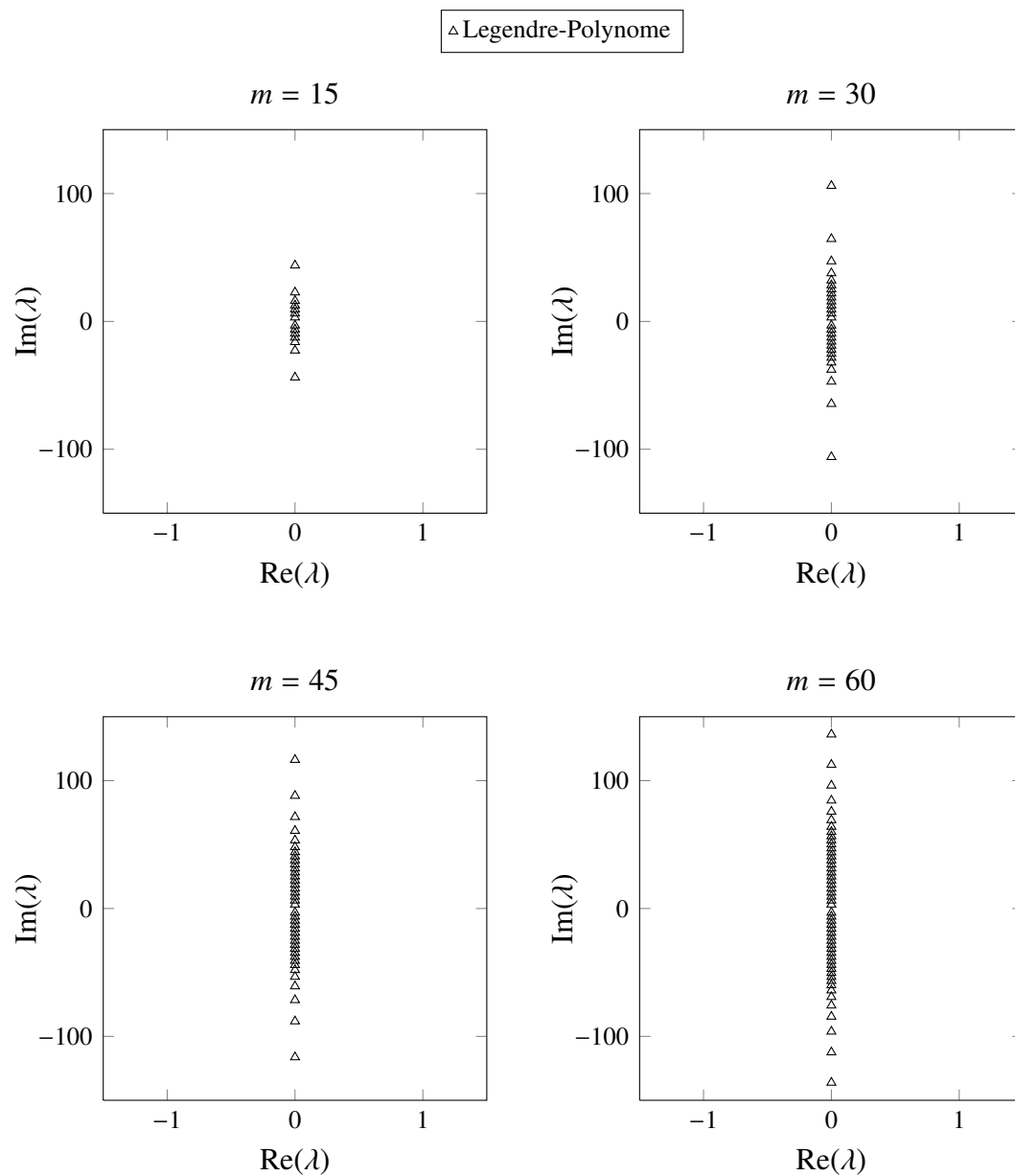


Abbildung 23: Endliche Eigenwerte des Matrixpaares $(-I_m, \mathcal{E})$ für Legendre-Polynome für $m = 15, 30, 45, 60$

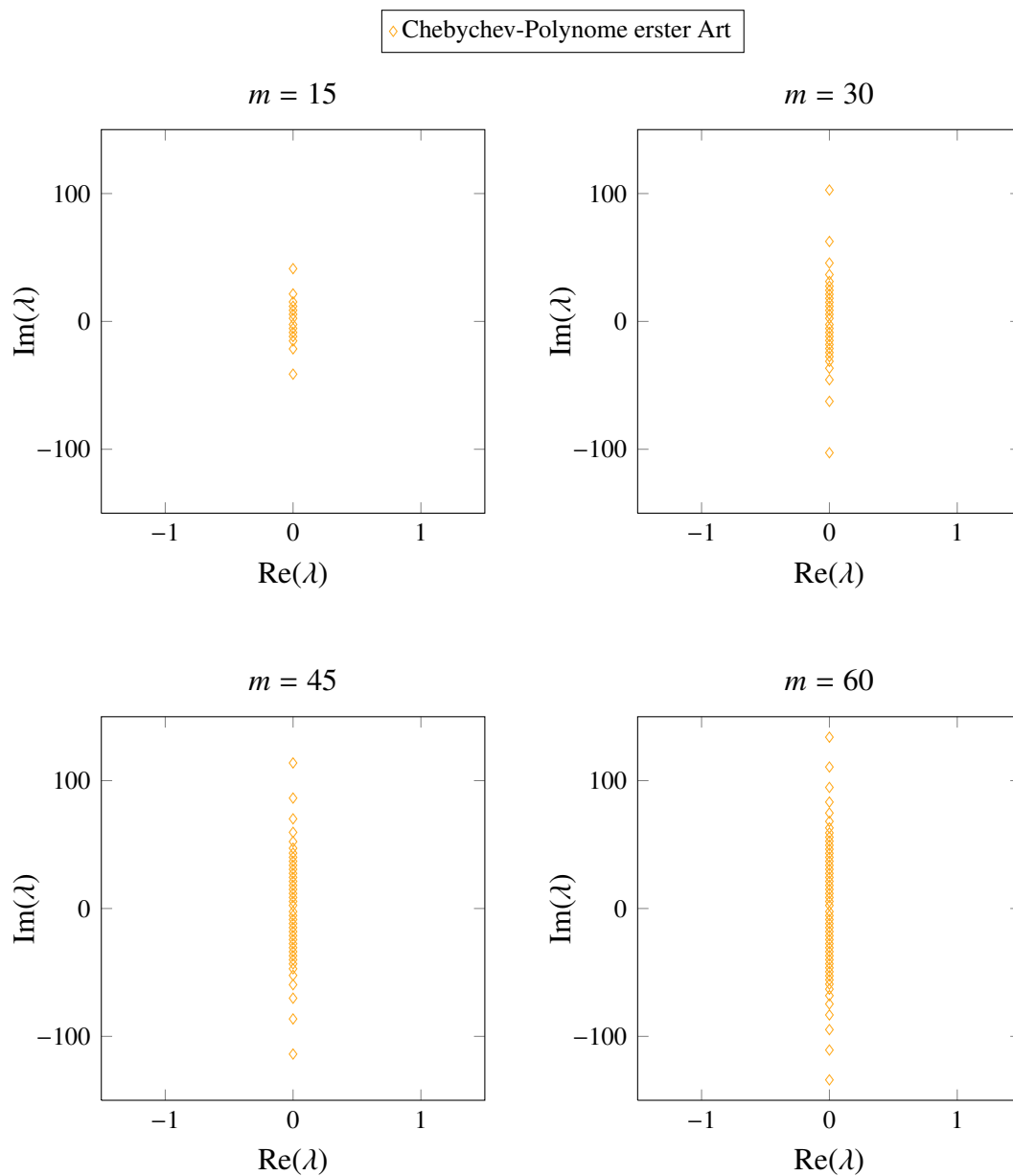


Abbildung 24: Endliche Eigenwerte des Matrixpaares für $(-I_m, \mathcal{E})$ Chebychev-Polynome erster Art für $m = 15, 30, 45, 60$

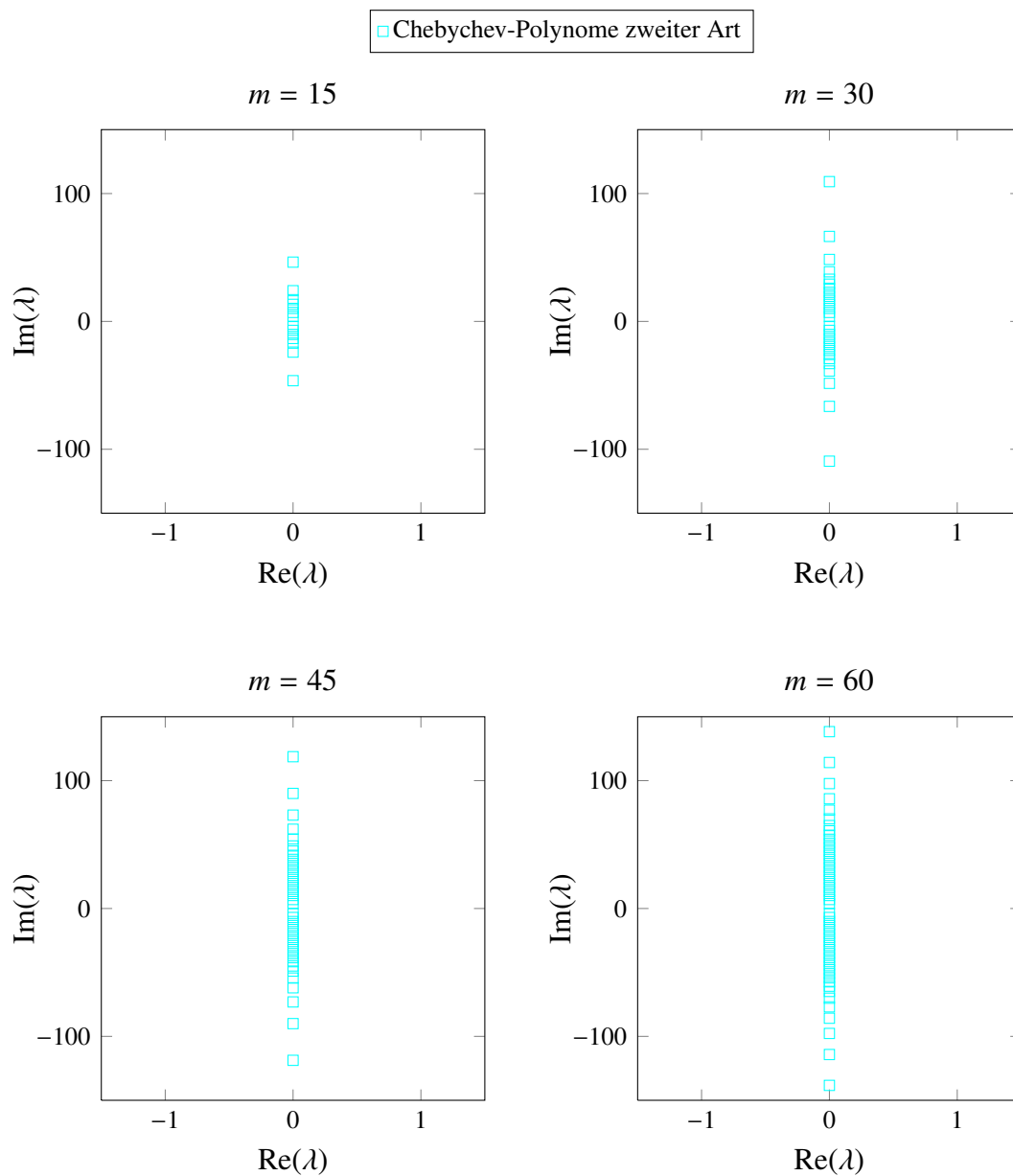


Abbildung 25: Endliche Eigenwerte des Matrixpaares $(-I_m, \mathcal{E})$ für Chebychev-Polynome zweiter Art für $m = 15, 30, 45, 60$

E CD

Auf dem Datenträger sind folgende Ordner zu finden:

- **Dokumente**
Dieser Ordner enthält zum einen die Masterarbeit einschließlich des Anhangs (Masterarbeit.pdf). Zum anderen ist ein Dokument Abbildungen.pdf enthalten, welches die Grafiken, die im Hauptteil der Arbeit stark verkleinert wurden, vergrößert und zusammen mit denen des Anhangs darstellt.
- **Matlab-Programme**
Dieser Ordner ist in weitere Unterordner strukturiert, in denen sich die verwendeten Matlab-Implementierungen, Testbeispiele und Ergebnisse befinden. Eine Readme.pdf erläutert die einzelnen Unterordner, sowie die darin enthaltenen .m - und .mat-Dateien und erklärt die Entstehung der Ergebnisdaten.

Fehlerkorrekturen zur Masterarbeit „Zeitbereichs- Modellreduktion und Sylvester-Gleichungen“

Änderung 20.04.2016, Seite 70

Für ein ungerades m lässt sich die Sylvester-Gleichung (4.4) wie folgt umformulieren:

$$\begin{aligned} & AV\tilde{E} + EV\tilde{A} = \tilde{R} \\ \Leftrightarrow & AV - EV \cdot (-\tilde{A}\tilde{E}^{-1}) = \begin{bmatrix} Ax_0 & Bw_1 & \dots & Bw_{m-1} \end{bmatrix} \tilde{E}^{-1} \\ \Leftrightarrow & AV - EV \cdot (-\tilde{A}\tilde{E}^{-1}) = B \begin{bmatrix} z_0 & w_1 & \dots & w_{m-1} \end{bmatrix} \tilde{E}^{-1} \\ \Leftrightarrow & AV - EVS = BL, \end{aligned}$$

wobei $z_0 = \begin{bmatrix} 0 & \dots & 0 \end{bmatrix}^T \in \mathbb{R}^p$ der Nullvektor ist.